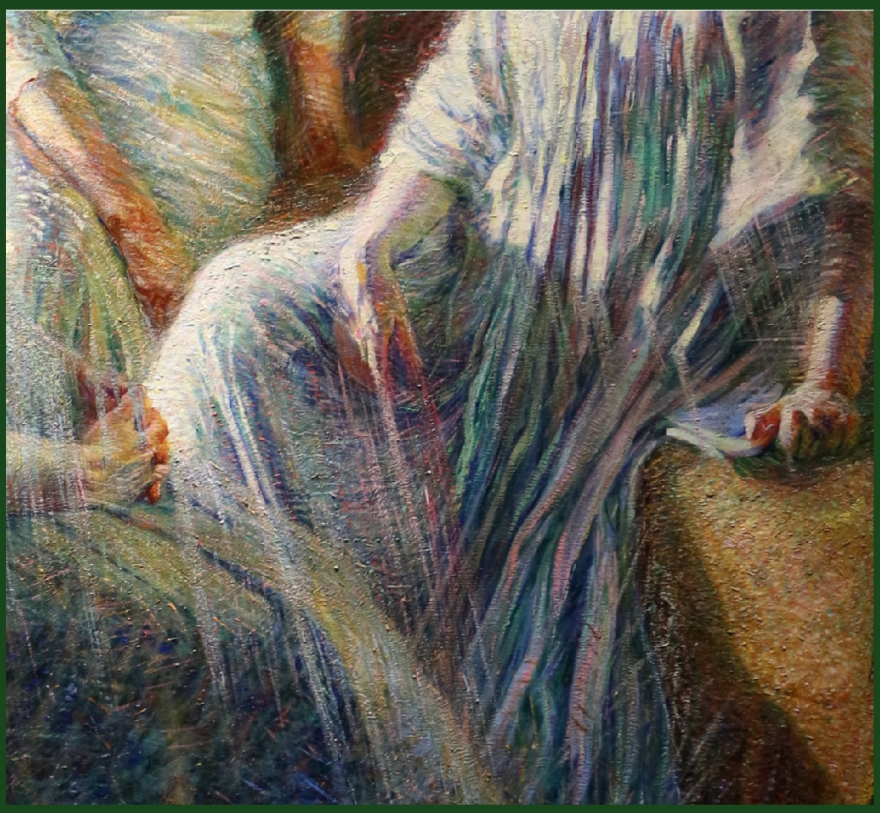
The Italian Law Journal

Vol. 07 No. 01 (2021)



EDIZIONI SCIENTIFICHE ITALIANE

www.theitalianlawjournal.it

THE ITALIAN LAW JOURNAL

An International Forum for the Critique of Italian Law

Vol. 07 – No. 01 (2021)

www.theitalianlawjournal.it

ISSN 2421-2156 June 2021 DOI 10.23815/2421-2156.ITALJ

Table of Contents

The Italian Law Journal
Vol. 07 – No. 01 (2021)

History and Projects

Essays

Hard Cases

Malebolge

Corporate and Financial Markets Law

Short Symposium: 'The Terrible Duty. Legitimacy and Usefulness of Punishment'

The Italian Law Journal is sponsored by 'Harvard Italian Law Association' (HILA), by 'Società Italiana degli Studiosi del diritto civile' (SISDIC) and by 'Società Italiana per la ricerca nel diritto comparato' (SIRD).

The Chinese Civil Code and 'Fascination' with Roman Law. A Conversation with Oliviero Diliberto

Camilla Crea and Oliviero Diliberto*

Abstract

The Civil Code of the People's Republic of China came into force on 1 January 2021 following a long and complex gestation lasting decades and involving many failed attempts at different times in Chinese history. The main focus of this short interview is an assessment of the possible significance that the legacy of Roman law (and its Italian scholarship) may have had on civil law codification in China.

I. The Possible Impact of Roman Law on Civil Law Codification in China

Law is both culture and politics, and, as such, is never without bias in its processes and what it produces. This would appear to be a fundamental underlying concept that invites us to reflect on diverse experiences of law, together with their interconnections and transitions. The Civil Code of the People's Republic of China (中华人民共和国民法典)[1] came into force on 1 January 2021 following a long and complex gestation lasting decades and involving many failed attempts at different times in Chinese history.[2]

Geopolitical and geo-economic interest in China goes hand in hand with a marked curiosity regarding historical and comparative discourse. There are many reasons for this: firstly, China's civil code shows the influence of various foreign legal models and their contamination through contact with local traditions, which led in turn to differentiation, adaptation, and/or 'domestication'.[3] Secondly, the

* Camilla Crea is Associate Professor of Private Law at the University of Sannio. Oliviero Diliberto is Full Professor of Roman Law at Sapienza University of Rome. Translation, questions and footnotes by C. Crea; answers by O. Diliberto.

[1] Civil Code of the People's Republic of China, adopted on 28 May 2020, effective from 1 January 2021 (https://tinyurl.com/p9sy4shh; for the English translation cf https://tinyurl.com/3ennvpd7 (last visited 30 June 2021)). See for the Italian translation of this code, O. Diliberto et al eds, *Codice civile della Repubblica Popolare Cinese*, trad. by M. Huang, intr. D. Xu (Pisa: Pacini Editore, 2021).

[2] R. Zimmermann, 'Codification: History and Present Significance of an Idea' *European Review of Private Law*, 95-120 (1995); P. Grossi, 'Codici: qualche conclusione fra un millennio e l'altro', in P. Cappellini and B. Sordi eds, *Codici. Una riflessione di fine millennio* (Milano: Giuffrè, 2002).

[3] M. Timoteo, 'China Codifies. The first book of the civil code between Western models to Chinese characteristics' *Opinio Juris in Comparatione*, 24-44 (2019).

choice to adopt a modern form of 'codification' appears to be quite significant as it clearly draws upon Western legal traditions as well as cultural and ideological phenomena from the past, far removed from those of today. Indeed, the system of Roman Law (directly, and indirectly, through the undoubted influence of the dominant German model)[4] is also known to have had some bearing on this codification.

A useful insight comes from an old book review that appeared in the Yale Law Journal in 1920. Its anonymous author recognized the key role played by continental Europe's legal systems ('Western jurisprudence')[5] in the early draft of the general principles of Chinese civil law, admitting – from the common law standpoint – that the translation into English and publication in a volume of Chinese Supreme Court decisions constituted a major step towards the civilization of China in comparison to other countries. This is why his or her reasoning concluded with a somewhat provocative aspiration:

> Apparently, the Chinese mind as a result of long centuries of civilization and philosophic study has acquired a nimbleness which enables its judges to apply with mastery the rules of the new *jus gentium*. May we not hope, however, that the legal structure to be erected will not be based exclusively upon the principles of continental law, but that it will appropriate also the good qualities of the Anglo-American legal system? May China be far-sighted enough to send more of her youth to study law in England and the United States, so that they may become acquainted with the spirit of Anglo-American law. If our young sister republic should succeed in blending the two great legal systems of the world – the Roman-continental and the Anglo-American – it would make a contribution to civilization, the effect of which can hardly be over.[6]

Some decades later (close to the proclamation of Mao's People's Republic of China, which occurred on 1 October 1949) Roscoe Pound, the great and renowned American scholar and thinker, among many other religious missionaries, went to China on his appointment as legal advisor. His expectation, shared by the larger American legal community, was to help to inspire, transform – and perhaps even

[4] J. Xue and A. Somma, 'La codificazione del diritto civile nel terzo millennio. Riflessioni storiche e politico-normative' *Materiali per una storia della cultura giuridica*, 329-343 (2004) (where J. Xue, interviewing A. Somma, underscored the strong interest in the German model. The Italian legal system was also of interest as it was a hybrid between the pandectistic tradition and the French civil code). See P.G. Monateri and J. Xue, 'Dialogo sulla codificazione del diritto civile in Cina' *Rivista critica di diritto privato*, 469-499 (2003).

[5] E.G.L., 'The Private Law of China' 30(2) *Yale Law Journal*, 180-184 (1920), (reviewing the English translation of 'The Chinese Supreme Court decisions: first instalment translation relating to general principles of civil law and to commercial law', translated by F.T. Cheng (and republished: Nabu Press, Charleston SC, United States, 2010).

[6] E.G.L., n 5 above, 184.

fix – Chinese society and its legal system along American lines.[7] By the end of his journey, however, he was forced to admit that China was well equipped with excellent codes inspired by Roman Law. Indeed, to a country whose culture was grounded principally on custom and morals, the systematic nature of this legal model appeared more appropriate and suitable than the Anglo-American one, thus proving particularly apt for its transition to a modern legal system.[8]

These insights explain the main focus of this short interview examining the possible significance that the legacy of Roman law (and its Italian scholarship) might have had on civil law codification in China, looking beyond the latest code, which has recently come into force.

*What are the historical reasons for China looking to the Roman law tradition? Would you say that the so-called philosophy of Roman law (*natura, ratio *and* aequitas *– to quote the exact words of Yang Zhenshan),*[9] *if such a thing really exists, may have played a role in this?*

Answer:

I would start from a general premise: Roman law, as the expression of a 'state reality' ('state' in the broadest sense, since the concept of State would emerge much later), ceased to exist, on the one hand, with the fall of the Western Roman Empire and, on the other, with the collapse of the Byzantine Empire in 1453. It might be useful to draw a parallel: although the Latin language gradually died out (albeit never completely: suffice it to recall the official – and, in its own way, 'universal' – language of the Vatican), many Romance languages sprang from it in Europe and later in Latin America through the Spanish and Portuguese conquistadors. These languages (eg French, Italian, Spanish, Romanian, Ladino, etc) are obviously different from each other, but they share a common syntactic and grammatical structure and many almost identical lexical items: their common origin in Latin represents a shared basis of communication, without prejudice to the evolution of each individual language and their differences.

As for Roman law, a somewhat similar process took place, but it was one of

[7] J. Kroncke, 'Roscoe Pound in China: A Lost Precedent for the Liabilities of American Legal Exceptionalism' 38 *Brooklyn Journal of International Law*, 77-143, 81 (2012). The story of Roscoe Pound symbolizes attempts to Americanize Chinese law and clarifies the role of Sino-American relations in the formation of modern American legal internationalism (for further reflections, Id, *The Futility of Law and Development: China and the Dangers of Exporting American Law* (New York: Oxford University Press, 2015)). See, also Z. Wang, 'The Roman Law Tradition and Its Future Development in China' 1 *Frontier of Law in China*, 72-78 (2006) (pointing out the Chinese preference for continental Roman law over common law in the early 20th century).

[8] R. Pound, 'Roman Law in China' *L'Europa e il diritto romano. Studi in memoria di P. Koschaker* (Giuffrè: Milano, 1954), 441.

[9] Z. Yang, 'La tradizione filosofica del diritto romano e del diritto cinese antico e l'influenza del diritto romano sul diritto cinese contemporaneo' 69(4) *Rivista internazionale di filosofia del diritto*, 582-599 (1992), (now in L. Formichella et al eds) *Diritto Cinese e sistema giuridico romanistico. Contributi* (Giappichelli: Torino, 2004), 29-43.

infinitely wider latitude: with the demise of the 'state' experience of Roman law, 'neo-Roman' legal systems were created all over Europe (these were initially the so-called Romano-Barbarian legal systems, adopted by the populations that took over the Western Roman Empire). The migrants of the time were fully aware of the cultural superiority of the empire they were conquering, immediately learning Latin, converting to Christianity, and assimilating Roman law, which they combined with their own customs and traditions. From that initial melting pot, a formidable phenomenon arose and grew: the *ius commune*, to all intents and purposes a neo-Roman law initially encompassing Central and Western Europe within its spectrum. Then, with the fall of Constantinople, the Orthodox patriarchate moved to Moscow – which, by no coincidence, became the third Rome – bringing with it Roman law, which thus acquired a territorial breadth infinitely greater than that of the Romance languages. Again, Roman law from old Europe reached Latin America and, to some extent, also one of the states in the United States, Louisiana, as well as Quebec in Canada. These are mixed jurisdictions that have adopted a code historically influenced by the French model.

In the late 19th century, Japan began its phase of 'modernization' and decided to adopt its own civil law legislation, directly inspired by the German legal system, reflecting the greatness of the Pandectics, the contemporary European doctrinal model par excellence. Through Japanese contamination, Roman law also reached China at the beginning of the twentieth century.[10]

A question thus arises: what is the common 'grammar' of the various neo-Roman laws? First of all, there is a shared exegetical technique that originated in Roman jurisprudence, ie the interpretation of legal texts, which is the same in all legal orders based on the Roman system. Secondly, the private law 'system' is also shared by the Roman one. If we think about it, this is one of the great paradoxes of history, albeit a fascinating one: classical Roman law, in fact, actually had no system (or almost none), being of a casuistic nature. However, in drafting the *Corpus Iuris*, Justinian, for the first time (apart from a few previous attempts) created a model from which the subsequent codices would stem, a systemic work. The 'technique' and the 'system' are therefore the same everywhere. An example may help to simplify and clarify: in 1930s Europe, Roman private law was applied in Stalin's Russia, Mussolini's Italy and in the France of Léon Blum and bourgeois representative democracy. The Roman matrix, its systematic layout, is the same in these very different countries, although, naturally, the content of each private law institution varies according to the political-ideological, economic, and social contexts of the various States. The epiphanies of property law are emblematic of this phenomenon.

The fundamental landmarks in this story are two epoch-making occurrences: firstly, the creation of Justinian's *Corpus Iuris Civilis*, which spread right across

[10] S. Schipani, 'Diritto romano in Cina' *Enciclopedia Treccani* (2009), available at https://tinyurl.com/99h8ne5m (last visited 30 June 2021).

Europe from the year one thousand (it should be recalled that the institution known as 'university' came into being at the *Studium* in Bologna, the first school of law, founded with the precise aim of promoting the study of Justinian's Roman law). Another fundamental step was the *Code Napoléon* of 1804. From a reading of Portalis's '*Discours préliminaire au premier projet de Code civil*',[11] and tracing the history of the code's development, which followed the French tradition while being imbibed with the new post-revolutionary individualist spirit, we may observe that the French jurists themselves constantly claimed to take their inspiration from Roman law. This is solemnly declared, starting with the right to property, and emphasized in Art 544 of the Civil Code as

> the right to enjoy and to dispose of things in the most absolute manner, provided that one does not make a use of them that is prohibited by laws (*lois*) or regulations (*règlements*).

This concept, originating with the bourgeois Enlightenment, does not exist as such in Roman law, which does not recognize the absolute right to property.[12] As I mentioned, I will return to this theme, which is obviously a central one in our reflection, later.

These two stages, the *Justinian Corpus Iuris* and the *Code Napoléon*, are in communication with each other. The *Corpus Iuris* transforms Roman case law into a 'code', and it is from this model of 'code' that the French protagonists of 1804 would draw direct (though largely misinterpreted) inspiration. Some authors have claimed that Roman law was so successful because of its philosophy: the *naturalis ratio*, the *ars boni et aequi* etc. The objective reality is that the system of Roman law has taken root in environments with very different ideologies and in equally diverse socio-economic environments (eg, monarchies, local seigniories, feudalism, the dictatorship of the proletariat, democratic republics): this means that the ideological component is of no consequence, otherwise it could not have worked.

The Roman law system was chosen because of its utility – its rationality – which, however, should not to be understood as *ratio naturalis*, invoked to affirm

[11] J.E.-M. Portalis, *Discours préliminaire au premier projet de Code civil (1801)* (original title: Motifs et discours prononcés lors de la publication du Code civil. Discours prononcé le 21 janvier 1801 et le Code civil promulgué le 21 mars 1804, with an introduction by M. Massenet (Bordeaux: Éditions Confluences, 2004), available at https://tinyurl.com/2h457k8m (last visited 30 June 2021). According to J. Gordley, 'Myths of the French Civil Code' 42(3) *The American Journal of Comparative Law*, 459-505, 489 (1994): 'For Portalis, law was founded on human nature, reflected in the laws of all civilized peoples but particularly those of the Romans, and discovered through the efforts of jurists and scholars over the centuries. 'Law (*droit*) is universal reason', he explained, 'supreme reason founded on the very nature of things. Enacted laws (*lois*) are or ought to be only the law (*droit*) reduced to positive rules, to particular precepts'. This higher law was reflected in those 'valuable collections for the science of laws' made by the Roman jurists'.

[12] O. Diliberto, 'L'eredità fraintesa. Il diritto di proprietà dall'esperienza romana al Code Napoléon (e viceversa)' *Rassegna di diritto civile*, 374-382 (2020).

that 'man is at the centre of law' as some have claimed. This statement fails to consider that over half of the human beings in ancient Rome were slaves and, therefore, a matter for law; women, moreover, could become *sui iuris* and enjoy legal capacity, but they had a reduced capacity to act. Essentially, I'd say the workings of Roman law, compared with modern codifications – had nothing to do with an alleged 'value system' of its own, but with its intrinsic, ductile and – so to speak – 'meta-temporal' nature.

II. Periodizations and the History of Codification: Academic and Institutional Dialogues with Italy

The history of Chinese codification is remarkably complex and strongly influenced by the political, social, and economic scenarios of each different period. It is recognized that the first draft of the civil code dates back to 1911, under the great *Qing* dynasty (*Da qing minlu cao'can*).[13] This project was never adopted, but a revised version based on the 'civil' parts extracted from the Qing code remained in use until the promulgation of the civil code drawn up by the Kuomintang (*Guomindang*) government in 1929-1930 during the Republican period. The German model, and the pandectistic school notably inspired both of these attempts to such an extent that the Chinese system started to be considered to belong to the civil law family, or to wear 'the civil law dress', and it was intellectualized within the framework of Roman law.[14]

Although this code is still in force in Taiwan,[15] it never took effect in mainland China, being formally abolished due to its incompatibility with the new spirit of the People's Republic of China (PRC), proclaimed in 1949, and the victory of the Chinese Communist Party (CCP) in the Chinese Civil War.

Professor Sandro Schipani proposed a possible division into periods: the pre-socialist period; the phase in the early 50s after the foundation of the People's Republic of China, shaped by the guiding role of Moscow and the inspiring

[13] P.C.C. Huang, *Code, Custom, and Legal Practice in China. The Qing and the Republic Compared* (Stanford: Stanford University Press, 2001) (comparing the Qing period with the Republican one); P.R. Luney, 'Traditions and Foreign Influences: Systems of Law in China and Japan' 52 *Law & Contemporary Problems*, 131 (1989) (pointing out the key role of this early codification in the history of China and subsequent attempts to draft a civil code); J. Zhang, 'On the Qing Civil Law (Qingdai minfa zonglun)' *Chinese University of Political. Science and Law Press*, 1998; L. Chen, 'The Historical Development of the Civil Law Tradition in China: A Private Law Perspective'78 *The Legal History Review*, 159, 161 (2010).

[14] Following the classification of R. David and J.E.C. Brierley, *Major Legal Systems in the World Today: An Introduction to the Comparative Study of Law* (London: Stevens, 2nd ed, 1978), 23-24. See J. Xue, 'Il diritto romano in Cina' 12 *Cardozo Electronic Law Bulletin*, 1-6 (2006).

[15] For further references see L. Zhang, 'Latest Development of Codification of Chinese Civil Law' 83 *Tulane Law Review*, 1000-1001 (2009). This codification was inspired by the BGB (Bürgerliches Gesetzbuch) but also by the Swiss and French legal systems: T.-F. Chen, 'Transplant of Civil Code in Japan, Taiwan, and China: With the Focus of Legal Evolution' *National Taiwan University Law Review*, 400 (2011).

example of the Soviet Union; the change in legal policy of 1978, with strong links to the socialist market economy and socialism with Chinese characteristics.[16] Here, legal discourse made its way back onto the political agenda after the 'nihilism' of the Soviet period.

Since 1978, with the opening of China to the world and under the political leadership of Deng Xiaoping, significant new steps have been taken in developing civil law: a first attempt, however, resulted simply in the enactment of the General Principles of Civil Law (GPCL) in 1986.

The law-making process started again in earnest in March 1998, when Wang Hanbin, Vice Chairman of the National People's Congress (NPC), created 'a Group for the Redaction of the Civil Code', comprising scholars from all over the world.

The aim shared by the members was to conduct feasibility studies for a future civil law code and to express their opinions on its possible contents and structure. At the very beginning, the commission focused on some preliminary questions that were incorporated into a questionnaire, analyzing the main Western legal models and evaluating their compatibility with the Chinese experience and tradition. It would seem no coincidence that one of the key questions of the questionnaire (which prof. Xue Jun distributed among Italian scholars) concerned 'the problem of assessing the pandectistic system and its modernity'.[17]

This additional phase, once again, did not lead to significant results, as the attempts at codification were intertwined, over the years, with a 'piecemeal approach'[18] where many special laws were passed in the different fields of private law (eg contract law, property and civil liability respectively, rights *in rem*, marriage, etc).

Actually, it was only after 2014 and the presidency of Xi Jinping that the idea of a domestic code was fully and effectively embedded in the Chinese political agenda, in line with the 'theory of rule of law with Chinese characteristics',[19] looking at foreign legal systems while comparing and experiencing them through a 'learning by doing' approach'.[20]

On 15 March 2017, the Fifth Session of the 12th National People's Congress passed the General Provisions of Civil Law, which represent an important step in Chinese civil law codification. They were incorporated in the first part of the

[16] Cf S. Schipani, 'Fondamenti romanistici e diritto cinese (riflessioni su un comune lavoro nell'accrescimento del sistema)' *Bullettino dell'Istituto di diritto romano*, XVI, (2016), (in line with the thinking of P. Jiang, 'Il diritto romano nella Repubblica Popolare cinese' 16 *Index*, 367 et seq (1988), (and in L. Formichella et al eds, n 8 above, 3)).

[17] J. Xue and A. Somma, n 4 above, 329.

[18] M. Timoteo, n 3 above, 28; L. Chen, 'Introduction', in Lei Chen and C.H. (Remco) van Rhee eds, *Towards a Chinese Civil Code: Comparative and Historical Perspectives* (Martinus Nijhoff Publishers: Leiden-Boston, 2012).

[19] H. Liang, 'The Reception of Foreign Civil Law in China' 1 *Shandong University Law Review*, 5 (2003).

[20] L. Wang, 'The Modernization of Chinese Civil Law over Four Decades' 14(1) *Frontiers of Law in China*, 39-72, 40-41 (2019).

civil code that was adopted on 28 May 2020.[21]

Within the framework of a contemporary Chinese legal system – complicated by the circulation of multiple foreign legal models and their contamination with the domestic reality and culture – one element seems to recur:

> In the long-lasting project of the codification of Chinese private law, the German model still plays a leading role as far as the structure of the code and the core component of its conceptual framework are concerned.[22]

And the German model – as has repeatedly been observed – is, in turn, profoundly influenced by Roman law.

The different periods in the codification process seem closely tied in with a different role of the inspiration by Roman law or its 'reception'.[23] *Within this framework, how did the dialogue with Italian scholars come about and how did it evolve?*

Answer:

At the turn of the 20th century, law was an alien concept in China, at least as it is understood in the West: the rules of civil coexistence were based on Confucianism and local customs and traditions, even though some provisions of rather elementary criminal law did exist. When China therefore decided to adopt a system of private law, and to start a process of modernization, it was inevitable that it would look to the most widespread system in the world (Roman law in the 'meta-temporal' sense), so that it might relate to it as Japan had done shortly before it. This country had a millenary culture, totally different from that of the West. As it opened to the world and abandoned feudalism (the *Shōgun*), Japan drew from the most solid and prestigious Western tradition, namely from the German model. Essentially, at that point in history, the common law was mainly applied in the United Kingdom, a colonial empire. This meant that it is was not particularly attractive to these Eastern countries. India had a common law system, but as a colony. This explains why other Eastern countries did not adopt the Anglo-Saxon system, preferring the continental one. Obviously,

[21] H. Jiang, 'The Making of a Civil Code in China: Promises and Perils of a New Civil Law' 96 *Tulane Law Review,* 777-819 (2021).

[22] L. Zhang, n 18 above, 1039: 'Actually, China's civil law is also a mixed jurisdiction, not only because of the great diffusion of the studies on the common law in China, but also because of the very special and important role of judicial interpretation in current legal practice. Today's Chinese civil law is based on Roman law and *Pandectenrecht*. However, by incorporating the common law experiences in the drafting of its civil code, Chinese legal scholars and the legislature are trying to exceed them and build a new and modern codification model in the world, mixed with the common law experience'.

[23] R. Li, 'The Reception of Roman Law', in Z. Yang and S. Schipani eds, *Roman Law, China Law and the Codification of Civil Law* (Beijing: Chinese University of Political Science and Law Press, 1995), 71.

we are talking about a period – the turn of the twentieth century – when the Chinese, for the first time, posed themselves the problem of constructing a civil code. Numerous commissions were set up but repeatedly failed. The last of these commissions – set up in the 1930s – succeeded in drawing up a code that, however, would never be applied in mainland China as war had broken out in Manchuria in the meantime; the area was occupied by the Japanese, and a very long period of aggression was to follow. Of course, the 1930s code would never be applied in China. We have a very well-researched work by Lara Colangelo,[24] a young scholar of Chinese language and law, who has produced a convincing chronology of all these attempts and the various commissions. Following the defeat of the nationalists, this code was therefore taken to Taiwan, where it is still in force today.

In any case, I'm not fully convinced by the periodization proposed by Jiang Ping. The Soviet period (from 1949 to the end of the 1950s) certainly existed, as Jiang Ping is aware, having graduated in law in Moscow and now being regarded as a sort of doyen among Chinese legal scholars. Jiang Ping, however, omits to mention – perhaps due to an understandable oversight – that after the Chinese Communist Party's break with the USSR, the Soviet period ended, and the season of 'legal nihilism' began. This phase coincided with the Cultural Revolution, when the Soviet model was abandoned, yet there was no consideration on law as such, as it was considered a bourgeois superstructure. Of course, after Mao's death, Deng Xiaoping came back on the scene. He seized power, and the era of the four modernizations began. It started in 1978 but took some time to become established.

The turning point in the development of the law actually occurred in 1988, when Ping Jiang came to Rome, invited by Professor Sandro Schipani, the eponymous hero of the construction of a cultural network and exchanges between Italian and Chinese academics. Schipani is credited with having foreseen a reality that, in the Italy of that time, was considered a sort of 'intellectual oddity', but which was, in reality, an absolute truth: China's opening to the market meant that it would soon need rules and to develop a civil law system. After Ping Jiang's visit to Italy, an initial cooperation agreement was drawn up, first with the University of Tor Vergata (in partnership with Beijing's CUPL), to begin translating Roman legal texts into Chinese, which would allow a direct approach to Roman law works. Thus, translations of some volumes on specific areas (ie the law of obligations, rights *in rem*, succession, the family) were published, with a selection of texts from the Digest. After this, the Institutes of Gaius and the Institutes of Justinian were translated into Chinese in their entirety. Then, in

[24] L. Colangelo, 'La traduzione delle fonti del diritto romano e la formazione di un linguaggio giuridico cinese: possibili interferenze grammaticali dal latino' *Rivista degli studi orientali. Nuova Serie*, 285-312 (2015); Id, 'La ricezione del sistema giuridico romanistico e la relativa produzione di testi in Cina all'inizio del xx secolo: le fonti del diritto romano in due dei primi manuali in lingua cinese' *Bullettino dell'Istituto di diritto romano*, 195 (2016).

the 1990s, a process to disseminate knowledge of Roman law began, one of dialogue and comparison in order to assess which legal model might best suit the Chinese context.

In all this, there is also a 'case' that concerns me and that helped the 'long march'[25] of Roman law in China to progress. In 1998, I became Minister of Justice in Italy; the Chinese were engaged in reflections on what the best model for their civil law might be, and I, in addition to being a politician and a Communist minister, was also a professor of Roman law. So, in 1999, in the company of Sandro Schipani and the Attorney General of the Italian Supreme Court, I went to China. The affair also assumed an 'institutional' – and no longer a solely academic – dimensions. In fact, the Chinese Minister of Justice was also present at that meeting. From then on, Party leaders, and later the People's Assembly, began to legislate according to the Roman private law model, mediated through a number of contemporary experiences of codification, starting with the BGB.

III. The 'Current' Sense of Roman Tradition

The Chinese Civil Code, approved on 28 May 2020, consists of 1260 articles and 7 books: the general provisions, property, contracts, personality, family law, succession and tort. A separate book deals with personality rights. The code stands as 'a milestone for both the protection of human rights and the promotion of rule of law in China'.[26] Compared to other codifications of the past, this code of the second millennium seeks to prioritize the human person and his or her dignity, promote core socialist values, and respond to the needs of the modern era (such as the digital revolution and ecological change), with a view to settling the practical problems arising from the Chinese context.

It is the result of legal transplants of foreign models and multiple contaminations with Chinese features and culture, which still create tensions for adaptation locally.[27] Among these models there is no mention of Roman law.

What remains, if anything, of this tradition/legacy of Roman law in the new code? Can we speak of current relevance, or is it simply an image of historical fascination?

Answer:

Rather than speaking of current relevance, we ought to discuss (use, employ, refer to) a somewhat stronger term: being in force. Roman law, as an expression of

[25] M. Timoteo, 'La lunga marcia della codificazione civile nella Cina contemporanea' *Bullettino dell'Istituto di diritto romano*, 35 (2016).

[26] Z. Huo, 'China Enters an Era with a Civil Code' *China Justice Observer* (May 29, 2020), available at https://tinyurl.com/2nw3s4ay (last visited 30 June 2021).

[27] H. Jiang, n 21 above, 777-919.

statehood (the Roman empire), is a definitively concluded experience. On the contrary, the 'meta-historical' or meta-temporal Roman law and its systemic structure is easily adaptable to other state realities. The Chinese relate to the German legal system, which is based on Roman law. On some matters, again because it is useful and practical, the Chinese have also drawn inspiration from some common law experiences and from the *lex mercatoria*, given their fundamental role in international trade.

In the Chinese collective imagination, their civil code was inspired by the Roman legal system. When President Xi Jinping came to Italy on a state visit to seal the New Silk Road agreement (the 'Belt and Road Initiative' (BRI)), he wrote an article that appeared in *Corriere della Sera*.[28] The Chinese President described the friendship between Italy and China as a phenomenon rooted in a prestigious historical legacy. There were two empires in the world, the Chinese in the East and the Roman in the West, and Italy is the heir of the latter. Italy still enjoys, undeservedly perhaps, the long wave of the Roman empire, which the Chinese recognize as being the only other one on the same level as the Chinese empire. We are faced with the recognition of a legal heritage, of which Italy is an expression.

There is more: another key element to consider is the progressive rediscovery of Western classical culture, beyond legal works, by the Chinese world. About fifteen years ago, some leaders of the Chinese Communist Party asked about Demosthenes, the well-known Athenian politician and orator. The question was justified by their interest in ancient Western rhetoric to train Chinese managers, given that learning the art of persuasion and argumentation is considered a fundamental skill, much more important than knowing how to do mathematics. Furthermore, the translation of works such as the Divine Comedy[29] shows an interest in the Western cultural tradition in the broadest sense, which necessarily

[28] 'La visita di Xi Jinping «Un patto strategico con l'Italia»' *Corriere della Sera*, 20 March 2019, available at https://tinyurl.com/ptbvzd6w (last visited 30 June 2021): China and Italy are respectively emblems of Eastern and Western civilization and have written some of the most important and significant chapters in the history of human civilization. Italy is the home of ancient Roman civilization and the cradle of the Renaissance, and its heritage of great monuments, artistic and literary masterpieces is now widely known in China. The contacts between the two great civilizations, the Chinese and Italian, have their roots in history. Already more than two thousand years ago, the ancient Silk Road connected ancient China and ancient Rome, despite the great distances that separated them. The Han dynasty sent Gan Ying on a mission in search of what they called '*Da Qin*' or 'Great Qin' which referred precisely to the Roman empire, while the writings of the poet Virgil and the Roman geographer Pomponius Mela contain multiple references to the 'Silk Country'. Later, Marco Polo's '*Milione*' triggered the first 'passion for China' in Western history and its author became a pioneer of contacts between Eastern and Western cultures, a model that still inspires ambassadors of friendship today (Authors' translation from Italian).

[29] On the various translations (among which, the one by Tian Dewang 田德望 (1997) stands out, as it would be the first complete translation from the original text) and their shortcomings with respect to an increasingly sophisticated and demanding Chinese public, cf K.P. Laurence, 'Translating the Divina Commedia for the Chinese Reading Public in the Twenty-First Century' 21(2) *Wong TTR: traduction, terminologie, rédaction*, 191-220(2008).

includes law.

The great difference in the Chinese civil codification process compared to other non-European civil codes, and unlike the Italian code, is that it lacks the political and cultural mediation of the *Code Napoléon*, although knowledge of French scholarship is still evident. All the codifications with a basis in Roman law have transposed it through its age-old tradition throughout the Middle Ages and the Modern Age. Contemporary civil codes, in essence, all have Roman foundations, but this results from a bourgeois Enlightenment mediation, which does not derive directly from Roman law as such but from the interpretation and use that the drafters of the French Civil Code of 1804 made of it.

An example may illustrate this phenomenon. The right to property is an absolute legal right in the Italian legal order (and in most others) to the point of being called 'the selfish right'. It builds on the work that the drafters of the first code of the modern age, namely the Napoleonic code, carried out on Roman legal sources. Indeed, in the first code of the united Italy, dated 1865, the definition of private property was literally and slavishly translated into Italian from the French code.

But in Roman law as such, the absoluteness of the right to private property is an unknown category. It was the bourgeois revolution to reconsider the Roman legal sources and draw concepts aiming to uphold the absolute and inviolable character of private property. Roman law, being flexible and adaptable, provided the Napoleonic codification with the framework, the system, but the contents – as already pointed out above – were determined by the lawmaker of the time.

By contrast, the Chinese codification has 'skipped' – so to speak – Napoleonic mediation, directly engaging in the appropriation and re-elaboration of the Roman system. The rules on property clearly demonstrate this process. Private property is not framed as a cornerstone (a sort of static engine) in the system of rights; no reference is made to its absoluteness (much less to its inviolability); no question is raised about its unity: in fact, different forms of property coexist on the same foot in Chinese statutes. Of course, the right to property is the foremost of real rights, but devoid of any sacredness or inviolability or absoluteness, as we are used to reading in contemporary civil codes.

A further aspect should be considered. In the Western experience, the codes were born before the Constitutions. Constitutions are a twentieth century phenomenon and present extraordinarily advanced concepts in terms of social rights. The Italian civil code dates back to 1942 and the Constitution to 1948. A few years passed between these two texts, yet they seem to belong to two different universes. It is not by chance that Italian private law scholars thought for some decades that the Constitution was a mere political-ideological manifesto, and it is only since the mid-1960s that scholars and courts have explored the relationship between the code and the Constitution. In so doing, they launched the season of constitutionalization of private law and promoted the interpretation of institutions

in the light of Constitutional principles.[30]

China reversed this process. The modern code was born after the Constitution of 1982, then amended several times in the following decades (1988, 1993, 1999, 2004 and 2018), thereby gradually and formally extending the rule of law and the protection of human rights.[31] The code appeared after the Constitution and stands as an instrument to recognize new rights, in implementation of the rule of law, while still within a socialist legal system. Art 1 of the Chinese Civil Code, not surprisingly, states

> This Law is formulated in accordance with the Constitution of the People's Republic of China for the purposes of protecting the lawful rights and interests of the persons of the civil law, regulating civil-law relations, maintaining social and economic order, meeting the needs for developing socialism with Chinese characteristics, and carrying forward the core socialist values.

IV. The Code of the New Millenium?[32]

Initial attempts at Chinese codification performed a 'defensive' function[33] vis-à-vis the local system, to react to the imposition of the law and jurisdiction of the courts as envisaged by international treaties. There was a tendency to 'imitate' foreign models and, in particular, Western legal traditions. Over time, the need to build a solid legal system, in line with domestic traditions and capable of 'contributing', with its own specificity, to transnational legal discourse, has emerged.

On a global level, will the new Chinese civil code represent a strategy of 'resistance' and identity, or will it actively seek to spread its paradigm within the international arena?

Answer:

The code should not be read in a (or at least not only) strictly political key,

[30] E. Navarretta, 'Diritto civile e diritto costituzionale' *Rivista di diritto civile*, 643 (2012); F. Macario, 'Autonomia privata (profili costituzionali)' *Enciclopedia del diritto. Annali* (Milano: Giuffrè, 2015), VII, 61. For a comprehensive study see the monumental work in five volumes, by P. Perlingieri, *Il diritto civile nella legalità costituzionale secondo il sistema italo-europeo delle fonti* (Napoli: Edizioni Scientifiche Italiane, 4th ed, 2020) (first published in 1983).

[31] Q. Zhang, 'A Constitution without constitutionalism? The paths of constitutional development in China' 8(4) *International Journal of Constitutional Law*, 950-976 (2011) (for more historical details, Id, *The Constitution of China: A Contextual Analysis* (Oxford: Hart Publishing, 2012)).

[32] J. Gordley and H. Jiang, *Part I: Will the Chinese Civil Code Become the Code of the Century?*, 16 November 2020, available at https://tinyurl.com/2m56uh69 (last visited 30 June 2021).

[33] S. Schipani, n 10 above.

as it essentially responds to a need for simplification and a *reductio ad unum*, given the many special statutes that have been enacted over time, in addition to the General Part.[34] But the intrinsic existence of a code changes society. The code is interpreted by giving specific form to the textual rules it contains, considering the millenary and extremely rich Chinese cultural tradition. And 'law in action' differs from the 'law in books'.

A recent case decided by the Court of Beijing is very significant. The Court recently applied Art 1088 of the Civil Code, holding that

> when one spouse is burdened with the additional duties to raise children, care for the elderly, or assist the other spouse in his or her work, he or she is entitled to receive due compensation in the divorce proceedings:

this is a recognition of domestic work, a decisive achievement.[35]

A political-ideological justification can probably be found in the 'non-choice' of the common law model – the system used in the UK and the US, China's main competitors on the global arena.

To conclude, I would like to add that, in my opinion, in a system aiming for the primacy of law, a code is to be preferred, in terms of certainty and predictability, to a law based essentially on judicial decisions.

[34] Y. Bu, *Chinese Civil Code: The General Part* (München: Beck; Oxford: Hart Publishing; Baden-Baden: Nomos, 2019).

[35] *La Repubblica*, available at https://tinyurl.com/5sez2hee; *The Guardian*, available at https://tinyurl.com/3an6cz64; *New York Times*, available at https://tinyurl.com/3an6cz64 (last visited 30 June 2021). See X. He, *Divorce in China: Institutional Constraints and Gendered Outcomes* (New York: NYU Press, 2021) (analyzing current divorce law practices).

Essays

Atheism and the Principle of Secularism in the Italian Constitutional Order

Francesco Alicino*

Abstract

More diverse and more militant nonreligious groups are contributing to change the socio-cultural landscape of a growing number of constitutional democracies. Many of these groups and their various components (hard and soft atheists, agnostics, rationalists, humanists, secularists) are claiming to enjoy the protection of religious freedom, while straightforwardly denouncing the legal tendencies that give traditional confessions distinct privileges against generally applicable laws. This also raises several questions about when, where, and how groups of atheists should engage with religious issues and the legal degree to which such engagement becomes 'religion-like'. On the other hand, this is even more evident in legal contexts where the model for managing the State-religions relationship and even freedom of religion are characterized by overt or implicit endorsements towards traditional confessions that, as such, enjoy special protected legal status. One of the most preeminent examples of that is given by the Italian association of atheists (also known as UAAR), who in the last years have launched judicial review proceedings against what they considered the Italian limited secularism. In this manner, nowadays Italian atheism is helping to shed light on the contradictions of the biased pro-religion interpretations of some important constitutional rules, including those related to the supreme principle of secularism.

I. Introduction

Religion has now taken centre stage in public debate worldwide. It is frequently identified as both the cause of large-scale global conflicts and a main source of transnational solidarities. Over the last decades, however, there has been a reduction in the amount of religious people who are active in devout practices. Furthermore, many of them happen to be part of a confession more as a result of their culture than for spiritual or ideological reasons. An emergent number of believers, for instance, affirm to be Roman Catholics because they 'feel at home' with the Church's culture and teaching, but it is highly improbable that they would believe in a divine Jesus, in hell, and in the original sin. At the same time, a growing number of people, particularly young adults, distance themselves from religion.

Nonreligious people are indeed the second largest group in North America and most of Europe. Today's East Asian societies have the highest proportion of people reporting 'no religion.' Australia is seeing an increase in the non-believers, while in Latin America younger are less likely than their elders to say that religion is

* Full Professor of Public Law and Religion and Constitutional Law, LUM University.

very important.[1]

Younger adults are less likely than older adults to consider religion very important in 46 countries

In just two countries – Georgia and Ghana – older adults (ages 40+) are less likely to say religion is very important in their daily lives

Source: Pew Research Center surveys, 2008-2017.
"The Age Gap in Religion Around the World"

PEW RESEARCH CENTER

Younger adults less likely to say that religion is very important across regions

Number of countries with each outcome, by region

	Religion less important to younger adults	Religion less important to older adults	No significant difference
Overall	46	2	58
Asia-Pacific	5	0	15
Europe	18	1	16
Latin America	14	0	5
Middle East-North Africa	4	0	5
North America	2	0	0
Sub-Saharan Africa	3	1	17

Note: Younger adults are those ages 18 to 39; older adults are those 40 and older.
Source: Pew Research Center surveys, 2008-2017.
"The Age Gap in Religion Around the World"

PEW RESEARCH CENTER

This proves that pluralism and multiple religious perspectives have increased dramatically not only through the proliferation of different confessions living in the same geopolitical milieu, but also through the rising presence of at least three socio-cultural categories: unaffiliated believers; believers who, although they remain faithful to their denominational religion, adopt forms of personal spirituality; and nonreligious people who assert patent claims against the public role of religions, as part of what they see to be the realization of the promise of the secular democratic State.[2]

In reality, the position of religious nones[3] takes different forms. Indifference to

[1] Pew Research Center, 'Young adults around the world are less religious by several measures' (2018), available at https://tinyurl.com/1q52olx7 (last visited 30 June 2021). See also F. Yang, 'Religion in the Global East: Challenges and Opportunities for the Social Scientific Study of Religion' 9 *Religions*, 1-10 (2018); D. Balazka, *Mapping Religious Nones in 112 Countries: An Overview of European Values Study and World Values Survey Data (1981-2020)* (Trento: Fondazione Bruno Kessler, 2020).

[2] R. Hirschl and A. Shachar, 'Competing Orders? The Challenge of Religion to Modern Constitutionalism' 85 *The University of Chicago Law Review*, 424-485 (2018); S. Ferrari 'Religion Between Liberty and Equality' 4 *Journal of Law, Religion and State*, 179-193 (2016); A. Jamal and J.L. Neo, 'Religious Pluralism and the Challenge for Secularism' 7 *Journal of Law, Religion and State*, 1-12 (2019).

[3] The notion of 'religious nones' indicates the category of people who select 'no religion' when surveyed asking their religious affiliations. It refers to lack of organizational affiliation rather than lack of personal belief. In origin this expression was used in the US. Now it is commonly used

religious belief on the one hand and the criticism of confessional traditions on the other exemplify various way of being nonreligious.[4] This also raises several questions about when, where, and how the groups of religious nones should engage with religious issues and the legal degree to which such engagement becomes 'religion-like'.[5]

In most constitutional democracies, nonreligious people are no longer 'excluded from religious interests and considerations'.[6] For example, they can refer to schools of thought that take positions 'on religion, the existence and importance of a supreme being, and a code of ethics'.[7] Nonetheless, instead of emphasizing the collective dimension of their attitude, nonreligious people are largely considered as individualistic with a relatively high score. Recent signs of revers still remain though. These signs see an expanding number of nonreligious people organize themselves into associations, which helps atheists fight for their rights, including the right to not believe in god(s) and propagate their arguments either alone or in community, public or/and private.

The example is given by Italy, where in the last three decades atheist organizations have evolved the ability to make their voices heard. It is important to note that they are successful in doing so through various forms of judicial activities, like those being prompted and promoted by the Union of Rationalist Atheists and Agnostics (*Unione degli Atei e degli Agnostici Razionalisti*) also known as UAAR.

One of these actions, for example, originated in 1996, when UAAR launched judicial review proceedings against the pro-religion *ex parte Ecclesia* method of bilateral legislation, as laid down in Arts 7.2 and 8.3 of the 1948 Italian Constitution. After a protracted legal battle, in 2016 this initiative resulted in the judgement (no 52/2016) of the Italian Constitutional Court, and it is now waiting for a decision of the European Court of Human Rights. We should also not forget that this initiative fits into a greater judicial enterprise, such as that pertaining to the displaying of crucifix in public spaces, religious teaching in schools, the system of 0,008 of the

throughout the world. See J. Thiessen and S. Wilkins-Laflamme, 'Becoming a Religious None: Irreligious Socialization and Disaffiliation' 56 *Journal for the Scientific Study of Religion*, 64-82 (2017).

[4] On the working definition of atheism A. Payne, 'Redefining "Atheism" in America: What the United States Could Learn from Europe's Protection of Atheists' 27 *Emory International Law Review*, 663-703 (2013); G.M. Epstein, *Good Without God: What a Billion Nonreligious People Do Believe* (New York: Harper Collins, 2009); R. Arons, *Living Without God: New Directions for Atheists, Agnostics, Secularists, and the Undecided* (Berkeley: Counterpoint Press, 2008).

[5] J. Thiessen and S. Wilkins-Laflamme, *None of the Above: Nonreligious Identity in the US and Canada* (New York: New York University Press, 2020); J. Schuh, C. Quack and S. Kind, *The Diversity of Nonreligion: Normativities and Contested Relations* (London: Routledge, 2020); S. Baldassarre, *Codice europeo della libertà di non credere. Normativa e giurisprudenza sui diritti dei non credenti nell'Unione Europea* (Roma: Nessun Dogma, 2020).

[6] As the Italian Constitutional Court stated in 1960 (no 58/1960 n 33 above). See A. Origone, 'La libertà religiosa e l'ateismo', in VvAa, *Studi di diritto costituzionale in memoria di Luigi Rossi* (Milano: Giuffrè, 1952), 417.

[7] The US's Seventh Circuit Court of Appeals, *Kaufman* v *McCaughtry*, 419 F.3d 678 (7th Cir. 2005).

income taxes in support of the Catholic Church and other few confessions,[8] and the right to freely propagate atheistic messages through communication campaigns, the so-called 'right to blasphemy'. Thus, even in a traditionally Catholic country and from a tiny minority that rarely have access to media megaphones, the Italian atheists are now marking a pivotal moment in their history.

In order to better understand this attitude, it is imperative to focus on two factors. First, the evolution of the way in which Italian legal system considers the right to freedom of religion. Second, the traditional roots and essential characteristics of the Italian State-Churches relationship. As we will see, in respect of these issues the judiciary courts are playing a crucial role, just when Italy is facing not only waves of new immigration, which are quickly changing the country's religious landscape – like it was in Great Britain and France during the 1960s and 1970s. Italy is experiencing a more flexible relation between people and religion, which is a typical feature of those who value the sense of belonging to religious communities for some of their precepts while interpreting others in a completely personal way.

The first part of this article focuses on the current religious multiplicity, which includes the rising presence of nonreligious people and the related organizations. This will give the opportunity to clarify the peculiar characteristics of the Italian legal order in relation to the current expression of collective atheism, which necessarily entails the ways of understanding and viewing religion and the religious experience in the country. From this point of view, it is important to consider the way in which the Italian legal system defines atheism, taking also into account the national and supranational provisions regulating the right to freedom of religion and belief.

Then, the article highlights the role played by judiciary courts, especially when considering the method of bilateral (State-Churches) legislation. This method seems attractive to some confessional organisations, while creating unfavourable distinctions for other groups, including those related to religious nones. In this manner, today's militant atheism is helping to shed light on the contradictions of the biased pro-religion interpretations of the Italian constitutional order, including the supreme principle of secularism.[9]

II. The Collective Forms of Atheism

More diverse and more militant nonreligious groups, also known as religious nones, are contributing to change the socio-cultural landscape of a growing

[8] On the 0,008 of taxes owed by natural persons, also known as *IRPEF*, see F. Alicino, 'Un referendum sull'otto per mille? Riflessioni sulle fonti' *Stato, Chiese e pluralismo confessionale*, 28 October 2013, 1-35.

[9] I refer to the '*principio supremo di laicità*', as the Italian Constitutional court calls it. See below, paras V and VI.

number of constitutional democracies. Many of these groups and their various components (hard and soft atheists, agnostics, rationalists, humanists, secularists) are claiming to enjoy the protection of religious freedom, while straightforwardly denouncing the legal tendencies that give traditional confessions distinct privileges against generally applicable laws. According to religious nones, this is in contrast with both the principle of equality and a coherent implementation of the concept of a secular democratic State.

For their part, religions, especially the most popular ones, continue to claim a peculiar role in society, which distinguishes traditional confessions from other 'common' associations. In their view, democratic pluralism and the State neutrality on religion infer neither hostility nor indifference to religions. Moreover, several religious representatives maintain that freedom of religion should be interpreted to mean that atheism has little or nothing to do with the collective dimension of religious experiences. As a matter of fact, atheist organizations do not merit the same level of legal guarantees that religious groups command. From that perspective, it is also interesting to observe that many atheists are also concerned by some legal systems, which require groups of religious nones to pose as 'religious' organizations to receive equal treatment. That is another piece of an already confused puzzle of constitutional law on what qualifies as 'religion'.[10]

In fact, all of this signals that traditional religions and today's atheist groups often collide on policy preferences and the true essence of contemporary constitutionalism. It is not by chance that these diverging viewpoints and interests also manifest themselves through high-profile legal clashes and court cases, in which the stakes for the competing parties they represent are both high and visible.[11] Protection of gender equality, abortion, reproductive freedoms, LGBTQ rights, same-sex marriage, the right to die with dignity, the display of religious symbols in public spaces, the right to ridicule and mock religion are considered

[10] See *ex plurimis*: *McCreary County, Ky.* v *American Civil Liberties Union of Ky.*, 545 US 844, 125 S.Ct. 2722, 162 L.Ed.2d 729 (2005); *Kaufman* v *McCaughtry*, US, August 19, 2005, no 04-1914; Supreme Court of Canada, *Mouvement laïque québécois* v *Saguenay* (City), 2015 SCC 16, [2015] 2 SCR 3; Canada Federal Court of Appeal, *Church of Atheism of Central Canada* v *Minister of National Revenue*, 2019 FCA 296, 2019, 29, 11; Supreme Court of Canada, *Syndicat Northcrest* v *Amseleum*, [2004] 2 SCR 551. On the relation between the definition of religion and atheistic organizations see also: R. Dworkin, *Religion Without God* (Boston: Harvard University Press, 2013); C. Miller, ' "Spiritual but Not Religious": Rethinking the Legal Definition of Religion' 102 *Virginia Law Review*, 833-894 (2016); D.H Davis, 'Is Atheism a Religion? Recent Judicial Perspectives on the Constitutional Meaning of "Religion" ' 47 *Journal of Church and State*, 707-723 (2005); L.H. Tribe, *American Constitutional Law* (New York: The Foundation Press, 1978), 417; C. Crockett, 'On the Disorentiation of the Study of Religion', in T. Idinopulos and C. Wilson eds, *What Is Religion? Origins, Definitions, and Explanations* (Leiden-Boston-Köln: Brill, 1998), 1-13; G. Laneve, 'Atheism as Part of Religious Phenomenon: Questions and New Challenges to Secularism' 25 *federalismi.it*, 154-181 (2020).

[11] A. Connaughton, 'Religiously unaffiliated people more likely than those with a religion to lean left, accept homosexuality' *Pew Research Center* (28 September 2020); H. Harting, 'Nearly six-in-ten Americans say abortion should be legal in all or most cases' *Pew Research Center* (17 October 2018).

some of the hallmarks of the current jurisprudence.

In this regard, it is useful to recall the 2011 *Lautsi and others* v *Italy*[12] decision of the European Court of Human Rights (ECtHR), which involved the human-rights claim of a mother residing in Italy who objected to the display of religious crucifixes in her sons' public schools. This is a case that was supported by UAAR and other European associations of religious nones and that, during the process, brought together strange bedfellows of religious groups like American Conservative Evangelicals, the Russian Orthodox Church, and the Vatican, all united by their advocacy of Christian symbols in the European public sphere.[13] One can also take into account the 2013 *Eweida* v *United Kingdom*[14] decision of the ECtHR (holding that Art 9 of the ECHR was violated when a British Airways flight attendant was prohibited from wearing a visible cross at work) and the 2014 *S.A.S.* v *France*[15] judgement of the ECtHR (ruling that the French laws banning the Islamic full-face veil did not breach the ECHR because the State autonomy and regulatory powers over attire in public spaces trump considerations of religion-based freedoms).[16]

Similarly, it is possible to point out with reference to the 2017 decisions of the Court of Justice of the European Union regarding *Achbita* v *G4S Secure Solutions NV* (affirming that, under certain conditions, employers may dismiss employees who refuse to comply with company policies concerning religious attire)[17] and the ECtHR's *Eweida* v *United Kingdom* judgement (ruling that a religious organization's claim to religious autonomy was sufficient to trump the claimant's right to respect for his private life).[18]

Finally, it is worth mentioning other important cases concerning the right to produce satire in France[19] and other sensitive issues, such as those referring to the 2015 US Supreme Court *Obergefell* v *Hodges* decision[20] (ruling that under the Equal Protection Clause of the Fourteenth Amendment marriage is a fundamental right guaranteed to all couples, including same-sex ones), the 2018

[12] Eur. Court H.R. (GC), *Lautsi and Others* v *Italy* App no 30815/06, Judgment of 18 March 2011.

[13] See P. Annicchino, 'Winning the Battle by Losing the War: The Lautsi Case and the Holy Alliance between American Conservative Evangelicals, the Russian Orthodox Church and the Vatican to Reshape European Identity' 6 *Religion & Human Rights*, 213, 215-18 (2011).

[14] Eur. Court H.R., *Eweida* v *United Kingdom* App nos 48420/10, 59842/10 and 36516/10, Judgment of 15 January 2013.

[15] Eur. Court H.R. (GC), *S.A.S.* v *France* App no 43835/11, Judgment of 1 July 2014.

[16] In this sense see also: Eur. Court H.R., *Dakir* v *Belgium* App no 4619/12, Judgment of 11 July 2017; *Belkacemi and Oussar* v *Belgium* App no 37798/13, Judgment of 11 July 2017.

[17] Case C-157/15 Samira Achbita e Centrum voor gelijkheid van kansen en voor racismebestrijding v G4S Secure Solutions NV, [2017] ECLI:EU:C:2017:203.

[18] Eur. Court H.R., *Eweida* v *United Kingdom* n 14 above.

[19] F. Alicino 'Freedom of Expression, Laïcité and Islam in France: The Tension between Two Different (Universal) Perspectives' 27 *Islam and Christian-Muslim Relations* 51-57 (2015); Id, 'The Italian legal system and imams. A difficult relationship', in M. Hashas, J.J. de Ruiter and N. Valdemar Vinding eds, *Imams in Western Europe. Developments, Transformations, and Institutional Challenges* (Amsterdam: Amsterdam University Press, 2018), 359-380.

[20] *Obergefell* v *Hodges* 576 US 14-556 (2015).

UK's *Alfie Evans* case (involving an infant with a GABA-transaminase deficiency),[21] and the 2019 *DJ Fabo's* of the Italian Constitutional Court (holding that assisted dying is not a crime if some persons wanting to end their life are experiencing intolerable suffering).[22]

These decisions are merely examples of the fact that often the stance of the traditional confessions is opposed to the secular view of religious nones. Nevertheless, and for the same reasons, this is even more evident in contexts where the model for managing the State-religions relationship and even freedom of religion is characterised by overt or implicit endorsements towards traditional confessions that, as such, enjoy special protected legal status. One illustrative and interesting example of that resides in Italy. In particular, it resides in the historical role played by the bilateral State-Churches normative instruments, as primarily affirmed in arts 7.2 and 8.3 of the 1948 Constitution.

III. Religion and the Italian Constitutional Order

The Italian population has many different ways of understanding and viewing religious belonging.[23] The tendency to think of oneself as Catholic, for example, is much more widespread than considering oneself unrelated to religious teachings. It is worth pointing out that this situation does not reproduce individualism in belief or the so-called *religion à la carte*, through which any person becomes the locus of his/her own religion. Despite uncertain and ambivalent convictions, most of the Italian citizens prefer to identify themselves as belonging to an official religion, primarily Catholicism.[24]

This explains the limited number of atheists and agnostics in the country.[25] People that do not believe or join any particular religion are constantly increasing in many European States: for example, they amount to thirty-five, forty percent of the population in France, Belgium and Germany, while nonreligious people have overtaken Christians as the majority position among white British population.[26] On the contrary, in Italy the corresponding number stands at around nine percent

[21] *In the matter of Alfie Evans* [2018] UKSC, 20 April 2018.

[22] Corte costituzionale 25 settembre 2019 no 242, *Il Foro italiano*, I, 829 (2020).

[23] F. Garelli, *Religion Italian Style. Continuities and Changes in a Catholic Country* (Farnham: Ashgate, 2014), 87.

[24] R.W. Bibby, *La religion à la carte: pauvreté et potentiel de la religion au Canada* (Montréal: Fides, 1988), 110; L. Witham, *Marketplace of the Gods. How Economics Explains Religion* (Oxford: Oxford University Press, 2010), 156-158; S. Lefebvre, 'Religion in Court, Between an Objective and a Subjective Definition', in L.G. Beaman ed, *Reasonable Accommodation. Managing Religious Diversity* (Vancouver-Toronto: UBCPress, 2012), 32-51.

[25] Of course, this picture changes in accordance with the socio-demographic characteristics of the population, or the different contexts where people live.

[26] It is interesting to note that the UK's 2010 Equality Act expressly states 'Religion means any religion and a reference to religion includes a reference to a lack of religion. Belief means any religious or philosophical belief and a reference to belief includes a reference to a lack of belief.'

and it has shown no particular growth trend over the last decades. This is because many Italians, including those who do not believe in God, consider religion in general and Catholicism in particular as reference in terms of their culture of origin and national identity.[27]

One result of this attitude is the low attendance to ordinary religious practices (Sunday worship service, private prayer, study and reading of the holy scriptures, etc), on the one hand, and the tendency to focus the attention on the great religious events (the proclamations of saints, the Pope's visits to local dioceses, the commemoration of charismatic religious figures), on the other hand. In addition, a vast majority of Italians participate in religious rites of passage (baptisms, church weddings, religious funerals), which are often seen as solemn celebrations of the most important moments in a person's life, as well as in the life of the local and, at times, even national community. Conversely, an important part of the Catholic world normally deserts parishes. From here stems one of the paradoxes of people's religious behaviour in Italy: it is still able to fill the public squares, while the churches remain substantially empty.[28]

It is important to note that this situation is also the result of the unique historical process, which has influenced the way the State effectively governs religious issues, including those referring to the right to freedom of religion and its relationship with atheism.

IV. Atheism and the Italian Constitutional Order

In Italy, freedom of religion is primarily regulated by Art 19 of the 1948 Constitution, which establishes that anyone is entitled to freely profess religious faiths in any form, individually or with others, and to propagate religions and celebrate rites in public or in private, provided they are not offensive to public morality. So far as the collective dimension of religious experience is concerned, this provision should be interpreted in combination with Art 8 of the Constitution, which states that all religious confessions enjoy equal freedom before the law. In addition, Art 20 of the Constitution affirms that no special legislative limitation or tax burden may be imposed on the establishment, legal capacity or activities of any association or institution on the ground of its ecclesiastical nature or its

[27] E. Drescher, *Choosing Our Religion. The Spiritual Lives of America's None* (New York: Oxford University Press, 2016), 16-52; G. Zurlo and T.M. Johnson, 'Unaffiliated, Yet Religious: A Methodological Demographic Analysis', in R. Cipriani and F. Garelli eds, *'Sociology of Atheism' Annual Review of the Sociology of Religion* (Leiden/Boston: Brill, 2016), 50-75; T. Cragun Ryan et al, 'On the Receiving End: Discrimination toward the Non-Religious in the United States' 27 *Journal of Contemporary Religion*, 105-112 (2012); L. Woodhead and A. Brown, *That Was The Church That Was: How the Church of England Lost the English People* (London: Bloomsbury Publishing, 2016); F. Garelli, *Religion Italian Style* n 23 above, 90.

[28] F. Garelli, *Gente di poca fede. Il sentimento religioso nell'Italia incerta di Dio* (Bologna: il Mulino, 2020), 3-7.

religious or worship purposes.

It can be easily noted that in these dispositions there is no reference to the freedom of thought and the freedom of conscience which, on the contrary, are expressly mentioned in many other national and supranational legal documents, including the Universal Declaration of Human Rights (Art 18) and the European Convention of Human Rights (Art 9).

As for the collective dimension of religious freedom, the Italian Constitution also refers to denominations (Art 8), religious faiths (Art 19) and the ecclesiastical nature, or religious or worship purposes of associations or institutions (Art 20). In these cases, the 1948 Constitution does not mention beliefs, associations or institutions other than denominational ones;[29] which may constitute an obstacle for nonreligious people who, for example, would want to see the promotion of atheism protected under the constitutional rules.

Once again, that is evident in the light of other legal documents, which are able to include groups of religious nones under the protection of the right to freedom of thought, conscience and religion. The example is given not only by the European Convention of Human Rights,[30] but also by the European Union (EU) law. According to this law, the Union equally respects the status of Churches and religious associations or communities and the status of philosophical and non-confessional organisations, while maintaining an open, transparent and regular dialogue with them.[31] Sometimes the EU's law goes even further, stating that the concept of religion 'shall include the holding of theistic, non-theistic and atheistic beliefs'.[32]

[29] In Germany, for example, the status of non-denominational organizations is established in the Constitution. In particular, Art 140 of the *Grundgesetz* states that associations pursuing philosophical ideology have the same status as religious groups. In other words, both religious groups and philosophical organizations (*Weltanschauungsgemeinschaft*, which includes humanistic and atheistic associations) may have the status of public law corporations (*Körperschaft des öffentlichen Rechts*, also known as *KdÖR*). It also means that each Lander-State is entitled to grant the *KdÖR* to atheistic associations that, in this way, may benefit some distinct rights against generally applicable laws. Thanks to such equal legal treatment between religious denominations and philosophical organizations, the Land of Lower Saxony, for instance, has signed an agreement with the *Freireligiösen Landesgemeinschaft Niedersachsen*, a local atheistic association.

[30] See Art 9 ECHR, which 'includes freedom to change his religion or belief and freedom, either alone or in a community with others and in public or private, to manifest his religion or belief, in worship, teaching, practice and observance'. At the same time, the ECHR declares that the State shall respect the right of parents to ensure right to education and teaching in conformity with their own religious and philosophical convictions' (Art 2 of the 1st Protocol to the Eur Court H.R.).

[31] Art 17 of the Treaty on the Functioning of the European Union.

[32] Council Directive 2004/83/EC of 29 April 2004 on minimum standards for the qualification and status of third-country nationals or stateless persons as refugees or as persons who otherwise need international protection and the content of the protection granted [2004] OJ L304/12, Art 10.1(b). In this vein, it is also important to note that the EU's Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation [2000] OJ L303/16 establishes a general framework for equal treatment in employment and occupation and prohibits direct and indirect discrimination, harassment, instructions to discriminate and victimization on grounds of religion or belief. On this see Case C-414/16 *Vera Egenberger* v *Evangelisches Werk fur Diakonie und Entwicklung e.V.*, [2018] ECLI:EU:C:2018:257,

To this regard, it is important to consider what the Italian Constitutional Court (ICC) affirmed in the 1960 decision (no 58), according to which freedom of religion in general and Art 19 of the Constitution in particular do not

> involve the protection of all forms of freedom of thought. Specifically, it does not imply atheism.

This is because atheism 'ends where a religious experience begins'.[33] It means that, since freedom of religion applies exclusively to persons who believe in a traditional religion, atheists cannot benefit from that constitutional protection.

However, in 1979, in the light of the pressing demand for the effective implementation of international human rights – which invariably refers to freedom of religion or belief, where belief includes not only religious but also nonreligious beliefs such as humanism, atheism and agnosticism – the Court reversed two decades of its own jurisprudence. The ICC affirmed that Art 19 of the Italian Constitution encompasses all manifestations of freedom of thought which, in a way of another, are correlated to religion. Moreover, the same Court stated that the Italian constitutional order

> does not legitimate differentiation of protection between the expressions of religious faith and the expressions of disbelief.[34]

Hence, since 1979 Art 19 of the Constitution has not only implied the protection of religious persons. It has also ensured an equivalent level of respect to religious nones. In other words, since 1979 religious freedom in Italy has inferred the protection of public and private phenomena that, from a philosophical and ideological point of view, could be located in-between two poles of legal concern: the positive pole, related to people who believe in a confessional organisation and the related precepts; and the negative pole, which may take the form of a sense of scepticism and realism suggesting, for example, that fear and superstition are mothers of all religions. In brief, since 1979 the Italian Constitution has been interpreted in a way that allows the protection of both religious people and atheists in their right to freely profess (religious or nonreligious) beliefs, whether they are acting individually or collectively, in private or in public.

It should be noted that when the ICC issued the 1979 judgement in Italy atheism existed but only in the form of individual attitude. In the late 1970s there was no organisation of religious nones capable of connecting isolated individual

where the European Court stated that the right of autonomy of Churches and the right of workers must be subject of an assessment aimed to ensure a fair balance between them.

[33] Corte costituzionale 13 July 1960 no 58 (translation mine), *Giurisprudenza costituzionale*, 752 (1960).

[34] Corte costituzionale 10 October 1979 no 117 (translation mine), *Il Foro italiano*, I, 625 (1981). See P. Bellini, 'L'ateismo nel sistema delle libertà fondamentali' 1 *Quaderni di diritto e politica ecclesiastica*, 85-90 (1985) and P. Floris, 'Ateismo e religione nell'ambito del diritto di libertà religiosa' *Il Foro italiano*, I, 5 (1981).

experience at the national level. Therefore, in 1979 the constitutional decision and its new way of considering atheism did not really involve the collective dimension of religious freedom. At the same time, we cannot forget that, since the Lateran Pacts were approved (1929), the collective dimension of religious freedom has been largely governed by the method of State-Churches bilateralism. Furthermore, this method has always referred to the special relationship between the State and the Roman Catholic Church. As such, it has affected the way in which atheism has been legally defined in the history of the Italian Republic.

V. The Principle of Secularism and Atheism

It is important to recall that the unification of Italy in 1871 abolished the secular-territorial power of the Catholic Church, which generated the hostility of the ecclesiastical hierarchy towards the newborn political entity. On the other hand, the predominantly moderate policy of the Italian State made its relationship with the Roman Church progressively less tense; so much so that, during the Fascist period, the Italian government and the ecclesiastical hierarchy were able to stipulate the Lateran Pacts, which was a turning point in the history of the Italian legal system.[35] Not only the 1929 Pacts were considered as a legal framework to reconcile two parties, the Roman Catholic Church and the Kingdom of Italy. The Pacts also established a proactive role for this Church in legally defining some religious issues, such as teaching of religion in schools, presence of ecclesiastical hierarchies in public debates, the State funding to the confessions, legal punishments for offences against religion,[36] and criminal law provisions related to blasphemy against the deity (*la divinità*), religious symbols and religious authorities.[37]

Most of all, the 1929 Pacts laid the foundation for a method of bilateral State-Church collaboration that, since the brand-new Republic of Italy entered into force in 1948, has partially been extended to religions other than Catholicism. This has been made possible by Arts 7 and 8 of the Constitution, which highlight the historical bonds between the State and the Catholic Church.[38]

[35] F. Ruffini, *Corso di diritto ecclesiastico. La libertà religiosa come diritto pubblico subiettivo* (Torino: F.lli Bocca, 1924); F. Margiotta Broglio, *Italia e Santa Sede dalla grande guerra alla conciliazione* (Roma-Bari: Laterza, 1966), 86; R. Pertici, *Chiesa e Stato in Italia. Dalla Grande Guerra al nuovo Concordato. Dibattiti storici in Parlamento* (Bologna: il Mulino, 2009), 189; A. Ferrari, 'The Italian Accommodations. Liberal State and Religious freedom in the 'Long Century' ', in L. Derocher et al eds, *L'État canadien et la diversité culturelle et religieuse 1800-1914* (Québec: Presses de l'Université du Québec, 2009), 143-153.

[36] Arts 402-406 of the 1930 Italian Penal Code.

[37] Art 724 of the 1930 Italian Penal Code. See C.A. Jemolo, *Chiesa e Stato negli ultimi cento anni* (Torino: Einaudi, 1971), 537. See also Corte costituzionale 18 October 1995 no 440, *Giurisprudenza italiana*, I, 178 (1996).

[38] C. Cardia, 'Concordato, intese, laicità dello Stato' 1 *Quaderni di diritto e politica ecclesiastica* 30 (2004); N. Colaianni, *Confessioni religiose e intese* (Bari: Cacucci, 1990), 35; P. Floris, 'Laicità e

Art 7 establishes that the State and the Catholic Church are independent and sovereign, each within its own sphere. Albeit weaker, this principle is also affirmed in Art 8.2 of the Constitution, which guarantees the free organisation of religious denominations other than Catholicism. At the same time, Art 7.2 declares that the Lateran Pacts regulate the State-Church relationships and that a change to these Pacts, when accepted by both parties, does not require the procedure of constitutional amendments.[39] It means that, when there is a State-Church bilateral agreement, a legislative (non-constitutional) act is sufficient in order to amend the 1929 Pacts. Another point of reference for the method of bilateralism is Art 8.3 of the Constitution, which affirms that legislative acts regulate the relationships between the State and minority religions. These acts, however, must be based on *intese*, which can be translated literally as 'understandings' between the State and denominations other than Catholicism.

Thus, once the Italian government and the representatives of a given denomination have signed an agreement (Art 7.2) or an *intesa* (Art 8.3), these documents need to be ratified (agreement) or approved (*intesa*) by specific acts of the Italian Parliament.[40] In this manner, the Catholic Church and minority religions holding an *intesa* have the guarantee that their legal status, benefits and privileges cannot be altered without considering their will. This also explains why, in order to keep the special status within the State's territory, some confessional organisations, in particular the Catholic Church and minority religions with *intese*, intensely support the principle of bilateralism.[41]

However, not all minority confessions are able to sign an understanding with the State. The method of bilateralism generates two main problems. First, it presupposes a relatively comprehensive religious institution capable of representing a denomination at the national level. This requirement was proved to be very challenging for some religious organisations, such as those referring to Islam.[42]

collaborazione a livello locale. Gli equilibri tra fonti centrali e periferiche nella disciplina del fenomeno religioso' *Stato, Chiese e pluralismo confessionale*, February 2010, 5.

[39] This procedure is provided by Art 138 of the Constitution.

[40] Concerning the recent relationships between the State and the Catholic Church, on 18 February 1984 the State and the Holy See signed an agreement, which was then ratified by the law of the Italian Parliament (legge 25 March 1985 no 121). This law is an atypical *sui generis* legislation because, once it enters into force, it can be amended only on the basis of a new agreement between the State and the Church: no amendment based on a unilateral legislation made by the Parliament is possible. The same can be said about the legislative acts approving *intese*: they can only be changed via additional legislative acts that, in turn, must be based on further understandings between the State and confessions concerned.

[41] G. Bouchard, 'Concordato e intese, ovvero un pluralismo imperfetto' *Quaderni di diritto e politica ecclesiastica*, 70 (2004); G.B. Varnier, 'La prospettiva pattizia', in V. Parlato and G.B. Varnier eds, *Principio pattizio e realtà religiose minoritarie* (Torino: Giappichelli, 1995), 8-13; S. Ferrari, 'Il Concordato salvato dagli infedeli', in T. Valerio ed, *Studi per la sistemazione delle fonti in materia ecclesiastica* (Salerno: Edisud, 1993), 127-158; M. Ventura, *Creduli e increduli. Il declino di Stato e Chiesa come questione di fede* (Torino: Einaudi, 2014), 58.

[42] C. Decaro Bonella, 'Le questioni aperte: contesti e metodo', in Id, *Tradizioni religiose e tradizioni costituzionali. L'Islam e l'Occidente* (Roma: Carocci, 2013), 34-35.

The second problem is caused by the excessive amount of discretion that the Government possesses in deciding whether to accept or reject the proposal made by an organisation to enter into negotiations for concluding an understanding.[43]

Besides, over the last thirty years the practical implementation of Art 8.3 of the Constitution has been characterised by the phenomenon of the so-called 'copy&paste understandings' (*intese fotocopia*); that is by the substantial similarity of all *intese* which have been signed by minority religions until now.[44] As a result, these *intese* have established a *de facto* common legislation, which is far from being considered general legislation: it is common to all religious denominations that have signed an understanding, but it cannot be applied to other organisations that do not have an *intesa* yet.[45]

As a matter of fact, religious groups without *intese* are subject to the 1929 law (no 1159) on 'admitted religions' that, approved during the fascist regime, legitimises an even greater discretionary power by the Italian Government.[46] On the contrary, religious groups possessing an understanding with the State are no longer subject to the 1929 law, whose provisions are entirely replaced by those (more favourable) affirmed in the legislative acts approving *intese*.[47]

These difficulties are even more evident in the light of the rules stated in Arts 2, 3, 19 and 20 of the Constitution that, together with Arts 7 and 8, in 1989 led the Constitutional Court to define secularism (*laicità*) as one of the supreme principles (*principi supremi*) of the Italian constitutional order. As such, secularism does not require indifference to religions. It requires the equidistance and the impartiality of the State law, especially when related to religious issues.[48] It also

[43] See Corte costituzionale 10 March 2016 no 52, *Il Foro italiano*, I, 1940 (2016). See also F. Alicino, 'La bilateralità pattizia stato-confessioni dopo la sentenza n. 52/2016 della Corte costituzionale' *osservatoriosullefonti.it*, 1-16 (2016).

[44] See https://tinyurl.com/smj6bw5t (last visited 30 June 2021).

[45] V. Crisafulli, 'Fonti del diritto (dir. cost.)' *Enciclopedia del diritto* (Milano: Giuffrè, 1968), XII, 948; F. Carnelutti, *Teoria generale del diritto* (Roma: Soc. ed. del Foro italiano, 1951), 35; M. Ricca, *Legge e Intesa con le confessioni religiose: sul dualismo tipicità-atipicità nella dinamica delle fonti* (Torino: Giappichelli, 1996), 35; B. Randazzo, *Diversi ed eguali. Le confessioni religiose davanti alla legge* (Milano: Giuffrè, 2008), 55.

[46] According to the 1929 law, the Minister of Interior will take into consideration the assets of the denomination or religious entity that claims recognition. For example, he will take into account: 1) the number of the claimants' members and how widespread they are in the Country; 2) the compatibility between the claimants' statute and the main principles of the Italian legal system; 3) the aim of the denomination that claims to be recognised by the State, an aim that has to be 'prevalently' of religion and worship.

[47] On this aspect see R. Zaccaria et al eds, *La legge che non c'è. Proposta per una legge sulla libertà religiosa* (Bologna: il Mulino, 2019); in particular see the following articles: P. Floris, 'Le istanze di libertà collettiva e istituzionale', 145-190, and F. Alicino, 'I problemi pratici e attuali della libertà religiosa', 235-246.

[48] See the following decisions of the Italian Constitutional Court: 12 april 1989 no 203, *Il Foro italiano*, I, 133 (1989); 25 May 1990 no 259, *Giustizia civile*, I, 2504 (1990); 14 January 1991 no 13, *Il Foro italiano*, I, 365 (1991); 27 April 1993 no 195, *Il Foro italiano*, I, 2986 (1994); 1 December 1993 no 421, *Il Foro italiano*, I, 14 (1994); 8 October 1996 no 334, *Il Foro italiano*, I, 25 (1997); 14 November 1997 no 329, *Il Foro italiano*, I, 26 (1998); 20 November 2000 no 508, *Il Foro italiano*,

means that, compared to the previous (Fascist) regime, there can no longer be an unreasonable (not constitutionally based) distinction. This is true not only with reference to the comparison between the Catholic Church and other confessional denominations. It is equally true when comparing the minority religions that have signed an *intesa* and those organisations that do not possess an understanding with the State.[49]

The implementation of the supreme principle of secularism has therefore revealed other interconnected problems, which are partly due to the pro-religion vision of the method of bilateralism. That is even more evident when considering today's neo-religious and cultural pluralism, which implies an increasingly important role for organisations of nonreligious people. After all, it is not by chance that many atheists and agnostics consider the method of bilateralism as the major driving force behind Italy's limited *ex parte Ecclesiae* secularism.

VI. The Italian Method of Bilateralism and Militant Atheism

In Italy the interpretation of the constitutional rules concerning the State-religions relationship remains tailored on the notion of traditional confessions. In turn, this notion is manly based on the Catholic Church's model of organisation. Thus, under the current unprecedented cultural pluralism, the Italian law does not seem to be consistent with a modern, secular democracy. On the contrary, it seems characterised by a limited secularism or, as some have said, 'a baptised *laicità*'.[50] The example is given by the method of bilateral State-Churches legislation, which is becoming increasingly difficult and, at times, harshly contested by many organisations. These include groups of religious nones that, in the meanwhile, are seeking a greater role in the public space as well as in the political arena.

It should be stressed that in the last decades the Italian atheism has

I, 26 (2002); 9 July 2002 no 327, *Il Foro italiano*, I, 2941 (2002). See also N. Colaianni, 'Laicità: finitezza degli ordini e governo delle differenze' *Stato, Chiese e pluralismo confessionale*, 9 December 2013, 39; G. Dalla Torre, 'Ancora sulla laicità. Il contributo del diritto ecclesiastico e del diritto canonico' *Stato, Chiese e pluralismo confessionale*, 3 February 2014, 4.

[49] V. Tozzi, 'Le confessioni religiose senza intesa non esistono', in *Aequitas sive Deus. Studi in onore di Rinaldo Bertolino* (Torino: Giappichelli, 2011), 1033-1055; G. Casuscelli, 'La rappresentanza e l'intesa', in Alessandro Ferrari ed, *Islam in Europa/Islam in Italia tra diritto e società* (Bologna: il Mulino, 2008), 285-322; N. Colaianni, *Diritto pubblico delle religioni. Eguaglianza e differenze nello Stato costituzionale* (Bologna: il Mulino, 2012), 68; F. Finocchiaro, *Diritto ecclesiastico*, updated by A. Bettetini and G. Lo Castro (Bologna: Zanichelli, 2012), 120; A. Bettetini, 'Commento all'art. 20 Cost.', in B. Raffaele, A. Celotto and M. Olivetti eds, *Commentario alla Costituzione* (Torino: UTET, 2006), 441-448; M. Ricca, 'Art. 20 della Costituzione ed enti religiosi: anamnesi e prognosi di una norma "non inutile" ', in *Studi in onore di Francesco Finocchiaro* (Padova: CEDAM, 2000), 1557-1580; P. Di Marzio, *L'art. 20 della Costituzione. Interpretazione analitica e sistematica* (Torino: Giappichelli, 1999); S. Fiorentino, 'Gli enti ecclesiastici e il divieto di discriminazione', in G. Casuscelli ed, *Nozioni di diritto ecclesiastico* (Torino: Giappichelli, 2006), 57-68.

[50] A. Ferrari, 'De la politique à la technique: laïcité narrative et laïcité du droit. Pour une comparaison France/Italie', in Basdevant-Gaudemet Brigitte and Jankowiak François eds, *Le droit ecclésiastique en Europe et à ses marges (XVIII-XX siècles)* (Leuven: Peeters, 2009), 333-349.

experienced a significant evolution. It has moved from a purely individual dimension to a rampant militant activism. As such, it has called into question the moral ascendancy of religion and its importance for civic belonging and national identity. Italian atheists has thus asked for the outlawing of many practices related to the privileged position of religions (especially of the Catholic Church) in public life, as demonstrated by several indicators (ie the display of the crucifix in classrooms, the legal impossibility of renouncing one's baptism,[51] the teaching of religion in classes, the system of 0.008 of the *IRPEF*).[52] In so doing, associations of nonreligious people argues that, even if they can enjoy many rights as individuals, it is difficult for them to identify with the State's law as a group. The legal system weakens the sense of belonging of many atheists, giving them the impression that they are condemned to remain eternally beyond the constitutional boundary of the Italian citizenry.[53]

Religious nones have consequently sued the State authorities on several occasions, challenging their activities on religious issues, including the method of bilateralism. Moreover, in this specific matter the Italian militant atheism has demonstrated its intention to take the bull by the horns. The most important example of this is the above-mentioned Italian Union of Rationalist Atheists and Agnostics (UAAR), which in 1996 requested the Government to initiate negotiations to sign an *intesa* with the State.[54] This was not possible, the President of the Council of Ministers replied, simply because the applicant was not eligible to be included in the national list of confessional beliefs. In addition, the President held that the refusal to accept an association's request to launch negotiations could not be subject to judicial review, as this would violate the sphere of constitutional powers vested in the Government.[55]

Nonetheless, UAAR decided to bring the case before the Court, which has

[51] In this sense UAAR offers the 'Debaptism Certificate,' see https://tinyurl.com/yuqtrns5 (last visited 30 June 2021), whose procedure has been partially validated by the Italian Data Protection Authority (*Garante per la protezione dei dati personali*), see https://tinyurl.com/1303glaa (last visited 30 June 2021). See also the Italian Bishops' Conference (CEI), 1999. Decreto Generale, Disposizioni per la tutela del diritto alla buona fama e alla riservatezza, Prot no 1285/99, Art 2, para 9.

[52] According to this system, all Italian taxpayers can participate to a sort of 'poll' to allocate 0.008 of their income tax (*IRPEF*) to the Catholic Church, the State and confessions holding an *intesa*: they can participate by signing under 'one of the others' in the tax form. The entire fund (ie the overall amount of 0.008 of the *IRPEF*) will then be divided proportionally among the choices selected by the taxpayer who signed to give 0.008 of all taxes to specific institutions (eg the Catholic Church, the State, one of the minority religions holding an *intesa*). In doing so, even the taxpayers who do not choose any denomination will end up funding one according to the selection made by those who have signed to give their taxes to a religious group.

[53] F. Garelli, *Religion Italian Style* n 23 above, 240-257.

[54] F. Alicino, La legislazione sulla base di intesa. I test delle religioni "altre" e degli ateismi (Bari: Cacucci, 2013), 218.

[55] See the President of the Council of Ministers of the Italian Republic, 'Atto protocollato DAGL 1/2.5/4430/23 e comunicato all'UAAR con lettera datata 20 febbraio 1996'. See also Consiglio di Stato, Parere 29 ottobre 1997 no 3048.

resulted in a long legal battle, marked by several judicial decisions. Some of them has been issued by the administrative courts (the regional administrative tribunals and the Council of the State).[56] Others by ordinary judges, including the Italian Supreme Court (*Corte di Cassazione*). In 2013, this Court held that the original goal of the *intese* is to make the constitutional right of religious freedom better implemented, more widely valued, and equally enjoyed by all.[57] The Italian Supreme Court, however, also affirmed that through the phenomenon of copy&paste understandings, the instrument of the *intesa* has been transformed into a sort of legislative framework, which is accessible only for few minority religions at the exclusion of all other groups.[58]

Another major problem concerning *intese* is that there is no formal procedure of using Art 8.3 of the Constitution, which can turn the discretionary power of the Government into unreasonable discrimination towards some minority groups. For this reason, in 2013 the Italian Supreme Court also stated that the decision to initiate negotiations could not be left to the absolute discretion of the Government: negotiations should be considered as a corollary of the equal freedoms guaranteed to all religious faiths. It follows that the Government's refusal to launch the negotiations cannot be considered as a political act. The refusal should instead be qualified as a legal act that, as such, is subject to judicial review.[59]

The Italian Constitutional Court intervened in the case in a different way in 2016, adopting the opposite approach: *intese* are no longer bound to equal freedom of all beliefs before the law.[60] According to the ICC, the significance of the provision under Art 8.3 of the Constitution consists in the extension of the bilateral method from the Catholic Church to non-Catholic faiths. This is possible only where the method reflects the common intentions of both religious minorities and the Government not only to conclude an agreement, but also to initiate negotiations.[61] As far as the supreme principle of secularism is concerned, the ICC affirmed that this principle certainly implies impartiality and equidistance with regard to each religious faith. However, the Court also ruled that the conclusion of an *intesa* does not involve the right to profess religious belief. This right, they clarified, is protected overall by other constitutional rules,[62] starting with those guaranteeing the right to profess individually or together with other any religion or to profess no religion at all.[63] It means that, along with the method of bilateralism, the Government holds a broad margin of discretion, which implies the power of defining what religion is, as well as the responsibility of deciding

[56] TAR Lazio (Rome) 5 November-31 December 2008 no 12539, *Rassegna Avvocatura dello Stato*, 324 (2008). Consiglio di Stato 18 November 2011 no 6083, *Il Foro italiano*, III, 632 (2012).

[57] Corte di Cassazione-Sezioni unite 28 June 2013 no 16305, *Il Foro italiano*, I, 2432 (2013).

[58] ibid.

[59] ibid.

[60] Corte costituzionale 10 March 2016 no 52 n 43 above.

[61] ibid.

[62] In particular those of Arts 3, 8.1, 8.2, 19 and 20 of the Constitution.

[63] Corte costituzionale 10 March 2016 no 52, n 43 above.

whether to initiate a negotiation with any religious group.

In other words, in this field the Government can do whatever it wants.[64]

VII. The Right to Freedom of and from Religion

It is important to note that the Constitutional Court supported the 2016 decision by a significant *obiter dictum*, for which

> the changing and unpredictable reality of national and international political relations, which may lead the Government to conclude that it is not appropriate to allow an association that requests it to launch negotiations. When confronted with this considerable variety of situations, the Government is vested with a broad discretion.[65]

Strangely enough, this passage of the decision has little to do with the Italian atheism and more to do with the confessional organisations that would subscribe an *intesa* in the near future. The *obiter dictum* is indeed important not only for UAAR case law, but also for the entire system of State-confessions relationship in Italy. Moreover, this passage uncovers another important aspect of today's new pluralism in Italy. More specifically, the 2016 constitutional decision can be fully understood when considering the presence of new religious creeds, such as those made up of Muslim immigrants.

This reveals that, along with new forms of militant atheism, Islam(s) is now the most illustrative example of Italy's current cultural-religious diversity.[66] On the other hand, the supreme principle of secularism implies the right to freedom of (and from) religion of atheists, which includes the right to manifest nonreligion or disbelief, either alone or in a community with others, in public or private.[67]

With reference to this aspect, it should be noted that on 17 April 2020 the Italian Supreme Court issued an interesting decision,[68] which reversed previous judgement by the Court of Appeal for the district of Rome.

The Court of Appeal had prevented UAAR from using the atheist campaign aiming to run buses around some cities with a peculiar slogan, which crossed out the letter '*D*' from the Italian word *Dio* (God). Therefore, in the slogan the only visible letters were '*i*' and '*o*', meaning *io* (myself). In this manner, the slogan read:

'Ten million of Italians live very well without *D* (which implicitly means

[64] F. Alicino, 'La bilateralità pattizia Stato-confessioni dopo la sentenza n. 52/2016 della Corte costituzionale' *Osservatorio sulle fonti*, 2 (2016).

[65] Corte costituzionale 10 March 2016 no 52 n 43 above, para 5.2 conclusion on points of law (translation of the author).

[66] F. Garelli, *Religion Italian Style* n 23 above, 170.

[67] F. Alicino, 'The Italian legal system and imams. A difficult relationship', in M. Hashas et al eds, *Imams in Western Europe. Developments, Transformations, and Institutional Challenges* (Amsterdam: Amsterdam University Press, 2018), 359-380.

[68] Corte di Cassazione 17 April 2020 no 7893, *Il Foro italiano*, I, 1538 (2020).

without *Dio*-God). And when they are discriminated, UAAR is at their side'.

According to the Court of Appeal of Rome,[69] the real goal of the bus campaign was not to promote atheism, but to offend all religions denominations. On the contrary, the Supreme Court held that Arts 19 and 21 of the Constitution provide wide-ranging forms of propaganda of free thought, which includes atheistic critique of religion. The expressions that constitute offences towards religions are prohibited by the law, the apex Court said. However, this is possible when offences are clear, direct and very serious. It is not the case of the bus campaign. The Court of Appeal failed to strike a fair balance between the protection of the rights of religions and the right to freedom of expression. In other words, the Court of Appeal gave absolute primacy to protecting feelings of religious people, without adequately taking into account the UAAR's right to freedom of expression. Therefore, the Court of Appeal was not able to explain how and why the above-mentioned slogan do not aim at promoting atheism. At the same time, the Court did not clarify how and why the slogan denigrates the concept of God, offending believers of all religions.[70]

Thus, rather than being offensive, the UAAR's slogan was the expression of the rights to be equally free before the law, to communicate freely one's own thoughts in both everyday speech and writing, as well as the right to profess freely nonreligious belief.[71]

VIII. Conclusion

Before the recent wave of immigration and the current process of globalization, cultural-religious landscape of many Western democracies was pluralist, but with a number of groups having similar traditions. Today, pluralism

[69] Corte di Appello di Roma 23 March 2018 no 1869, unpublished.

[70] In this same vein see Eur. Court H.R., *Sekmadienis Ltd.* v *Lithuania*, App no 69317/14, Judgment of 30 January 2018.

[71] N. Colaianni, 'Propaganda ateistica: laicità e divieto di discriminazione' *Questione giustizia*, 10 June 2020, available at https://tinyurl.com/1vhc6mz4 (last visited 30 June 2021); M. Miele, 'La Cassazione e il «credo ateo o agnostico»' *La Nuova Giurisprudenza Civile Commentata*, II, 1133-1136 (2020); M. Croce, 'Opportune (e ovvie) precisazioni della Cassazione in tema di propaganda del non credere' 2 *Quaderni costituzionali*, 401-404 (2020); J. Pasquali Cerioli, ' "Senza D". La campagna Uaar tra libertà di propaganda e divieto di discriminazioni' *Stato, Chiese e pluralismo confessionale*, 4 May 2020, 50-56.

indicates the presence of people from very different cultures that, compared to the traditional ones, involve distinctive customs, peculiar value systems and unique practices. This trend is even more evident in legal systems with a history of religion-based influence, where the laws regulating some sensitive matters (blasphemy, proselytism, personal status, etc) and State-religions relationship still remain largely grounded on the needs and views of traditional confessions. That often clashes with a secularized attitude of atheists claiming equal treatment, freedom of expression, and freedom of religion, in both the individual and the collective sense of the terms.

The example is given by the Italian legal system, within which the interpretation of constitutional rules frequently promotes religious arguments for the implementation of the principle of secularism. For the same reasons, this explains the contrast between those who support the atheistic ideal of a 'true' secular democracy and those who sustain confessional viewpoint in governing today's socio-cultural landscape, which is also characterized by the emerging presence of new strong religious actors, like Islam(s). Thus, most of the current questions involving atheism are strictly related with at least two main factors: the historical roots of the system of State-Churches relationship; the presence of some different conspicuous forms of religious affiliation. These two factors make difficult the interpretation of the separation between religion and State, as requested by what the Italian Constitutional Court calls the supreme principle of secularism (*principio supremo di laicità*).[72] This principle remains, not by accident, largely undefined.[73]

For all these reasons, the study of the Italian atheism is extremely interesting. Even though they are a minority among minorities, atheists are able to challenge many intricate contradictions of the constitutional domain. They use the judiciary machine as precisely as possible in order, for example, to test the incongruities at the heart of the bilateralism (State-Churches) method and biased readings of the principle of secularism.

In particular, they bring these incongruities under the stricter control of both the national and supranational legal systems, which are informed by a multi-faceted conceptualisation of constitutional democracy.[74] According to religious nones,

[72] See para VII.

[73] F. Alicino, 'La libertà religiosa', in F. Buffa and M.G. Civinini eds, *La Corte di Strasburgo* (Roma: Questione giustizia, 2019), 458-467.

[74] See, for example, the following Eur. Court H.R.'s judgements, whose relative actions have been brought by UAAR: Eur. Court H.R., *Pellegrini* v *Italy* App no 30882/96, Judgment of 20 July 2001; Eur. Court H.R., *Lombardi Vallauri* v *Italy*, 20 October 2009; *Lautsi* v *Italy* App no 30814/06, Judgment of 3 November 2009; Eur. Court H.R. (GC), *Lautsi and Others* v *Italy* n 12 above. At the end of the day, one of the main ambitions of the Italian atheists is 'the concrete recognition of the supreme constitutional principle of secularism, especially with the reference of public schools and institutions, as well as the full equality before the law of all persons, regardless of their philosophical and religious beliefs'. In this perspective, the current form of Italy atheism calls for the abolition of every privilege or benefit granted, in law or in fact, to any religion' (art 3b of UAAR's Statute, which was approved during the national congress of 2 July 2006, translation mine).

this is the first step towards much more ambitious targets, such as thoroughly secular environment where there is no longer a need to be atheists, at least in the militant sense of the term.

The Protection of Choreographies Under Copyright Law: A Comparative Analysis

Andrea Borroni and Giovanna Carugno*

Abstract

The legal literature on intellectual property has rarely focused on choreographies. Choreographic works are different from other works protected under copyright law, because they consist in a limited number of standardised building blocks (musical notes, dance steps and movements) which are then each time arranged in an original, creative, and reproducible combination. The questions for lawyers are whether the combination of these elements is deemed worthy of protection by the legal domain in its entirety, or whether the musical part and the movement sequence can only find protection as separate components of the choreographic work; in either case, the question arises as to what are the thresholds for protection, and what remedies are available. In this paper, the author examines the legal issues related to choreographies through a comparative approach, considering concrete cases related to this matter as well as national legislation and international IP treaties.

I. Introduction

1. Research Question and Methodology

This article aims at providing an interdisciplinary contribution to the legal investigation of choreographic works, trying to offer a new comparative perspective. The protection of choreographies under copyright law will be analyzed through a diachronic and synchronic approach and the dialogue between legislation, legal case law, and legal scholarship will constitute a key-point to reveal the changing perspective in the protection of these works.

It will be underlined how the formants interact in order to frame the operational rules in the world of dancing, tackling the issues raised by choreographers and performers. In this respect, the literature relating to the history of dance and choreographic expression cannot be neglected.

Particular attention will be given to the United States' legal system, since it

* Andrea Borroni is Assistant Professor of Private Comparative Law, University of Campania 'Luigi Vanvitelli', 'Jean Monnet' Department of Political Sciences. Giovanna Carugno holds a PhD in Comparative Law and Processes of Integration from the University of Campania 'Luigi Vanvitelli', 'Jean Monnet' Department of Political Sciences. While this article is the product of common critical analysis and considerations, Giovanna Carugno authored paragraphs I.1.; I.2; II.1; II.2; II.3; II,4 whereas Andrea Borroni III.1; III.2; III.3; IV; V.

was the first one to adopt legal rules on choreographies; since then and until now, its example has served as an inspiring model for other States, even those belonging to the civil law family.

The comparison between the provisions and legal solutions offered by scholars and judges in different countries will be studied in light of the harmonizing role played by the international Treaties in the field of intellectual property, with a particular focus on the 1886 Berne Convention for the Protection of Literary and Artistic (the 'Berne Convention'). In conclusion, as it often happens in the legal domain, it will be highlighted that the traditional divide between common and civil law countries has becomes nuanced.

2. Brief Historical Notes

The legal literature has only recently focused on choreography. Until the middle of the 20th century, choreographies played a marginal role amongst the creative works considered worthy of copyright protection in *both* the *civil law and common law* traditions.[1]

This lack of protection, in terms of providing a legal status for choreographies and consequently, recognising exclusive rights to the choreographer, was due to many reasons.

Primarily, before the technological revolution, it was difficult to fix the choreography in a tangible support,[2] but later it became possible to record a dance

[1] Since the interest of the legislator was mainly focused on the protection of musical works, from the early 19th century, choreographers started to set limits to the use of their creations by concluding agreements with publishers, theatrical managers, dancers, and other artists (first among the others, the composers who wrote the musical scores for the choreography). This practice was especially developed in the Italian system, where historical sources show that the most common contract in the dance world was the *locatio operis*. As underlined by the scholars, this kind of agreement was primarily used to govern the relationship between the choreographer and the other protagonists of the theatrical arena. G. Azzaroni, *Del teatro e dintorni: una storia della legislazione e delle strutture teatrali in Italia nell'Ottocento* (Roma: Bulzoni, 1981), 90.

[2] Cf L.I. Mirrel, 'Legal Protection for Choreography' 27 *New York University Review,* 792 (1952). As correctly noted, another factor which *hindered* the legal protection of choreography was represented by the *specific nature* of this creative work; in fact, 'unlike literature or music, dances are intangible work of art that lives primarily through performance rather than through recording' (L.B. Cramer, 'Copyright Protection for Choreography: Can It Ever Be 'En Pointe?' Computerized Choreography or Amendment: Practical Problems of the 1976 U.S. Copyright Act and Choreography' 1 *Syracuse Journal of International Law and Commerce*, 145 (1955). The American anthropologist J.W. Keliinohomoku, 'An anthropologist looks at ballet as a form of ethnic dance', in A. Dils and A. Cooper eds, *Moving History/Dancing Cultures: A Dance History Reader* (Middletown: Wesleyan University Press, 2001), 38, emphasises this difference between dance and other forms of art, by identifying the performance movement as *the heart* of a dance work. He defined dance as 'a transient mode of expression performed in a given form and style by the human body moving in space. Dance occurs through purposefully selected and controlled rhythmic movements'. From this difference derives the difficulty in materially fixing the choreographic work: 'fixation in tangible form (…) presents a problem in the protection of choreography because movement is not susceptible of fixation as are other art forms (…) A choreographer's finished product is ephemeral, lasting only the length of the dancer performance. Music has similar

performance which is easier than writing.[3]

Since the Renaissance, notation was the most used method to give tangible form to a choreography, with the primary aim to keep it preserved and transferred to the performers.[4] Through the notational method, the dancing masters wrote symbols to indicate the position of the feet and the sequence of step. Until then, the 'recording' of choreographies heavily relied on the memory of the dancers;[5] this

> caused the loss of many great choreographic works when either the author died, or his memory failed without the works having been passed on to another by word of mouth and by demonstration.[6]

The practice of writing dances has continued to increase over the centuries[7] due to the growth of the dance community. This facilitated the recognition of authorship in relation to choreographic work, in the presence of a fixed score of the movement patterns.[8]

problems, but recording dance is much more difficult than recording music because dancers move in space as well as time'. M. Cook, 'Moving to a New Beat: Copyright Protection for Choreographic Works' 24 *UCLA Law Review*, 1294, 1287-1312 (1977).

[3] In effect, 'it is doubtful whether copyright protection would be afforded to a work which was not recorded in some tangible form but was merely performed' (J.E. Fitzgerald, 'Copyright and Choreography' 5 *CORD News*, 26, 25-42 (1973). Modern technologies give the possibility to incorporate a dance work in a material support, by recording it in a video, as a more flexible, less expensive and faster solution compared to the use of the notation system. For more details, see para 4 of this paper.

[4] For an introduction to the history of the dance notation system, see A.G. Hutchinson, *Dance Notation. The Process of Recording Movement on Paper* (London: Dance Books, 1984); Id, *Labanotation. The System of Analyzing and Recording Movement* (New York: Routledge, 1991).

[5] There were many notation systems developed from the mid-15th century, that differently combined letters, representative figures and symbols of the pathways. The first evidence of dance notation comes from Spain and consists in a manuscript which recorded a typical popular dance called 'low' or 'bass' dance, for the particular position of the feet, which were not to be lifted from the floor. In the opinion of M. Bourgat, *Technique de la Danse* (Paris: Presses Universitaires de France, 1986), 18, '*(é)crire la danse, c'est définir dans le temps et dans l'espace, par des lettres, des chiffres, et des signes appropriés, une succession d'attitudes du corps permettant la succession d'un thème dansant*'. Over the centuries, each choreographer has used his own method to write steps and body movements. Such a fragmentary approach has led to the development of many different ways and styles to notate dance, each of which was not fully shared among the choreographers or accepted by the entire dance community. This contributed to making dance as a form of art, 'unstable, depending on generations of dancers whose uncertain memories are associated with their own styles and body habits', F.E. Sparshott, *A Measured Pace, Toward a Philosophical Understanding of the Arts of Dance* (Toronto: University of Toronto Press, 1995), 199.

[6] G.D. Ordway, 'Choreography and Copyright' 15 *Copyright Law Symposium*, 174, 172-189 (1965).

[7] The milestones of dance notation were the handbooks written by R.-A. Feuillet, *Chorégraphie; ou, l'art de décrire la danse* (Paris: M. Brunet, 1700), A. Saint-Léon, *La Sténochorégraphie* (Paris: A. Saint-Léon, 1852), F.A. Zorn, *Grammatik der tanzkunst* (Leipzig: J.J. Weber,1887) and M. Morris, *Notation of movement* (London: Kegan Paul, Trench, Trubner & Co, 1928).

[8] F. Yeoh, 'The Value of Documenting Dance', available at https://tinyurl.com/sszrf5fe (last visited 30 June 2021), points out that the reason for recorded dance goes beyond the need to obtain a fixed product as empirical evidence to prove authorship. In fact, 'the value of documenting of

II. Building a Legal Framework

1. The United States Experience

In 1952, the US copyright office admitted for the first time the registration of a choreography, *Kiss me Kate* by Hania Holm, written in the well-known Labanotation.[9]

This case opened the doors to the recognition of the copyrightability of a dance not only in the American legal system, but also in other countries around the globe.[10]

Thus, choreographic works, along with pantomimes, were explicitly included as one of the categories of copyrightable subject-matter in the United States.

This was in response to the needs perceived by legal scholars and choreographers to give protection to those creations of the mind. Before then, choreographies were not deemed worthy of protection *per se*, but could rather be included in the larger set of dramatic compositions protected since the Act to Amend the Several Acts Respecting Copyrights, 3 February 1831 (the '1831 Copyright Act') and defined as including

> all manner of compositions in which the story is represented by dialogue or action instead of narrative, and a scene or composition in which the author's ideas are conveyed by action alone, is within the term.[11]

Though, all dances could not be included, but only those that met the standard laid out in the case *Fuller* v *Bemis*, ie those dances that told a story, portrayed a character or depicted an emotion.[12]

dance works in its various manifestations is evident not only for the purposes of copyright, but for preservation and scholarship. Developments in recording process that will make the art form of dance more accessible will only enhance its status'.

[9] Labanotation is a system of notation invented by the Austro-Hungarian choreographer Rudolf Laban (1879-1958) that permits to record not only the position of the feet, but also every human gesture and motion. This system is described as follow: 'Labanotation involves a staff that is divided vertically by a center line to represent the two sides of the body. The staff is divided further into two to twelve vertical columns. The complex symbols in these columns of the staff represent the positions of all parts of the body at a given point in space and time. The center line represents the spine and the right and left lines correspond to the right and left sides of the body. The staff, which is read bottom to top, contains symbols which convey specific movements. The length of these symbols signifies the length of time allotted for that movement'. A.K. Weinhardt, 'Copyright Infringement of Choreography: The Legal Aspects of Fixation' 13 *Journal of Corporation Law*, 839, 836-891 (1988).

[10] As L. Wilder, 'U.S. Government Grants First Dance Copyright' 19 *Dance Observer*, 69 (1952) underlined, thanks to Holm's claim, 'the battle of choreographers for legal recognition and protection passed into history. From now on, dance works are to be considered artistic property and must be protected as such'.

[11] *Daly* v *Palmer*, 6 Fed. Cas. 1132, 286 (1868).

[12] *Fuller* v *Bemis*, 50 Fed. 926 (1892). In that judgment, the court held that '(a)n examination of the description of complainant's dance, as filed for copyright, shows that the end sought for and accomplished was solely the devising of a series of graceful movements, combined with an attractive

In Fuller, the plaintiff's famous serpentine dance[13] failed the test and was ineligible for copyright protection for its non-figurative character.

In fact, before Public Law no 94-553, 90 Stat. 2541, 19 October 1976 (the '1976 Copyright Act'), choreographies were conceived as dramatic compositions, a large category of works protected under Section 5 of Public Law 60-349, 35 Stat. 1075, (the '1909 Copyright Act').

The dramatic component was presented only in a dance that

> tells a story, develops a character, or expresses a theme or emotion by means of specific dance movements and physical action.[14]

This component should have been recognised by the administrative authority and, lastly, by the courts. However,

> although the copyright office will allow registration of certain choreographic works, this does not guarantee that the courts will enforce protection against unauthorized use of such works.[15]

In other words, the dramatic character of the dance was ascertained on a case-by-case ground; this clearly represented a glitch in the copyright law system and limited the effective defense of the rights of the choreographers. Therefore, the creation should have passed a preliminary 'copyrightability test', based on the presence (or absence) of the dramatization elements.

Following the enactment of the 1976 Copyright Act, however, the situation changed, and choreographies could be protected *suo jure*. Various scholars had,

arrangement of drapery, lights, and shadows, telling no story, portraying no character, depicting no emotion, adding that (s)urely, those (movements) described and practiced here convey and were devised to convey, to the spectator, no other idea than that a comely woman is illustrating the poetry of motion in a singularly graceful fashion. Such ail idea may be pleasing, but it can hardly be called dramatic'. On the topic, see also *Copyright Law Revision: Studies Prepared for the Subcommittee on Patents, Trademarks, and Copyrights of the Committee on the Judiciary, United States Senate, Eighty-Sixth Congress, Second Session Pursuant to S. Res. 240, studies 26-28* (Washington: US Government Printing Office, 1960-1961).

[13] 'In that dance, Fuller, wrapped in veils, moved as if under hypnosis (…) and became a flower, a butterfly, or a flame, thanks to the use of sticks that extended the movements of her arms and of colored and illuminated silk fabric'. F. Rosso, *Cinema e Danza. Storia di Un Passo a Due* (Torino: UTET, 2008), 11. Kraut underlined the relevance of the Fuller case '(a)s an early attempt by a white woman to use the legal system to secure ownership of a choreographic work (…). Viewed from this perspective, the lawsuit offers a case study of a white, female, early modern dancer's endeavors to harness the racial privileges of whiteness and establish herself as a property-holding subject. Accordingly, this essay approaches the circulation of the Serpentine Dance and Fuller's lawsuit against Bemis as the story of a gendered struggle to attain proprietary rights in whiteness'. A. Kraut, 'White Womanhood, Property Rights, and the Campaign for Choreographic Copyright: Loïe Fuller's Serpentine Dance' 43 *Dance Research Journal*, 4, 3-26 (1975).

[14] Copyright Office, Circular 41: Choreographic Works 1 (April 1977); accord, 37 C.F.R. 202.7 (1976).

[15] G.D. Ordway, n 6 above, 178.

after all, been arguing in favour of such an approach. In 1959, for instance, Martin had observed that '(t)he choreographic field cannot by any possible manipulation be forced into the category of dramatic works'.[16]

In the same way, Chujoy emphasised that considering choreographies as a type of dramatic composition was anachronistic and added that

> (t)he problem of storytelling or dramatic qualifications of a dance work submitted for copyright is a serious one, albeit antiquated. A quarter of a century ago a ballet without a story was an exception, today it is quite often the prevailing fare eg, most ballets in the repertoire of the New York City Ballet.[17]

In order to extend the scope of copyright on choreographies to those dances that cannot be reduced in a dramatic form – as the case of the so-called 'abstract dance, in which, aside from their esthetic appeal, no story or specific theme is readily apparent'[18] – choreographies and dramatic works are named separately in para 102(a) of the 1976 Copyright Act:

> (w)orks of authorship include the following categories: (1) literary works; (2) musical works, including any accompanying words; (3) dramatic works, including any accompanying music; (4) pantomimes and choreographic works; (5) pictorial, graphic, and sculptural works; (6) motion pictures and other audio-visual works; (7) sound recordings; and (8) architectural works.

The necessity to reform copyright legislation was pointed out also by the President Franklin Delano Roosevelt, who remarked that the drafters of the 1909 Copyright Act

> are imperfect in definition, confused and inconsistent in expression; they omit provision for many articles which, under modern reproductive processes, are entitled to protection; they impose hardships upon the copyright proprietor which are not essential to the fair protection of the public; they

[16] *Copyright Law Revision* n 12 above, 111.

[17] ibid 115.

[18] M.B. Nimmer, *Nimmer on Copyright* (New York: Bender, 2002), 2. According G.D. Ordway, n 6 above, 181, 'even though some choreography may qualify as dramatic composition, it is equally obvious from the foregoing that not all dance is embraced within that concept. The traditional ballets which are commonly noted for conveying a storyline would obviously qualify. It is in the area of the 'modern and abstract' dances, where the dramatic content is questionable, that the real problem lies'. Before the entry into force of the 1976 Copyright Act, '(t)o secure and retain statutory copyright, one must register his work in one of the registration classes set out in §5 of the Act. But since choreographic works are not mentioned in §5, to establish eligibility for statutory copyright, a choreographer must convince the Copyright Office (and possibly the courts) that his dance composition fits in one of the classes that is mentioned. This often results in attempting to register choreographic works in Class D – dramatic or dramatico-musical compositions'. J.I. Roth, 'Common Law Protection for Choreographic Works' 5 *Performing Arts Review*, 75, (1974).

> are difficult for the courts to interpret and impossible for the Copyright Office to administer with satisfaction to the public. Attempts to improve them by amendment have been frequent, no less than twelve acts for the purpose having been passed since the Revised Statutes. To perfect them by further amendment seems impractical. A complete revision of them is essential.[19]

It is important to point out that, despite the inclusion of choreographies amongst the protected works of authorship within the US system, the 1976 Copyright Act does not clarify the characteristics of a choreography, or of a pantomime, as to how it differs from a dramatic work. The House Report points to the fact that it was a deliberate choice, for choreography and pantomime to 'have fairly settled meanings'.[20]

This lack of a definition allows the courts to consider worthy of protection a great number of diverse types and styles of dance. Indeed,

> if Congress were to stipulate a narrow, precise definition in the legislation, according to today's understanding of dance, it would restrict future choreographers from copyright protection for developments in dance that cannot be foreseen today.[21]

Such an approach is therefore commendable, for it allows for an evolution of the concept of choreography, giving the chance to

> (d)ance critics, theorists, philosophers, and historians (to play) a continuing role in this dialogue as we broaden and improve our understanding of 'dance'. Imposing a narrow codification in the Copyright Law would curtail this process unnecessarily.[22]

A definition is, instead, contained in the Copyright Office Practices, Compendium II (1984), where a choreography is defined as

> the composition and arrangement of dance movements and patterns, and is usually intended to be accompanied by music. Dance is static and kinetic successions of bodily movement in certain rhythmic and spatial relationships. Choreographic works need not to tell a story in order to be protected by copyright.[23]

[19] A. Latman et al, Copyright for the Eighties: Cases and Materials (*Charlottesville*: The *Michie* Company, 1985), 7.

[20] H.R. Rep. no 1476, 94th Cong., 2d Sess. (1976), 53.

[21] ibid.

[22] ibid.

[23] Copyright Law Reporter, 1991, no 625.

2. The US Influence on the Categorization of Choreographies as Dramatic Works

The protection of choreographic works in common law countries took inspiration from the US model, that classified this kind of creations as dramatic works.

In the United Kingdom, choreographies are mentioned by Section 35(1) of the Copyright Act 1909, under which

> (d)ramatic work includes any piece for recitation, choreographic work or entertainment in dumb show, the scenic arrangement or acting form of which is fixed in writing or otherwise, and any cinematograph production where the arrangement or acting form or the combination of incidents represented give the work an original character.

A similar provision is contained in section 3(1) of the Copyright Act 1988 (R.S.C., 1985, c. C-42), which mentioned choreographies in the list of dramatic works as well.

The same approach is taken in Canada, where choreographies have been considered a type of dramatic work since the Copyright Act, 1921. Although, it was only in 1988 that choreographies were given a statutory definition as a result of the Copyright Act 1988.

And that is still the case in India, where choreographies are not autonomously protected, but rather as dramatic works, pursuant to Section 2(h) of the Act to amend and consolidate the law relating to copyright, 4 June 1957 (the 'Indian Copyright Act 1957'), under which dramatic work

> includes any piece for recitation, choreographic work or entertainment in dumb show, the scenic arrangement or acting form of which is fixed in writing or otherwise but does not include a cinematograph film.[24]

Therefore, since the Indian Copyright Act 1957 does not explicitly deal with choreographies, in order for them to be protected, they must meet the same requirements as all other dramatic works. Namely, they must be original and fixed in a tangible medium of expression, so as not to be fleeting.[25]

However, to be eligible for protection, a choreography must feature 'dramatic

[24] See also Academy of General Education, *Manipal* v *B. Malini Malia*, AIR 2009 SC 1982, where the Supreme Court confirmed that '(k)eeping in view the statutory provisions, there cannot be any doubt whatsoever that copyright in respect of performance of dance would (...) come within the purview of the definition of dramatic work'.

[25] U. Srivastava, 'So You Think You Can (Copyright) Dance? An Analysis of the Copyrightability of Choreographic Works in India' 12 *Journal of Intellectual Property Law and Practice*, 43 (2017). The very same approach is also adopted in South Africa, where Section 1 of Act no 98/1978, clarifies that dramatic work 'includes a choreographic work or entertainment in dumb show, if reduced to the material form in which the work or entertainment is to be presented'.

action', a requirement frequently criticised; Vaver, for instance, wrote that

> (o)nly a farfetched interpretation of the old Act could produce the result it claimed. Choreography, included as a species of "dramatic work", may take some colour from its genus, but obviously extends to other things than Othello on point.
>
> Two other major genera in the Indian Copyright Act 1957, literary and artistic works, also non-exhaustively list a number of miscellaneous species in their definitions, but do not require them to have all the characteristics of the genus.[26]

Considering choreographies as literary works could be a useful way to overcome the limitations deriving from the idea of dramatization. Under Section 2 of the Copyright Act 1988, literary works are 'other than a dramatic and musical work, which is written, spoken or sung'; the article then states that dances are included into the category of dramatic works. In this case, choreography can meet the requirements of a literary work only if it is conceived as a narrative product, 'as a book or article (…) – for example, a history of the dance or a critical appraisal of a particular dance or style of dancing', unlike the case where a choreography is written down in a descriptive way, with the aim of fixing it in a tangible form and enabling dancers to perform it.[27]

3. The Civil Law Classification of Choreographic Works

In the civil law family, choreographic works possess an independent status from dramatic compositions. Under French law, a clear-cut distinction is drawn between choreographies and dramatic works. In particular, Art 3 *Loi no 57-298 du 11 mars 1957 sur la propriété littéraire et artistique* (the '1957 Copyright Act'), states that

> *(s)ont considérés notamment comme des œuvres de l'esprit au sens de la présente loi: (...) les œuvres dramatiques ou dramatico-musicales; les œuvres chorégraphiques, les numéros et tours de cirques et les pantomimes dont la mise en œuvres est fixée par écrit ou autrement.*

[26] D. Vaver, 'The Canadian Copyright Amendments of 1988' 4 *Intellectual Property Journal*, 144-145, 121-155 (1989). Cf A.G. DeMille, who writes that '(c)horeography is neither drama nor storytelling. It is a separate art. It is an arrangement in time-space, using human bodies as its unit of design. It may or may not be dramatic or tell a story. In the same way that some music tells a story, or fits a 'program', some dances tell stories-but the greater part of music does not, and the greater part of dancing does not' (*Copyright Law Revision: Studies Prepared for the Subcommittee on Patents, Trademarks, and Copyrights of the Committee on the Judiciary, United States Senate, Eighty-Sixth Congress, Second Session Pursuant to S. Res. 240, Studies 26-28* (Washington: US Government Printing Office, 1960-1961)).

[27] *Copyright Law Revision* n 12 above, 97.

In addition, Section 112-2 of the 1992 *Code de la propriété intellectuelle* clarifies that

> *(s)ont considérés notamment comme des œuvres de l'esprit au sens du présent Code: (…) 3° Les œuvres dramatiques ou dramatico-musicales; 4° Les œuvres chorégraphiques, les numéros et tours de cirques, les pantomimes, dont la mise en œuvres est fixée par écrit ou autrement*.[28]

In the Italian law, under Art 2 para 3 of legge 22 April 1941 no 633, *Protezione del diritto d'autore e di altri diritti connessi al suo esercizio* ('Law no 633/1941') both choreographies and pantomimes constitute a separate and autonomous category of works, which need to be fixed in a tangible support to be eligible for protection.[29]

Instead, in Germany, the legislator provides protection to the works of the art of dance within the category of pantomimes ('*pantomimische Werke einschließlich der Werke der Tanzkunst*')[30] and the choreographies are not explicitly mentioned.

In Austria, choreographies fall only within domain of literary works. This is self-evident by reading Art 2 para 2 of the Austrian Copyright Act, under which choreographies (and pantomimes) are theatrical works, 'expressed by gestures or other movements of the body' and described in a written form.[31]

Choreography is incorporated, for copyright purposes, in a literary work, when it is described by means of words and/or symbols. Registering choreography as a theatrical work – a sub-category of the literary ones – is the way offered not only by the Austrian Copyright Act, but also by other legal provisions in force in some extra-European countries[32] to protect this kind of creation. The same

[28] These works are performed by the artist through movement, gestures, and steps; in light of this, they were described by the scholars as '*œuvres gestuellesm*'. C. Caron, *Droit d'auteur et droits voisins* (Paris: Lexis Nexis, 2006), 138.

[29] On the difference between pantomime and choreography see M. Fabiani, *Diritto d'Autore e Diritti degli Artisti Interpreti o Esecutori* (Milano: Giuffrè, 2004), 64, who underlines that both, however, have in common the element of movement of the body; the imitation and the expression of eyes and arms are, according to the literature, the distinctive elements of pantomime when confronted to the choreography. See also M. Pasi, *Danza e Balletto* (Milano: Jaca Book, 1993), 107, quoting an excerpt of 'Dissertazione programmatica' from the ballet *Don Juan* by Gasparo Angiolini (1761), and V. Buonsignori, *Precetti sull'arte mimica applicabili alla coreografia ed alla drammatica divisi in quattro lezioni teoriche* (Siena: Tipografia dell'Ancora di G. Landi e N. Alessandri, 1854).

[30] See Art 2, para 1, no 3, of the 1965 Copyright Act (*Urheberrechtsgesetz Gesetz über Urheberrecht und verwandte Schutzrechte (Urheberrechtsgesetz*), enacted on 9 September 1965, and lastly modified on 4 April 2016).

[31] Federal Law on Copyright in Works of Literature and Arts and on Related Rights (*Bundesgesetz über das Urheberrecht an Werken der Literatur und der Kunst und über verwandte Schutzrechte*), in the last version published in the Federal Gazette I, no 99/2015.

[32] For instance, in South Korea, where the Supreme Court assessed the possibility to consider choreographies as independent works (Decision E 639, 10 April 2004), but in its judgement opted for a literal application of the law. In fact, Art 4 para 1 of Law no 3916 of 31 December 1986 stated that these creations are protected under the category of theatrical works, along with pantomimes,

attempt was made in common-law countries, but without a positive outcome. In fact, in few cases literary property rights were recognized as applicable to a choreographic work, especially in the past, when dances were incorporated through notational systems by using written instructions and verbal illustration of the movements assigned to the dancers.[33]

Choreographies can be registered as non-dramatic literary work, but this does not ensure protection against the risk of an i*nfringement* on the rights of the author regarding the dramatic representation of his own creations.

Like other literary works, choreography would be protected in its description, as a non-dramatic expression, but not for any potential (dramatic) performance of it.[34]

4. The Berne Convention and the Requirement of Fixation

Looking at international instruments, on the other hand, choreographies are not mentioned amongst the creations of the mind deemed worthy of protection in the original version of the Berne Convention, the first treaty dealing with copyright protection. Art 3 of the Berne Convention defines the expression 'literary and artistic works' as including

> books, pamphlets, and all other writings; dramatic or dramatico-musical works, musical compositions with or without words; works of drawing, painting, sculpture and engraving; lithographs, illustrations, geographical charts; plans, sketches, and plastic works relative to geography, topography, architecture, or science in general.

The choice not to include choreographies in the set of copyrightable works was because it would be difficult to define the characteristic of this specific creation of the mind, which combines music and movements, and to distinguish it from others, such as pantomimes.

In particular, Germany was against the inclusion, whereas Italy was in favour

dramas and other (unlisted) operas. G. Choi, 'A Study on Copyright Protection of Choreographic Works' 64 *Law Journal*, 204, 203-234 (2019).

[33] This was possible first because choreography, as mentioned above, was not deemed worthy of inclusion as an autonomous work (neither dramatic nor literary) in the US until the 1976 Copyright Act's revision. Moreover, the category of literary works is too broad to contain many different creative expressions, including choreographies. As correctly underlined by a commentator reassessing the role of Art 17 of the US Copyright Act, para 101, 'literary work is one expressed in words, numbers, or other verbal and numeric symbols or indicia, regardless of the nature of the material objects (…) in which they are embodied' and of their artistic merit or aesthetic value. D.E. Bouchoux, *Intellectual Property: The Law of Trademarks, Copyrights, Patents, and Trade Secrets* (Boston: Cengage Learning, 2000), 199.

[34] See in particular, the case of the choreography titled *Beethoven Sonata*, recorded by the choreographer Ruth Page as a book in 1953 and not as a dramatic work. A. Chujoy, 'New Try to Copyright Choreography' 22 *Dance News*, 4 (1953).

of protecting choreographies as an autonomous category of artistic work.[35]

In 1908, therefore, the Berne Convention was amended *(Berlin revision), and as a result, choreographies were included in the list of* literary and artistic works protected under copyright law, as long as they are fixed in a tangible support. In fact, Art 2, para 1, states that:

> the expression 'literary and artistic works' shall include any production in the literary, scientific or artistic domain, whatever may be the mode or form of its reproduction such as (…) dramatic or dramatico-musical works, choreographic works and entertainments in dumb show, the acting form of which is fixed in writing or otherwise.

The expression 'or otherwise' was inserted as a compromise between the German and the Italian positions to ensure that copyright protection was afforded to the largest extent possible.[36]

It is a common prerequisite for both the common and civil law systems that choreographies be fixed in a tangible support for them to be protected under the copyright law.[37]

As is commonly known after all, dance involves a sequence of many different movements which, to the layman, may not be immediately distinguishable. This can create difficulties when it comes to the determination of whether there was an infringement.[38]

[35] D. Howland, 'The International Movement to Protect Literary and Artistic Property', in Id et al eds, *Art and Sovereignty in Global Politics* (New York: Palgrave Macmillan, 2017), 69. Italy was in favour because it had a long tradition of choreographers who had been creating on commission many choreographies, since the end of the eighteenth century, and requesting that their authorship be recognized on the *libretto* with a specific phrasing. The first-time choreographies were autonomously mentioned among the creations of the mind worthy of protection was the Regio decreto legge 7 November 1925 no 1950 which at Art 1 specifically includes choreographies in the list of the artistic works protected by this decree.

[36] See 'Études générales – La convention de Berne révisée du 13 novembre 1908' *Droit d'Auteur*, 78 (1909).

[37] As far as the United States are concerned, two specific legal documents deal with this issue. See, the US Copyright Act 1976 requires choreographic works to be 'fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device' (title 17(b), Copyright Act). And, the US Copyright Office, Circular no 51b: 'The particular movements and physical actions of which the dance consists must be fixed in some sort of legible written form, such as detailed verbal description, dance notation pictorial or graphic diagrams, or a combination of these (…) Even a textual description of a dance would not seem to constitute (…) a work of choreography if the description is so general and lacking in detail that the dance could not be performed', cf B. Häger, *The Dancer's World: Problems of Today and Tomorrow* (New York: International Dance Council and UNESCO, 1978), 96-97.

[38] On the issue of copyright infringement in relation to choreographies in the US, the leading case is *Horgan* v *MacMillan Inc.* In that case, Macmillan Inc. had published a book titled *The Nutcracker*, containing photographs of George Balanchine's copyrighted choreography of the famous ballet. Balanchine's estate had then sued the publisher for copyright infringement. The US District Court for the Southern District of New York, however, concluded no infringement had

For this reason, the requirement of fixation was first established by the Berne Convention, since it is evident that the tangible support represents the most immediate way to prove authorship and copyright protection starts from the moment of fixation which, conventionally, is deemed to coincide with the moment of creation.[39]

It is clear that when the dance steps are fixed in writing, through symbols or words, it does not automatically entail that the choreography is a literary work. The classification of the work depends on the choice made by the choreographer, whether an alternative is provided by national law to register it as a literary work or as a choreography *ex se*.

In any case, the act of fixating a choreography in tangible support has a practical use because it makes it possible to ascertain if infringement occurs and gives certainty to the act of creating the choreography.[40]

Scholars have long written on the requirement of fixation. Most of the scholars, on that issue, affirm that fixing constitutes a condition for the existence of the choreographic work.[41] In particular, there is also who deems mandatory for the author to fix all aspects of choreography, beside the plot, traces, and the screenplay including also the specification of all constitutive elements, such as dance movements, plastic, and figurative figures, colors of the costumes, scenarios, etc.[42]

Other scholars underlined the practical difficulties related to the fixing

occurred, 'because (t)he still photographs in the Nutcracker book, numerous though they are, catch dancers in various attitudes at specific instants of time; they do not, nor do they intend to, take or use the underlying choreography. The staged performance could not be recreated from them' (*Horgan* v *MacMillan, Inc*., 621 F. Supp. 1169 (1985).

The Second Circuit Court of Appeals overruled the judgment, finding in favour of the plaintiff, for, they held, 'the standard for determining copyright infringement is not whether the original could be recreated from the allegedly infringing copy, but whether the latter is 'substantially similar' to the former, confirming that the proper test consists in determining whether the ordinary observer, unless he set out to detect the disparities, would be disposed to overlook them, and regard their aesthetic appeal as the same, and adding that (e)ven a small amount of the original, if it is qualitatively significant, may be sufficient to be an infringement, although the full original could not be recreated from the excerpt' (*Horgan* v *MacMillan, Inc*., 789 F.2d 157, 2d. Cir. (1986)). On the case, see J. Hilgard, 'Can choreography and copyright waltz together in the wake of Horgan v Macmillan, Inc.?' 19 *UC Davis Law Review*, 757-789 (1994); P.S. Gennerich, 'One Moment in Time: The Second Circuit Ponders Choreographic Photography as a Copyright Infringement Horgan v Macmillan, Inc.' 53 *Book Law Review*, 379-407 (1987).

[39] 'Regardless of the number of times a dance has been publicly performed, a choreographic work is created when it is fixed in a copy for the first time'. K. Abitabile and J. Picerno, 'Dance and the Choreographer's Dilemma: A Legal and Cultural Perspective on Copyright Protection for Choreographic Works' 27 *Campbell Law Review*, 44, 39-62 (2004).

[40] It is noteworthy that '(f)ixation in express detail is also beneficial in proving that an infringer 'copied' from the original work as opposed to creating the work itself. The unlikely similarity of specific movements and details cuts against the possibility that two choreographers independently created the movements' (M. Cook, n 2 above, 1296).

[41] S. Ercolani, *Il Diritto d'Autore e i Diritti Connessi: la Legge n. 633/1941 dopo l'Attuazione della Direttiva n. 2001/29/CE* (Torino: Giappichelli, 2004), 105.

[42] P. Greco, I Diritti sui Beni Immateriali: Ditta, Marchi, Opere dell'Ingegno, Invenzioni Industriali (Torino: Giappichelli, 1948), 170.

procedure in works that, for their own nature, live through performance, as product of a 'dynamic art', like dance.[43] It was highlighted that a choreography 'vanishes promptly upon performance' and remains 'impermanent'.[44]

Thus,

> (f)or working choreographers more interested in copyright protection for economic control than for preservation of the art form, written notation is an unattractive option.[45]

One way to fix the work without missing the performative features of the dance is the recourse to audio-visual devices, *in lieu* of writings.

Consequently, some specialists reminded that fixation could take the form of photographic, cinematographic, and television recording of the work. In doing so, the criticalities related to the transposition of what is performed in writings are reduced and the fixation, as far as it is possible, more authentic.

In this regard, it is undeniable that technological instruments facilitate the making and circulation of unauthorized copies of the choreographic work. As it was written,

> performance theory, which describes the development of individual agency through physical "embodiment" in the cultural worlds (…) has important lessons for crafting limits on property rights in experience, especially in cyberspace, where embodiment is the primary mode of experience and play;

in this area, 'dancing online' becomes 'a commodity, to the tune of literally billions of dollars'.[46]

For this reason, the battle of choreographers to protect their works is not over and copyright law should be ready to reshape and tune its solutions in relation to the digital advancement.

[43] The divide between 'dynamic' and 'permanent' has been clearly highlighted in the literature. If the former is 'is unstable or ephemeral, and that may invite unpredictable change though the influence of natural or human forces', the latter is an art 'that has and is meant to have weak, unclear boundaries – art that blurs text and context'. R. Brauneis, 'How Much Should Being Accommodate Becoming? Copyright in Dynamic and Permeable Art' 43 *Columbia Journal of Law and the Art*, 381 (2019).

[44] J. Taubman, 'Choreography Under Copyright Revision: The Square Peg in The Round Hole Unpegged' 10 *Performing Arts Review*, 241, 219-256 (1980). According to Anthea Kraut, '(t)he irony for dance is that copyright, with its requirement that works be 'fixed in a tangible medium of expression', has represented the temporal solidity – the past and the future – which is supposedly lacks; choreographic copyright is not an 'apparatus of capture". A. Kraut, *Choreographing copyright: race, gender, and intellectual property in American dance* (Oxford: Oxford University Press, 2016), 232.

[45] J.M. Lakes, 'A pas de deux for Choreography and Copyright' 80 *New York University Law Review*, 1854, 1829-1861 (2005).

[46] A. Chander and M. Sunder, 'Dancing on the Grave of Copyright?' 18 *Duke Law & Technology Review*, 149, 143-161 (2019).

Therefore, nowadays, the requirement of fixation appears outdated and should be totally re-thought.[47]

III. The Originality of a Choreographic Work

1. Premise

For a choreography to be eligible for protection, the requirement of originality must be met.

This element is present both in common law and civil law systems.

For a work to be deemed original a two-pronged test is applied: (i) the novelty and (ii) a minimum level of creativity.

In the US, according to the leading precedent,

> original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity.
>
> To be sure, the requisite level of creativity is extremely low; even a slight

[47] The criticism towards the requirement of fixation increased under the influence of some lawsuits that revealed the dangers of incorporating a sequence of movements through audio-visual recordings. In a leading case, some artists (namely, the singer and hip-hop performer Terrance Ferguson, with the actors Alfonso Ribeiro and Russell Horning) sued the company Epic Games, to the district court of the Central District of California, in December 2018. The plaintiffs alleged that the movements performed by a digital character in the videogame named Fortnite (distributed by the respondent), were based on their choreographies and, so, it ended up being copied and used without any kind of permission. As '(t)he suits seek to block Epic Games from using the dance moves, awards of money earned off the moves purchased in Fortnite, punitive damages and attorney's fees' to restore the moral and economic damages suffered by the plaintiffs. Z. Crane, 'Fortnite Is "Dropping" Into Legal Land: A Proposal to Amend the Copyright Act to Address Artists' and Game Developers' Concerns Over Dance Moves as Purchasable Emotes in Video Games', 6, available at https://tinyurl.com/6zn93nwz (last visited 30 June 2021). See the text of the judgments: *Ferguson* v *Epic Games*, No. 2:18-cv-10110 (C.D. Cal. 2018), *Ribeiro* v *Epic Games, Inc.*, No. 2:18-cv-10412 (C.D. Cal. 2018), and *Redd* v *Epic Games, Inc.*, (C.D. Cal. Dec. 17, 2018). Cf also E. Hack, 'Milly Rocking through Copyright Law: Why the Law Should Expand to Recognize Dance Moves as a Protected Category', 88 *University of Cincinnati Law Review*, 651, 637-651 (2019), who argued that the decision of the court to exclude the copyrightability of hip-hop dance movements invented by Terrance Ferguson, interpreting them as simple routines *per se* not worthy of protection, 'encourages the intellectual theft' and disincentives the creative activities of choreographers and dancers. To solve this crucial issue, taking into consideration the technological and artistic evolution, the scholar suggests amending the 1976 Copyright Act, enlarging the scope of the notion of choreographies up to the inclusion of hip-hop sequences. The same conclusion is drawn also by A. Chander and M. Sunder, 'The Romance of the Public Domain' 92(5) *California Law Review*, 1331 (2004).

Even if, as it was stated, 'if too much material is protected, choreographers will lack incentives to create new pieces as a result of a shrinking public domain, and there will consequently be fewer jobs for dancers'. K.M. Benton, 'Can Copyright Law Perform the Perfect Fouetté? Keeping Law and Choreography on Balance to Achieve the Purposes of the Copyright Clause' 36 *Pepperdine Law Review*, 114, 59-128 (2008).

> amount will suffice. Most of the works make the grade quite easily, as they possess some creative spark, 'no matter how crude, humble or obvious' it might be.[48]

Similar requirements are also applicable in Italy.[49]

The same holds true for France as well where the courts have phrased the requirement as '*l'empreinte de la personnalité de l'auteur*'[50] o '*l'empreinte du talent créateur personnel*',[51] and for Germany as well where it is defined as *individualität*.

As stated in recital 17 of the preamble to Directive 93/98 'an intellectual creation is an author's own if it reflects the author's personality'.[52] The European Court of Justice in the Case C-*145/10 Eva-Maria Painer* v *Standard VerlagsGmbH and others* stated that a creation is original also when its author is 'able to express his creative abilities in the production of the work by making free and creative choices'.[53]

[48] Feist Publ'ns, *Inc.* v *Rural Tel. Serv. Co.*, 499 US 340 (1991).

[49] See, among others, V.M. De Sanctis, *I Soggetti del Diritto d'Autore* (Milano: Giuffrè, 2005), 172, who clearly states that, to meet this standard, the choreographer needs to add something recognizably *per se,* by expressing his creative ability in an original manner and transferring his personal footprint to the work. It is required, in other words, to point out a personal contribution to the work in order to have the direct paternity without any mediation of preexisting work. On the topic see also the following judgments: Corte d'Appello di Milano 8 July 1988 and, previously, Tribunale di Roma 12 May 1951, *Il Foro Italiano*, I, 1425 (1951) expressing the necessity to recognise in the work of creation feelings of the author. Cf also P. Zatti and G. Alpa, *La Nuova Giurisprudenza Civile Commentata*, 795 (1989).

[50] Cour d'appel de Paris 21 November 1994, in *Revue Internationale du Droit d'Auteur*, 381 and 243, (1995).

[51] Cour de cassation 13 November1973, *Dalloz*, 533, (1974). See also A. Lucas and H-J. Lucas, *Traité de la Propriété Littéraire et Artistique* (Paris: Litec, 2001), 72-87.

[52] Council Directive 93/98/EEC of 29 October 1993 harmonising the term of protection of copyright and certain related rights, [1993] OJ L290, Recital, 17.

[53] Case C-*145/10 Eva-Maria Painer* v *Standard VerlagsGmbH and others*, Judgment of the Court of 1 December 2011, para 88, available at www.eurlex.europa.eu. See also Case C-5/08 *Infopaq International A/S* v *Danske Dagblades Forening*, Judgment of 16 July 2009, [para 33 et seq and, in particular, paras 37-38, available at www.eurlex.europa.eu: 'In those circumstances, copyright within the meaning of Art 2(a) of Directive 2001/29 is liable to apply only in relation to a subject-matter which is original in the sense that it is its author's own intellectual creation. As regards the parts of a work, it should be borne in mind that there is nothing in Directive 2001/29 or any other relevant directive indicating that those parts are to be treated any differently from the work as a whole. It follows that they are protected by copyright since, as such, they share the originality of the whole work'. See also, the Opinion of Advocate General Trstenjak delivered on 12 February 2009 Case C-5/08 *Infopaq International A/S* v *Danske Dagblades Forening*, para 58 stating that '(t)he interpretation of 'reproduction in part' must not however be an absurd or excessively technical one according to which any form of reproduction of a work would be included no matter how minimal or insignificant a fragment of the work it is. I believe it is necessary, in interpreting that concept, to strike a balance between a technically inspired interpretation and the fact that the reproduction in part must also have a content, a distinctive character and – as part of a given work – a certain intellectual value, for which reason it is necessary to give it copyright protection. I consider that, to determine whether in a given case there is reproduction in part, it is appropriate to take two aspects into account. First, it is necessary to establish whether the

2. The Minimum Creativity Requirement: Threshold for Creativity?

Choreography relies on movements and steps which can be considered raw building blocks for the choreographer. According to a scholar,

> (a) choreographer of classical ballet has a specific movement vocabulary to work with. Like notes of music, however, these same steps can be put together in an infinite number of combinations. The prescribed steps can also be modified, as in contemporary ballet and modern dance, or repeated in different directions or done by a variety of dancers. In other words, the same step will look different in a dance depending on what step comes before and after it; the direction or tempo in which it is executed; whether it is performed while turning or leaping; what the rest of the body is doing at the same time; and how many dancers are doing it simultaneously. In short, what makes choreography interesting – instead of repetitive and boring – is the combination of the steps.[54]

However, these elements (steps and movements) cannot *per se* be copyrightable because they are standardized and so fall into the so-called 'public domain'.[55] These basic elements are used by the choreographer in the same way words are used by writers and notes by musicians.

As Traylor explains

> (it) is very different from an author writing words on paper. A choreographer works with a group of dancers who are trained in the discipline, and with a skeleton music source. The intellectual act of creation

reproduction in part is actually identical to a part of the original of the work (element of identification). In the case of reproduction in part of a newspaper article, that means specifically that it is necessary to determine whether the same words are found in the reproduction as in the newspaper article and whether those words are in the same order. Second, it must be established whether one can, on the basis of the reproduction in part, recognise the content of the work or determine with certainty that it is an exact reproduction in part of a given work (element of recognition)'.

[54] See M. Kerner, *Barefoot to Balanchine: How to Watch Dance* (New York: Doubleday, 1991), 132-133. See also Case C-5/08, *Infopaq International*, n 53 above, at paras 45-46: 'Regarding the elements of such works covered by the protection, it should be observed that they consist of words which, considered in isolation, are not as such an intellectual creation of the author who employs them. It is only through the choice, sequence, and combination of those words that the author may express his creativity in an original manner and achieve a result which is an intellectual creation. Words as such do not, therefore, constitute elements covered by the protection'.

[55] For an example of what is copyrightable see the following example reported by Schulman: 'during a visit to India I had the occasion to see dancing which was so ritualistic and stylized that there could be no doubt that the various dancers and groups followed set and identical patterns. However, these patterns, I am told, were traditional and accordingly no choreographer could claim originality for them' (*Copyright Law Revision: Studies Prepared for the Subcommittee on Patents, Trademarks, and Copyrights of the Committee on the Judiciary, United States Senate, Eighty-Sixth Congress, Second Session Pursuant to S. Res. 240, Studies 26-28* (Washington: US Government Printing Office, 1960-1961), 109).

> occurs when movements are conceived by the choreographer and directed into the trained bodied and intellects of the dancers. Only the thoughts and artistic concepts of the choreographer are manifested (…). The dancers' role is to follow the directions of the choreographer.[56]

As an example, Carrière remarks that

> ballet classical movements such as *arabesque*, *assemblé*, *cabriole*, *entrechat*, glissade, *jeté*, *pirouette* or *sissonne* are not by themselves copyrightable.[57]

Furthermore, the choreographic work is deemed original regardless of its aesthetic value.[58]

Similarly, traditional dances are based on the repetition of standardised movements.

As clearly pointed out

> (s)ocial dance steps and simple routines are not copyrightable under the general standards of copyrightability. Thus, for example, the basic waltz step, the hustle step, and the second position of classical ballet are not copyrightable. However, this is not a restriction against the incorporation of social dance steps and simple routines, as such, in an otherwise registrable choreographic work. Social dance steps, folk dance steps, and individual ballet steps alike may be utilised as the choreographer's basic material in much the same way that words are the writer's basic material.[59]

An analogous approach is embraced by the Italian courts, with a particular emphasis on the originality of a choreography to be eligible for protection.[60]

[56] M.M. Traylor, 'Choreography, Pantomime and the Copyright Revision Act of 1976' 16 *New England Law Review*, 234, 227-255 (1980).

[57] L. Carrière, 'Choreography and Copyright. Some Comments on Choreographic Works as Newly Defined in the Canadian Copyright Act', 14, available at https://tinyurl.com/5ae52nz9 (last visited 30 June 2021).

The same author adds that 'they do not represent the right kind of creativity. In other cases, features are not protectable because they are not original or are insufficiently creative'. Leistner maintains that in Member States where a higher test applied, Art 6 of Directive 93/98 and of Directive 2006/116 lowered the level of originally required to comply with the directive. Then, he traces the comparison between the criterion of 'sweat of the brow', which derives from common law and the parameters of *originalité* and *Schöpfungshöhe*, which are familiar to civil law systems. M. Leistner, 'Copyright Law in the EC: Status Quo, Recent Case Law and Policy Perspectives' 46 *Common Market Law Review,* 847-884 (2009).

[58] H.R. Rep no 941476, at 51 (1976), reprinted in 1976 U.S.C.C.A.N. 5659, 5664.

[59] The Compendium of Copyright Office Practices, Compendium II, para 450.

[60] A choreography that uses the steps of another dance, which constitutes a consolidated genre, as is the case for salsa, can only be copyrighted when it is original and, therefore, noticeably distinct from the genre whose steps it uses (Tribunale di Roma 18 March 2004, *Annali italiani del diritto d'autore, della cultura e dello spettacolo*, 493 (2005)).

In the old times, the issue related to the copyrightability of a choreography rarely arose since dances consisted in the replication of simple steps and figures, in order to make them easier to memorize and to adapt to the different music tunes.

In the 18th and 19th centuries, there appeared the so-called dance tune books and dance figures books. The former were collections of standardized steps (able to be used with different musical representations) whereas the latter contained original and new ones pictured dance figures available for different musical works (without any connection with the music).

To be eligible for protection, a combination of steps must include a *quid pluris* and a *quid novi*, resulting in a homogenous creative combination. For a practical example see the US Copyright Office Statement of Policy issued on June 18, 2012, where it is stated that

> a claim in a choreographic work must contain at least a minimum amount of original choreographic authorship. Choreographic authorship is considered, for copyright purposes, to be the composition and arrangement of a related series of dance movements and patterns organised into an integrated, coherent, and expressive whole.[61]

According to the Copyright Office the standardised steps fall within the

[61] See 77 Fed Reg 37605, 37607 (June 22, 2012): 'a mere compilation of physical movements does not rise to the level of choreographic authorship unless it contains sufficient attributes of a work of choreography. And although a choreographic work (…) may incorporate simple routines (…) exercise routines as elements of the overall work, the mere selection and arrangement of physical movements does not in itself support a claim of choreographic authorship'. It was correctly underlined that 'the work must be an 'original work of authorship' – the choreographer cannot simply copy a dance or performance and then seek copyright protection for it. The basis for originality lies in the physical setup, composition, and execution of the choreography. The choreographer must use his own creativity and imagination to use the basic dance steps, while simultaneously formulating his own unique creation. This new creation is what will be eligible to gain copyright protection' (K. Abitabile and J. Picerno, n 39 above, 7). Then again, 'it would seem possible, at least, that combinations of steps could be original, just as could combinations of words, for which there is strong support from decisions involving literary works. Whether or not the elements are original, the combination could be 'new and novel'. Combinations of dance steps also would seem analogous to a distinctive melody in music, for which there is considerable precedent for meeting the requirement of 'originality'. However, many combinations clearly belong to the public (eg, a series of turns à la seconde followed by multiple pirouettes, common in so many male solo variations in ballet), and some skeptics wonder whether any combinations could meet the statutory requirement of originality (namely, that only 'original works of authorship' are eligible for copyright)'. J. Van Camp, 'Copyright of Choreographic Works', in J.D. Viera and S. Breimer eds, *Entertainment, Publishing and the Arts Handbook 1994-95* (New York: Clark Boardman Company, 1994), 59-92. This concept can be extended to all genres of choreography, including those of cheerleaders, wrestlers, artistic gymnasts and skaters. See, among others, H.M. Abromson, 'The Copyrightability of Sports Celebration Moves: Dance Fever or Just Plain Sick' 14 *Marquette Sports Law Review*, 571-601 (2003), and L.J. Weber, 'Something in the Way She Moves: The Case for Applying Copyright Protection to Sports Moves' 23 *Columbia-VLA Journal of Law and the Arts*, 317-361 (1999).

domain of 'commonplace movements or gestures',[62] and so they are not eligible for copyright protection. For sake of precision, in this class, sports routines[63] (such as yoga sequences),[64] acrobatic exercises, and classical ballets movements are mentioned as well.

Even a mere rearrangement of these existing elements can be copyrightable as long as it consists in a new combination never seen before.[65] When it comes to music and dance, authors can rely on a wide set of material, which they can rearrange in ever changing ways to come up with an original result.

To sum up, choreographers take individual movements and steps which are not copyrightable in and of themselves[66] and come up with a choreography, which is copyrightable as a whole.

This is particularly important in light of the observation that choreographers

[62] See U.S. Copyright Office, *Copyright Registration of Choreography and Pantomime,* circular n. 52, 2017, available at https://www.copyright.gov/circs/ (last visited 30 June 2021).

[63] On the similarities between choreographies and standardized sequences of the athletes, see *ex multis* T. Griffith 'Beyond the Perfect Score: Protecting Routine-Oriented Athletic Performance with Copyright Law' 30 *Connecticut Law Review*, 689, 675-695 (1998): '(r)outine-oriented athletic performance (…) is most similar to (…) choreographic works. (…) Both tend to exhibit a planned and prepared routine, the result of which entertains the audience, displays the performer's athletic abilities, and gives the performer herself (or himself) a great deal of self-gratification. Additionally, both rely greatly upon creativity and artistic expression'.

[64] An important judgment that applied the principle expressed by the Copyright Office to the yoga was delivered by the US *Court* of Appeals for the *Ninth Circuit* in 2015 (*Bikram's Yoga College of India, L.P.* v *Evolation Yoga, LLC*, 803 F.3d 1032 (9th Cir. 2015)). The judges clarified that the sequences *(asana)* of the Bikran yoga could not be copyrightable as choreographies, because they were repetitive movements based on standardized blocks, which had not an artistic value: they lacked a creative coordination proposed by the choreographer as something new from what was previously performed. In the opinion of the court, 'because the sequence was primarily influenced by functional concerns about physical and mental well-being, it is entirely disqualified from copyright protection. Any aspects of the sequence that were motivated by aesthetic concerns are, thus, bound up with the sequence's function and are unprotected'. C. Buccafusco, 'Authorship and the Boundaries of Copyright: Ideas, Expressions, and Functions in Yoga, Choreography, and Other Works' 39 *Columbia Journal of Law & the Arts*, 425, 421-435 (2016). Another scholar observed: 'While choreographed dance is an expressive art, the Bikram Yoga series is a functional system. As has been discussed at length, the Bikram Yoga series is essentially a functional work, 'discovered' and 'researched' by Bikram, intended to be used to derive certain physical and mental benefits in the body, as Bikram himself has admitted. Bikram has never claimed that there is nothing artistic or expressive about the series. Choreographed dance, on the other hand, is primarily an expressive, artistic work. Although a dancer may benefit in certain ways from choreographed dancing-by improving his health and fitness level, increasing his flexibility, or deriving pleasure from the experience-this is certainly not the intended purpose of choreographed dance. Rather, a copyrightable, choreographed dance is intended to express the original, creative talent of the choreographer and is valued primarily for this reason'. K. Machan, 'Bending Over Backwards for Copyright Protection: Bikram Yoga and the Quest for Federal Copyright Protection of an Asana Sequence' 12 *UCLA Entertainment Law Review*, 57, 29-61 (2004).

[65] *Stanley* v *Columbia Broadcasting System* 35 Cal. 2d 653, 664, 221 P. 2d 73, 79 (1950).

[66] Therefore, choreography is not different from the other creative works, 'created from uncopyrightable component parts or formal elements – colors, notes, words, shapes, chemicals, and other substances'. C. Buccafusco, 'A Theory of Copyright Authorship' 102 *Virginia Law Review*, 1274, 1229-1295 (2016).

are often influenced by a dance school or by specific techniques as a result of their training.

Therefore, requiring choreographers to be novel, rather than original, could result in a lack of protection in the event that an author was deemed to be too faithful to the conventions of his dance style. Even the use of standardised steps belonging to a specific dance style could turn out to be eligible for copyright protection as long as the combination of these standardised steps are original.

In this, dance is similar to music where it is possible for composers to arrange existing chords and tunes to create innovative melodies. In other words,

> (b)allet and modern dance vocabularies contain basic movements which can be used by anyone and incorporated into an original choreographic work, but it is the unique combination of dance steps that determine originality.[67]

For these reasons, it is important for the courts to be extremely careful when determining whether a choreography is original or not.

3. The Domain of Originality

It is important to point out that the determination of originality must be carried out in relation to the choreographic work rather than its performance, considering the differences between the way one performer interprets the dance as opposed to another.

As clearly stated by Carrière,

> (the performance of (the) steps may greatly vary from one dancer to another according to their own interpretation. Therefore, the steps may be quite similar but their rendering by a dancer be so different that the copying choreography may be perceived as different from the copied one. It is submitted however that under the Copyright Act, it is not the performance of a work that is protected but rather the work itself.[68]

After all, the performance of a choreography influences the creative process of a choreographer because it is by seeing his creation actually performed by dancers that the author can realise if the execution of abstract idea has been successful and consistent with what he had exactly in mind.[69]

It is relevant to take into consideration that the concept of originality is flexible, and it can be applied more or less strictly. However, there is a distinct lack of case law on this specific issue.

[67] N. Arcomano, 'The Copyright Law and Dance' *The New York Times* available at http://www.nytimes.com.

[68] L. Carrière, n 57 above, 14.

[69] L.E. Wallis, 'The Different Art: Choreography and Copyright' 33 *UCLA Law Review*, 1459, 1442-1471 (1986).

It is up to choreographers and lawyers to fill this void.[70]

If one applies the stricter test, one only looks at the movements of the dancers, whereas if one applies the more extensive one, all the elements of the choreography must be evaluated which means movements, steps, the relationships with the settings and the spaces, costumes. In the latter case,

> (t)he choreographer distributes predefined elements such as steps, jumps, spins and transitions uniformly, using the entire space. The elements are stylishly and harmoniously connected with each other under consideration of different tempos.[71]

The choice of one approach over the other bespeaks different attitudes towards choreographic works which can be considered either a coherent whole or the combination of different elements, in particular movements and music.

IV. Choreography: Unique Whole or Combination of Separate Creative Blocks?

Although various legislations consider choreographies copyrightable, none of them define their nature and characteristics. Etymologically, choreography is Greek in origin and means, literally, the art of writing ballets (*χορεία* = dance and *γραφή* = writing).[72]

Consequently, a choreography is a composition created to be danced and, as such, it consists of steps and movements. However, traditionally dance was never an autonomous art form but rather one that evolved in parallel to music, from which it was originally indistinguishable being part of the *mousiké* practiced by ancient Greeks and consisting in harmonious arrangement of words, melody, and dance.

The question for the lawyer is if the combination of dance movements and music is deemed worthy of protection by the legal domain in its entirety, or whether the musical part and the step sequences can only find protection as

[70] J. Haye, 'So You Think You Can Steal My Dance? Copyright Protection in Choreography', available at https://tinyurl.com/a32kwem9 (last visited 30 June 2021).

[71] M. Kerner, n 54 above, 132-133. It is necessary to add that 'the choice of performing space might be original and an integral part of the work. If a choreographic work involves the design of movement, the choice of location for that movement seems to be part of the design. For example, the use of ramps running across the audience or the steps leading to a public monument could be considered an original element in the design of the work. However, mere use of the performing space itself probably would be excluded from protection as a 'procedure', although the pattern of movement combined with the design for the space could be protected as original' (J. Van Camp, n 61 above, 59-92). Finally, also 'the choice of a particular musical accompaniment for a certain combination of steps might be considered an original element of a choreographic design'. See, on the use of the space, among others, D.S. Palmer, *Light, Scenography and the Choreographic Space* (Leeds: University of Leeds, 2015).

[72] F. Pompei, *Le Parole del Teatro: Glossario* (Roma: Aracne, 2008), 27.

separate components of the choreographic work. As far as the nature of choreographic is concerned, its performance consists in the assemblage of different elements where music and (or, *rectius*) dance are the essential elements.[73] If, however, one considers the creative act of a choreography, the situation changes, and two different hypotheses emerge.

On the one hand, the choreographic work can be seen as a combination of dance steps and body movements, resulting from the single creative act on the part of the choreographer who is using an already existing tune, whose license must be obtained to be allowed to play it.

In this case, the two components, ie music and choreography, are different and each one enjoys copyright protection autonomously.

As a result, if a producer wants to use a choreography in a movie or a musical, he will have to obtain the permission of and pay royalties to the creator of the melodies, the author of the lyrics, and the choreographer.[74]

On the other hand, the choreography may also be the result of a dual and integrated creative act on the part of the choreographer and the composer of the music who cooperate to create a composite artistic creation.

In this case, a choreography is the product of a collaborative act of creation, which is subject of rights belonging to choreographer and the composer.

This coexistence takes different shapes in the US where a choreographic work is considered a *unicum* made up of elements which cannot be separated without distorting the nature of the whole work,[75] as opposed to Italy where it is considered a composite work and the different creative contributions can be separated even if the performance is indivisible.

In the American legal system, in fact, such a choreography can enjoy protection as a joint work, which is

> the one prepared by two or more authors with the intention that their contributions be merged into inseparable or interdependent party of a unitary work.[76]

The authors become owners as equals, which means that each of them exercises independently the right to use and commercially exploit the work, remaining 'subject only to the obligation to account to the other joint owner for any profits that are made'.[77]

[73] P. Cuoco and M. Gallina, 'In principio era il testo (autori e diritto d'autore)', in M. Gallina ed, *Organizzare teatro. Produzione, Distribuzione, Gestione* (Milano: Franco Angeli, 2014), 193.

[74] J. Van Camp, n 61 above, 59.

[75] P. Greco, 'Collaborazione creativa e comunione di diritto di autore' *Il Diritto d'Autore*, 12, 1-50 (1952).

[76] US Copyright Act, title 17, para 101.

[77] *Thomas* v *Larson* 147 F. 3d 195 (2d Circ. 1998) and, in analogy, *Community for Creative Non-Violence* v *Reid* 846 F.2d 1485, 1498 (D.C.Cir. 1988). A consequence of the joint authorship is the fact that the duration of economic rights is extended after the death of the last remaining author.

Joint works are different from the so-called collective works, constituted by separate and autonomous copyrightable works put together in a whole,[78] and from the above-mentioned category of composite works, which also includes for the Italian lawmaker the choreography created by a music composer and a choreographer, involved in a single intellectual process developed within the same timeframe.

In this latter hypothesis, in the absence of particular agreements between the two authors governing the rights on the composite work – for instance, by sharing them through specific percentages of ownership – (Art 33 of Law no 633/1941), – only the choreographer is entitled to enjoy the exploitation rights, because the legal principle to be followed in this case is the one of 'the most important contribution'.

In this sense, Art 37 of Law no 633/1941 specifies that in a choreographic work, where the music has not a major function or value, the dance prevails, so, the choreographer is the primary author and the music composer has only the right to receive a remuneration for his creative effort.

Another problem is how to legally classify the contribution given by the dancers in the choreographic works. Dancers, as performers, add 'interpretative elements' to the choreography and this was seen by some commentators as a circumstance that can lead to the recognition of a form of joint authorship between dancers and the choreographer.[79]

However, in the opinion of this author, it is necessary to distinguish the case in which the dancer can enjoy rights as a co-choreographer, because he concretely contributes in the intellectual creation, by adding something new to the work, and not only in the practical performance of the dance – from the one where the dancer is *stricto sensu* a performer,[80] in which case, like a musician or an actor, his performances are protected under the so-called related or neighboring rights, as lastly established by the 1996 *WIPO* Performances and *Phonograms Treaty* (WPPT).

In particular, the dancer is entitled to moral rights (Art 5 of WPPT) – such as the rights to be recognized as the performer of the dances and 'to object to

[78] By virtue of title 17, para 102, letter c) of the US Copyright Act, the author of one of these works has only 'the privilege of reproducing and distributing the contribution as part of that particular *collective work,* any revision of that *collective work,* and any later *collective work* in the same series'.

[79] It was observed that the recognition of such a joint authorship 'would be inconsistent with apparent understanding in the dance community, as well as with practices of choreographers in registering their copyrights' (J. Van Camp, n 61 above, 59). But it is also true that 'dance is the dancer'. For this reason, 'the relation of dancer to choreographer is not just that of executant or performer to auteur – which, however creative, however inspired the performer, is still a subservient relation. Though a performer in this sense, too, the dancer is also more than a performer'. S. Sontag, 'Dancer and the Dance' 9 *London Review of Books*, 9-10 (1987).

[80] In fact, in this case, the choreography has already been created by the choreographer as a complete work, with its meanings and its unique characteristics. Cf V.M. De Sanctis and M. Fabiani, *I Contratti di Diritto d'Autore* (Milano: Giuffrè, 2007), 77.

any distortion, mutilation or other modification of his performances that would be prejudicial to his reputation' – and economic rights (Arts 6, 7 and 8 of WPPT), such as the right to authorize the broadcasting, communication to the public and the fixation of his unfixed performance,

> the direct or indirect reproduction of their performances fixed in phonograms, in any manner or form', the distribution and the commercial rental 'of the original and copies of their performances fixed in phonograms.

V. Final Remarks

Since the middle of the 20th century, choreographies have been protected by law both by statute and case law, in many countries, as a specific artistic manifestation, with the recognition of moral and economic rights in favour of the choreographer, as long as the work has been fixed in a tangible medium and is original. Both the common and the civil law recognize the copyright of the choreographer, which can be a joint copyright with the author of the music in the event of a composite work.

From a systemic perspective, furthermore, it is possible to draw a parallel between choreography and the other arts, especially literature and music. Hence, each one of these arts is based on its own peculiar vocabulary and expressive grammar, which constitute the building blocks (steps, words, musical notes) that are combined by the artists to come up with their creative work. So far, these building blocks have not been considered copyrightable *per se*, because, on the one hand, they are usually not original, and, on the other hand, they are too small to reveal an author's personal touch, while the combination of these elements is copyrightable, provided that the standard requirements are met. This final outcome is, even if through different legal paths and formulas, widely embraced in the legal systems belonging to the Western legal tradition.

Choreography gradually emerged as an autonomous legal construct in the US, where its independent recognition was the consequence of the suggestions of the scholars and rulings of the courts, but similar patterns have also been followed in civil law countries. Therefore, the divide between the two legal families of the Western Legal Tradition is not so evident in the solutions adopted, so much so, that, when dealing with the concepts of dramatization, fixation and originality, the traditional grouping of legal systems may not be particularly accurate or useful.

On the one hand, in Italy and France, where choreographies have a long tradition, parliament failed to regulate the subject since the praxis has been, *de facto*, governing since time immemorial all cases, courts provided some degree of certainty while legislation only came in recent times. On the other hand, in the common law countries and in the Germanic area, choreographies were

confined to the dramatic and theatrical arena, for they were not classified by copyright laws as independent works.

In conclusion, the common law Countries and the solutions of Central Europe, after having enacted a legal framework, relied on the courts to outline the rules on choreographies; whereas, in France and Italy, parliament only intervened after judges had reached a mature and stable regulation on the matter.

Essays

Advancing the Rule of Law: Creating an Independent and Competent Judiciary

Larry A. Di Matteo*, Galyna Mykhailiuk** and Nina Mykhailiuk***

Abstract

An independent and competent judiciary is an essential element in rule of law systems. The rule of law continues to be tested, even in countries where the principle has been firmly entrenched as in the United States. The judicial reform movement in Ukraine offers a case study in the creation of such a system. The government and civil society recognize the necessity of developing a rule of law culture as a precursor to economic development. The judicial reform movement has resulted in new laws that include revisions to the qualifications and evaluation process for judicial appointments. Recent Constitutional amendments have given foundational authority for a wide-ranging assessment process requiring judges to meet standards of competency, professionalism and integrity. The core belief is that any approach to improve the quality of the judiciary needs to be ambitious enough to create public trust and confidence in the courts. This article analyzes the current status of reform in Ukraine, its shortcomings, and suggests how the judicial reform process may be improved. It is a case study relevant to countries transitioning from former autocratic regimes to rule of law systems.

'Power corrupts and
absolute power corrupts
absolutely'.[1]

* Huber Hurst Professor of Contract Law & Legal Studies, University of Florida, Warrington College of Business and Levin College of Law.

** Dr Habil, PhD, LLM (the University of York, the UK), Professor in the Department of European and International Law, National University 'Kyiv Mohyla Academy' and Lead Scientific Advisor of the Scientific-Research Institute of Intellectual Property, National Academy of Law Sciences of Ukraine. Member of the Parliament of Ukraine, Deputy Head of Law Enforcement Committee in the Parliament of Ukraine.

*** PhD, Associate Professor in Department of Foreign Languages for Professional Communication, International Humanitarian University and also an Academic Secretary of the Specialized Academic Council K 41.136.02 for PhD degree in philological specialties 10.02.04 – Germanic Languages; 10.02.17 – Comparative Historical and Typological Linguistics; 10.02.15 – General Linguistics.

[1] The phrase 'Absolute power corrupts absolutely' is traced to the 1887 Lord Acton's writing of 1887. See https://tinyurl.com/ycqbyjke (last visited 30 June 2021).

I. Introduction

Contemporary Western society places high value upon two ideals: individual liberty and rule of law.[2] Another tenet of democratic societies is the principle of separation of powers, in order to provide checks and balances against the arbitrary use of power by a single person or group. The judiciary's role is to prevent the capricious use of power by other branches of government and act as the protector of individual rights.

It is generally conceded that a system of democracy coupled with a market economy is the most efficient creator of economic growth. The two are intimately connected as proved by the relative wealth of Western democracies. One of the key characteristics of the democratic, market economy is the free and fair election of government officials. This article will argue that popular election of officials is not sufficient to create a fair and just society. As important as the right to vote, is a system that distributes power so that no institution of government becomes all-powerful, where no two branches of government may collude to consolidate power. The most important protector of personal rights and safeguard against government corruption is a fair, independent, and competent judiciary.[3] This article focuses on all three of these elements of an effective judiciary.

As a case study on the attempt to create Western-style judiciaries in countries formerly under autocratic rule, this article focuses on the judicial reform movement in Ukraine as it struggles to create an independent judiciary. Ukraine's economic stagnation is linked to its failure to obtain sufficient funding from the EU and foreign investors due to widespread governmental corruption.[4] It has achieved the status of a democratic, free market system but, this is not enough for the rule of law requires a separation of powers guarded by an independent and competent judiciary, and a mostly corruption-free government.

Ukraine is an ideal case study for a number of reasons. It is a country evolving out of years of authoritarianism and the yoke of corruption. The transition to democracy has been successful given that different Presidents have been elected in the last two elections.[5] The problem with corruption is seen as a major societal issue. In response, a host of judicial reform laws have been enacted. Despite their shortcomings, Ukraine's path to create an independent judiciary and rule of law system should not be viewed as a hopeless quest. Hope can be seen in the creation

[2] K. Wangmo, 'Rule of Law – A Comparative Analysis of the Rule of Law in Australia and Bhutan' *JSW Law Research* Paper no 18-6, 24 October 2018, available at https://tinyurl.com/y6dqa2sf (last visited 30 June 2021, rule of law and liberty are closely related).

[3] I. Kaufman, 'The Essence of Judicial Independence' 80(4) *Columbia Law Review*, 671, 671-701 (1980) (meaning of judicial independence); S. Burbank, 'What Do We Mean By 'Judicial Independence?' 64 *Ohio State Law Journal*, 323, 323-330 (2003).

[4] European Court of Auditors, 'Special Report EU Assistance to Ukraine', available at tinyurl.com/720t531j (last visited 30 June 2021).

[5] On 25 May 2014, Petro Poroshenko won the Presidential elections and on 21 April 2019, Volodymyr Zelensky was elected as his successor.

of civil society groups and the existence of political will to follow through on judicial reform.

An independent and competent judiciary is pivotal in order for the courts to enforce anti-corruption laws. There has been substantial progress in reforming the law related to the selection and competence of judges at all levels of the court system, as Ukrainian reform laws have followed international standards.[6] Unfortunately, the lack of an overall framework of reform and the haste in the implementation of the reforms has produced numerous legal problems and false starts.

Despite good intentions, the article will show that the initial attempts at judicial reform, 2015 to 2020, have been only partially successful. Their failures were due to hastily drafted reform laws, whose implementation was problematic since the laws were insufficiently comprehensive. The article poses that reformers in Ukraine need to adopt an evolutionary or progressive approach to the improvement of its judiciary and reject the more radical approach of the mass replacement of all judges attempted in the initial reform laws.[7] There is evidence that Ukraine will persist in reforming its judiciary. At the end of 2019 and in early 2020, the new government has recognized the weaknesses of the initial judicial reform effort by beginning to amend its reform laws in order to fill in gaps and create a more general framework for reform. This trial and error process is inherent in an evolutionary or progressive approach.

This article will look at the initial steps in the implementation of judicial reforms in Ukraine including creation of two self-governing judicial bodies, implementation of a new judicial appointment process, and appointment of a new Supreme Court. It will look at the shortcomings of the appointment process including constitutional law issues. It will review the most recent reform of creating and appointing the High Anti-Corruption Court. Part two examines the definitional issues relating to the meaning of the rule of law and its numerous elements. Part three examines the obstacles faced by countries transitioning from autocratic or corrupt legal regimes to rule of law systems. An analysis will be undertaken of the patterns of judicial reform found across a variety of countries, such as former Soviet-bloc countries trying to shake long

[6] See European Commission for Democracy through Law (Venice Commission) (2015a), Joint Opinion on the Law on the Judicial System and the Status of Judges and Amendments to the Law on the High Council of Justice of Ukraine, CDL-AD (2015) 007, 23 March 2015; European Commission for Democracy through Law (Venice Commission) (2015b), Preliminary Opinion on the Proposed Constitutional Amendments regarding the Judiciary in Ukraine, CDL-PI (2015) 016, 24 July 2015, paras 19-20; European Commission for Democracy through Law (Venice Commission) (2015c), Opinion on the Proposed Amendments to the Constitution of Ukraine Regarding the Judiciary as Approved by the Constitutional Commission on 4 September 2015, CDL-AD 027.

[7] European Commission for Democracy through Law, 'Venice Commission Welcomes Judicial Reform in Ukraine' (2 June 2018), available at https://tinyurl.com/y2axn8px (last visited 30 June 2021).

histories of autocratic rule. Most of these judicial reform movements have failed to create truly independent court systems. In most of these countries, judicial reforms have been enacted into law but not functionally implemented and in some cases, implementation has been followed by retrenchment. Part four analyzes the judicial reform program underway in Ukraine since the 2014 Maidan Revolution.[8] Part five discusses the weaknesses of judicial reform in Ukraine and provides recommendations based on the earlier review of the essential elements of the rule of law. It concludes with the most recent changes on the reform agenda that recognizes judicial reform as a long-term project, needing formal institutional change and the creation of a rule of law culture.

II. Rule of Law

The judiciary acts as a check against self-interested use of government resources by the executive and legislative branches of government. Judicial independence enables courts to 'serve as an institutional check on the legislative and executive branches and is essential for the judiciary to protect the rule of law'.[9] The American federal judiciary provides a benchmark for judicial reform because of its recognition as being insulated from corruption due to provisions found in Article III of the US Constitution.[10] First, the pool of federal judges appointed ideally represents the best legal minds in the country with the highest ratings given by the American Bar Association.[11] Second, from the very beginning of the Republic it was acknowledged that federal courts were the sole arbiters of the constitutionality of government actions and laws. This power of judicial review is ensconced in American legal tradition. Third, federal judges are appointed for life freeing them from political pressure. Fourth, federal judges are well compensated and have access to substantial resources. Finally, any hint of judicial corruption would attract an immediate investigation by the Federal Bureau of Investigation, Justice Department, and so forth.

1. Definition and Components

The rule of law by itself is a vague concept. Numerous definitions have been offered some more expansive than others. Brian Tamanaha provides a short

[8] On 1 December 2013 hundreds of thousands of people protested pro-Russian President Viktor Yanukovych refusal to sign a long-anticipated agreement to become an EU associate member (Euro Maidan or Revolution of Dignity).

[9] E. Larkin, 'Judicial Selection Methods: Judicial Independence and Popular Democracy' 79(1) *Denver University Law Review*, 65, 65-90 (2001).

[10] Art III §1: 'judges shall hold their offices during good behaviour'.

[11] 'ABA Standing Committee rates nominees 'Well Qualified', 'Qualified' or 'Not Qualified' ABA, 'Ratings', available at https://tinyurl.com/y3wln5lr (notion of 'well qualified' has been questioned in the Trump Era where judicial qualifications have not always been the measure for judicial appointments, last visited 30 June 2021).

and simple definition: 'The rule of law means that government officials and citizens are bound by and generally abide by law'.[12] This simple definition provides the core idea of the rule of law, but it fails to capture the complexity of the different elements that make up such a system. The elements associated with the rule of law include the recognition that the

> exercise of power arbitrarily cannot be conferred or upheld by law... the rule of law connects in ... different ways to a collection of institutional, formal, and procedural requirements – powers of government must be separated, laws are public, stable and non-retroactive, and courts are accessible and governed by principles of due process and justice.[13]

American legal philosopher Lon Fuller spoke of the inner morality of law:

> the principles of legality often thought to form the core of the rule of law – generality, publicity, non-retroactivity, clarity, non-contradiction, possibility of compliance, stability, and congruence between official action and declared rule constitute the 'inner morality' of the law.[14]

Legal rules that meet the inner morality of law are complimented by procedurally just administrative and judicial systems.[15] For example, anti-corruption laws may be enacted, but are of little practical significance if the processes of rule of law are not available.

2. Judicial Independence

One of the core principles of the rule of law is a 'diverse, competent, independent, and ethical lawyers and judges'.[16] An essential element of an independent court system is the insulation of judges from political and corruptive influences.[17] An independent judiciary is characterized by decisional independence, institutional independence, competency, and accountability.

[12] B. Tamanaha, 'The Rule of Law and Legal Pluralism in Development' 3(1) *Hague Journal Rule of Law*, 2, 1-17 (2011).

[13] L. Austin and D. Klimchuk, *Private Law and the Rule of Law* (New York: Oxford University Press, 2014); see also, UN Security Council, 'The rule of law and transitional justice in conflict and post-conflict societies: report of the Secretary-General' S/2004/616 (23 August 2004), available at https://tinyurl.com/y4troclf (last visited 30 June 2021).

[14] ibid 3, citing Lon Fuller (1964, 3).

[15] D. Wood, 'The Rule of Law in Times of Stress' 70 *University Chicago Law Review*, 455, 455-470 (2003).

[16] T. Banducci, 'Rule of Law and the Judiciary that Upholds It' 50(3) *Advocate*, 6, 6-7- (2017).

[17] A. Hamilton, 'Federalists Papers, No. 78 (1787-1788)', available at https://tinyurl.com/px3yxtf ('The complete independence of the courts of justice is peculiarly essential in a limited Constitution', last visited 30 June 2021). See also, C. Montesquieu, Spirit of Laws (Amherst-New York: Prometheus Books, 2002, first published in 1748), 181 ('there is no liberty, if the power of judging be not separated from the legislative and executive powers').

Decisional independence involves judicial actions unaffected by personal interest, threats or political pressure. Decisional independence is measured by individual decisions and whether they are fair and impartial.[18] *Institutional independence* refers to the constitutional and political acknowledgement of the judiciary as an equal branch of government. Institutional or structural independence requires that:

> The judiciary to be organized, governed, and funded in an autonomous manner. The *competency* of judges has a direct impact on the rule of law and requires the selection of individuals based on merit. Finally, the integrity of the judiciary requires *accountability* including the establishment of codes of ethics, impartial disciplinary boards, decisions that adhere to the constitution, and transparency.[19] The area of judicial accountability relates to both decisional and insular independence.[20]

Oversight of the judiciary is needed to make sure that decisions are free of illicit influences.

a) Selection of Judges

How should judges be selected has been a long running debate in most legal systems. In the eighteenth and early mid-nineteenth century, appointments were restricted to the elite; in the middle of the 19th century democratic elections of judges became popular; and by the end of the 19th century countries experimented with the concept of 'merit selection' by establishing judicial appointment commissions.[21] In some countries, executives appoint higher court judges, sometimes with the aid of judicial commissions. Some countries have retention systems in which judges serve an initial term but, additional terms require further assessment.

The goals of merit selection are to appoint independent, competent, and diverse judges. The rationale for merit selection is to 'de-emphasize politics while stressing qualifications' and increase diversity.[22] The general consensus is that these outcomes are best achieved through the use of independent nominating commissions. However, this begs the question of whether nominating commissions or judicial councils are any less political than other means of appointment? In a recent study, Greg Goelzhauser concluded that the method of judicial appointment

[18] R. Souders, 'A Gorilla at the Dinner Table: Partisan Judicial Elections in the United States' 25 *Review Litigation* 532, 519-574 (2006).

[19] L. Arkfeld, 'The Rule of Law and an Independent Judiciary' 46 *Judges Journal*, 13, 12-15 and 46-47 (2007).

[20] J. Tunheim, 'Challenges to Judicial Independence in Our World' 84 *Hennepin Lawyer*, 5, 4-6, (2015).

[21] G. Goelzhauser, *Judicial Merit Selection: Institutional Design and Performance for State Court* (Philadelphia: Temple University, 2019), 1-2.

[22] ibid 2 and 128.

is less important than the transparency of the process and accountability for making purely political nominations: 'The emphasis should be on issues concerning transparency, applicant pool composition, and commission decision-making'.[23] One step towards greater transparency is the keeping of merit selection records made available for inspection and review by citizens and civil society groups.

The greatest safeguard is the creation of a rule of law culture that sees an independent judiciary as indispensable.[24] The culture of judicial independence, like most cultural norms, takes a long time to be acculturated into society. The first step is 'creating adjudicative arrangements and jurisprudence and maintaining ethical traditions and codes of judicial conduct'.[25] In transitioning countries, a progressive plan to implement the various elements of an independent and competent judiciary needs to be set in place from the beginning.

b) Judicial Diversity

Ukraine recognizes the underrepresentation of women in the judiciary. The United Kingdom, previously had hoped to advance the quality and diversity of its judiciary, through the creation of the Judicial Appointments Commission (JAC) in 2005.[26] Despite fifteen years of existence the overall outcomes in diversifying the judiciary have been minimal.[27] Neutrality and impartiality on the surface leads to the preservation of the status quo and discrimination in application. For example, judicial experience is considered a prerequisite to judicial appointment. But this provides an obstacle for underrepresented groups, since their lack of experience is replicated throughout the levels of the court system.[28] In sum, diversity is a type of qualification that should be recognized independently in the appointment process.

3. Impartiality and Judicial Conduct

The other side of the coin of independence is accountability. A judiciary with unchecked power can in the wrong hands become the thing that it was established to prevent. Therefore, it is important for the judiciary to build a culture of impartiality and when individual judges fail to honor this standard they need to be held accountable.[29] 'There is an inextricable link between judicial

[23] ibid 136.

[24] See generally Int'l Association Judicial Independence and World Peace (2008, 1).

[25] ibid 3.

[26] The Constitutional Reform Act of 2005 created the Judicial Appointments Commission to review judicial applications and make non-binding recommendations, based 'solely on merit'.

[27] B. Karemba, 'Debating Judicial Appointments in an Age of Diversity' 39 *Legal Studies*, 358, 358-360 (2019).

[28] ibid 121.

[29] J. Moliterno et al, 'Independence and Accountability: The Harmful Consequences of EU Policy toward Central and Eastern Europe Entrants' 42 *Fordham International Law Journal* 481, 480-552 (2018).

ethics and judicial independence'.[30] Prompt publication of judicial decisions, media access, removal of corrupt or incompetent judges, and high standards of ethics enhance judicial accountability.[31]

4. Benefits of Rule of Law

The benefits of a rule of law system include lower levels of corruption, with attendant efficient use of scarce resources; trust in government; enhancing economic growth, and the protection of human rights. In its preamble, the Universal Declaration on Human Rights states that 'human rights should be protected by the rule of law'.[32] The United Nations noted that

> transparency and accountability in both the development and application of the law are powerful tools for ensuring public oversight of the use of public resources and preventing waste and corruption.[33]

The most debilitating influence on the Ukrainian economy and its ability to attract foreign investment has been government corruption. An independent and competent judiciary, along with independent government prosecutors, is pivotal in fighting the country's war on corruption.

The separation of powers and an independent judiciary are essential to maintaining a democratic system. The separation of powers helps ensure that law creation is based upon the building of consensus by democratically elected representatives. The judiciary functions as a protector of individual rights and as a check on the other two branches. The recurring problem in new democracies and countries transitioning to rule of law systems is that the political branches (executive and legislative) attempt to enhance their power (for personal gain) by co-opting the power of the courts.[34]

Another obstacle to independent judiciaries is a weak constitution that fails to provide adequate judicial powers or is easily amended by ruling parties.

[30] M. Greenstein, 'The Challenge of Maintaining Confidence in a Judiciary Lacking in Diversity' 55 *Judges' Journal*, 40 (2016).

[31] American Bar Association, 'Judicial Reform Index Factors' (7 January 2019), available at https://tinyurl.com/y63awovq (last visited 30 June 2021). See also, United Nations High Commission on Human Rights, 'Basic Principles in the Independence of Judiciary', UN Resolutions 40/32 (29 November 1985) and 40/146 (13 December 1985), available at https://tinyurl.com/y2ju848t (last visited 30 June 2021); Council of Europe, 'Judges: independence, efficiency, and responsibilities', Recommendation CM/Rec (2010) 12 (17 November 2010), available at https://tinyurl.com/y6qk3ntp (last visited 30 June 2021).

[32] Universal Declaration of Human Rights (10 December 1948), available at https://tinyurl.com/ybqpqebq (last visited 30 June 2021).

[33] United Nations General Assembly, 'Strengthening and coordinating United Nations Rule of Law Activities: Report of the Secretary-General: Addendum', A/68/213/Add. 1 (11 July 2014), available at https://tinyurl.com/yyh6gdfx (last visited 30 June 2021), hereinafter, A/68/213/Add. 1.

[34] See, eg, J. Rankin, 'EU Challenges Poland over Judicial Independence' *The Guardian* (10 October 2019), available at https://tinyurl.com/y3g7u2du (last visited 30 June 2021).

Many countries allow for amending national constitutions by simple votes of the legislature. In China, constitutional amendments require a two-thirds vote of the National Peoples' Congress (NPC).[35] The Polish Constitution requires a two-thirds vote of parliament.[36] A strong constitution is one that is extremely difficult to amend in order to relocate the balance of powers among the branches of government. Under Art V, amending the US Constitution requires approval of two-thirds of Congress and three-quarters of fifty state legislatures. Governments ultimately work and survive based upon the stability and legitimacy of their foundational laws.[37]

The democratic, market economy has been firmly recognized as the most efficient way of ordering societies. A United Nations Report states that:

> The rule of law and development are strongly interrelated and mutually reinforcing, that the advancement of the rule of law is essential for sustained economic growth.[38]

The rule of law provides the context for fair and predictable legal environments where businesses and entrepreneurship flourish. In the case of Ukraine, developing a rule of law system is a precondition imposed by the EU, IMF and other potential donors and investors.

An independent and competent judiciary advances the due process rights of a fair hearing under law. Without due process, personal and human rights are subject to abuse by corrupt or authoritarian governments. The enforcement of rules that conform to procedural justice norms is the means of protecting constitutional freedoms and fundamental rights. The United Nations recognizes that:

> The rule of law provides a structure through which the exercise of power is subjected to agreed rules, guaranteeing the protection of human rights.[39]

Without due process rights there is no assurance a government will not move towards authoritarianism through the suppression of fundamental rights.

[35] 'China to Amend Constitution for Fifth Time' *The NPC Observer* (15 Jan 2017), available at https://tinyurl.com/y669qjhh (last visited 30 June 2021).

[36] Polish Constitution of 1987, Art 235, available at https://tinyurl.com/y2572w5f (last visited 30 June 2021).

[37] See generally, J. Locke, *Two Treatises on Government* (Cambridge: Cambridge University Press, 2013, first published anonymously in 1689), (government legitimacy through social contract of the people); C. Montesquieu, n 17 above, 181 (theory of separation of powers); D. Hume, *Essays: Moral, Political and Literary* (New York: Wallachia Publishers, 2015, first published circa 1776), (government based upon the rule of law); J.J. Rousseau, *The Social Contract* (Amsterdam: M.M. Rey, 1762), available at https://tinyurl.com/q7lhx9y (last visited 30 June 2021, legitimate political order within a framework of classical republicanism; sovereignty is in the people).

[38] A/68/213/Add. (1 and 12).

[39] ibid 3.

5. Importance of Judicial Reputation

This section examines the particularly important role of judicial reputation in the building of trust in government institutions[40] and the elements associated with the development of a positive collective judicial reputation.[41]

a) Judicial Councils and Merit Selection

The use of judicial councils or merit commissions has become common. Judicial reform is in continuous flux routinely done in common and civil law systems, as well as in developed and underdeveloped countries, and in countries with long traditions of judicial independence and those just beginning to create such independence.[42] Judicial councils have been recognized as an international best practice.[43] Garoupa and Ginsburg estimate that sixty percent of countries, mostly within civil law systems, have adopted judicial councils.[44] The reputation of a judiciary is dependent on the reputation of the appointing councils. Judicial councils can enhance their reputations by encouraging public participation in council activities. Civil society groups can be used as a tool to monitor judicial councils, ensuring transparency, which is a key to building public trust.[45]

b) Judicial Selection in Context

Judicial reform does not transpire in a vacuum. The use of judicial councils or merit commissions in themselves does not ensure the selection of a quality judiciary. The council must be placed in the context of the politics and legal tradition of each country. This is especially true in countries with little history of merit based judicial selection and in countries bereft by corruption. A quality judiciary is a product of the quality of lawyers seeking appointment, systems of accountability, rendering of well-reasoned opinions, and when judicial corruption is thoroughly investigated.

The quality of the judicial council is linked to the number of high-ranking judges on the council. Israel is an example of how tradition and customs play an important role in the selection of quality judges. The Judicial Selection Committee is composed of four political appointments, two members of the bar

[40] N. Garoupa and T. Ginsburg, *Judicial Reputation: A Comparative Theory* (Chicago: Chicago University Press, 2015), 16 (Judicial reputation 'conveys information to the uninformed public about the quality of the judiciary and the legal system').

[41] Any reference to judicial reputation refers to the reputation of the judiciary as a whole.

[42] N. Garoupa and T. Ginsburg, n 40 above, 10.

[43] See V. Autheman and S. Elena, 'Global Best Practices – Judicial Councils: Lessons Learned from Europe and Latin America' *IFES White Paper Series* (April 2004), available at https://tinyurl.com/y3ncemvr (last visited 30 June 2021). See also, Palermo Doctrine (Elements of a European Statute of the Judiciary), available at https://tinyurl.com/y3y7uvxs (last visited 30 June 2021. In the United States, the judicial council movement began in the 1920s).

[44] N. Garoupa and T. Ginsburg, n 40 above, 101.

[45] V. Autheman and S. Elena, n 43 above, 15-16.

association, and three sitting justices. Even though the justices are in the minority they play an outsized role in the appointment process; no new justice has never been appointed without the approval of the three justices.[46]

Judicial selection of judges in the United Kingdom has gone through a unique evolution. The high court was historically part of the English House of Lords and was never viewed as a separate branch of government.[47] This changed with the Constitutional Reform Act of 2005, which created an independent UK Supreme Court. The Reform Act also established the Judicial Appointments Commission (JAC), Judicial Appointments and Conduct Ombudsman (JACO), and the Directorate of Judicial Offices for England and Wales (DJO). These various judicial oversight entities ensure a transparent judicial appointment process and accountability through continuous monitoring of judicial conduct.

III. Struggle for Rule of Law in Transitioning Countries

The court system in England is a product of centuries of evolution from the ecclesiastical courts to the royal courts to the modern unified law courts.[48] The story of judicial reform in Ukraine is part of a broader movement being replicated in numerous countries in former or current authoritarian countries. The success in these countries transitioning to rule of law systems has been mostly a record of failure.[49] The substantive failures include hasty attempts to transport Western-style laws and systems into countries with little experience in Western-style legal traditions.[50]

John Tunheim discusses a number of countries where the creation of an independent judiciary has been problematic.[51] In Uzbekistan, the government provides little job security with judicial terms lasting five years and reappointments delegated to an executive branch committee. Thus, judges 'simply do not know how to be independent'.[52] He notes that Kosovo has a well-constructed formal law of judicial independence, with appointments lasting to the age of retirement, but that in practice, 'courts remain weak in the face of corruption based on

[46] E. Salzberger, 'Judicial Appointments and Promotions in Israel: Constitution, Law and Politics', available at https://tinyurl.com/y5emggut (last visited 30 June 2021).

[47] The breadth of judicial review was shown in the court's holding Prime Minister Boris Johnson's suspension of Parliament was unconstitutional. See R (on the account of Miller) v Prime Minister, et al, (2019) UKSC 41 (24 September 2019), available at https://tinyurl.com/y4p6mxsf (last visited 30 June 2021).

[48] See A. Hogue, *Origins of the Common Law* (Bloomington: Indiana University Press, 1966), XII, 276 (history of English common law and the evolution of court system).

[49] B. Tamanaha, n 12 above, 1 (referring to countries in Africa and Asian).

[50] L.A. Di Matteo, 'Rule of Law in China: The Confrontation of Formal Law with Cultural Norms' 51 *Cornell International Law Journal*, 393, 391-444, (2018).

[51] J. Tunheim, n 20 above, 4.

[52] ibid.

family or friendship relationships'.[53] Tunheim assessment of Ukraine in 2015 is similar to Kosovo, with formal law in favor of judicial independence at odds with the judiciary playing little role in combatting corruption and the public's perception that judges are tools of the wealthy and powerful.

1. Commitment to Creating an Independent Judiciary

The success of creating a rule of law system is anchored in an independent judiciary, granted with the power to interpret and enforce constitutional rights and preserve the allocation of power among the three branches of government. Like Ukraine, 'the justice reform going on in Albania aims for a total reformation of the judicial system and the functioning of the courts', emphasizing criteria for the selection of judges of high quality.[54] The new Albanian Constitution decrees that the President's rejection of a candidate has no effect if the majority of the members of High Judicial Council vote for appointment.[55] As in Ukraine, this is an important first step but the success of the Albanian judicial reform process is far from being secured.

The experience in Serbia over the past few decades illustrate that judicial reform laws hastily designed and too slowly implemented fail in their goal of creating an independent judiciary. The result 'has not been an evolutionary process, but instead a vicious circle ... leading to serial reforms of the judiciary'.[56] Thus, it is important that reform laws be well crafted and comprehensive before implementation. Sadly, this has not been the case for Ukraine but, there is recent evidence that the government is learning from past mistakes.

2. Eternal Vigilance Needed to Maintain Rule of Law Systems

Some countries initially created independent systems but have retreated towards authoritarianism by limiting that independence. This has been the case in countries where the independence of the judiciary is a relatively new phenomenon (Poland), but also in countries with a long tradition of judicial independence (Turkey).[57] The maintenance of an independent judiciary once established requires external vigilance. The best example is the destruction of the independence of the Turkish judiciary, which is traced to the creation of a

[53] ibid.

[54] ibid 47.

[55] B. Bara and J. Bara, 'Rule of Law and Judicial Independence in Albania' 2(1) *University of Bologna Law Review*, 32-33, 23-48 (2017).

[56] V. Petrov, 'Constitutional Reform of the Judiciary in Serbia and EU Integration', 2 *EU & Comparative Law Issues & Challenges*, 4, 8, 1-9 (2018).

[57] In Turkey, an independent judiciary hardly exists. MEDEL-Association, 'La Justice en Europe: II n'y a plus de Justice en Turquie', Magistrats Europeens pour la Democratie et les Libertes (MEDEL), La justice en Europe, MEDELNET.EU 29, 36 (23 May 2017), available at https://tinyurl.com/y5743au9 (last visited 30 June 2021).

secular state by Kemal Ataturk in 1922.[58] The judiciary has long served as a vanguard against autocratic rulers and enforced the separation of religion from government. Almost a century of judicial independence was quickly washed away due to the election of a populist autocrat, resulting in the arbitrary dismissal of 4,400 judges.[59]

After the fall of communism, Poland used the government model of the United States by establishing three separate and equal branches of government including an independent judiciary.[60] Unfortunately, the current ruling party has passed laws allowing executive branch 'capture' of the judiciary.[61] A new law went into effect on 15 January 2018 that introduced a retirement age for Supreme Court judges forcing numerous existing judges off the court.[62] The law also fixed the terms of existing district court judges to four years. More importantly, the power to appoint and dismiss judges was transferred to the Ministry of Justice in the executive branch without review by the National Council of the Judiciary.[63]

The events in Turkey and Poland

> should give pause to states such as Slovakia, Czech Republic, Croatia, Montenegro as well as Ukraine, Serbia, and Kosovo.

The creation of an independent judiciary remains vulnerable in countries if authoritarian, populist's parties come to power. Maintaining judicial independence requires persistent monitoring from private actors and civil society:

> Private actors can have a significant impact on the promotion of judicial independence by utilizing both economic threats and investments in NGOs that promulgate judicial independence.[64]

The following section examines the judicial reform movement in Ukraine.

[58] A. Bali, 'The Perils of Judicial Independence: Constitutional Transition and the Turkish Example' 52 *Virginia Journal International Law*, 235, 235-320 (2012) (reviews the history of the Turkish judiciary).

[59] E. Felter and O. Aydin 'The Death of Judicial Independence in Turkey: A Lesson for Others' 38 *Journal National Association Administrative Law Judiciary*, 42, 34-56, (2015).

[60] M. Zimmer, 'Judicial Independence in Central and East Europe: The Institutional Context' 14(1) *Tulsa Journal Comparative & International Law*, 85, 101-132, (2006).

[61] See G. Goelzhauser, n 21 above, 103-127.

[62] A. Sanders and L. von Danwitz, 'Selecting Judges in Poland and Germany: Challenges to the Rule of Law in Europe and Propositions for a New Approach to Judicial Legitimacy' 19(4) *German Law Journal*, 779, 769-816 (2018).

[63] See Amnesty International (10 August 2017). The new laws have been condemned by the Venice Commission. Venice Commission, Opinion 904/2017), available at https://tinyurl.com/yxo6th2m (last visited 30 June 2021).

[64] R. Stopchinski, 'Enforcement Mechanisms for International Standards of Judicial Independence: The Role of Government and Private Actors' 26(2) *Indiana Journal Global Legal Studies*, 693, 673-694, (2019).

IV. Rule of Law in Ukraine

The above section illustrated that building a rule of law legal system has been problematic in many countries. However, a bit of caution is needed here since, the history, culture, role of law in society, and evolution of government and social institutions in Ukraine are unique in themselves.[65] Insights can be gained from more developed rule of law systems and from the failures of other countries in their attempts to establish such systems, but these insights or rules of thumb need to be tailored to the uniqueness of Ukraine.

The history of the Ukrainian government and court system is one characterized by corruption. The public perception of the Ukrainian judiciary as independent and competent has been overwhelmingly negative.[66] As of 2014, a majority of Ukrainians (fifty eight percent) saw corruption as a fact of life.[67] The weakness of the court system has resulted in under-enforcement of anti-corruption laws.[68] Transparency International's 2018 Corruption Index rated Ukraine one hundred and twenty out of one hundred and eighty countries and territories. The situation in Ukraine was summarized as follows:

> Four years since anti-corruption legal and institutional frameworks were introduced, progress is too slow. The newly established anti-corruption bodies have not succeeded.[69]

Transparency International, however, listed Ukraine as a 'Country in Transition' and noted that things may change with the growth of civil society organizations to combat the country's vested interests.[70]

1. Past: Lurching Toward the Rule of Law

An early attempt at reforming the judiciary after the Maidan Revolution of 2014 failed.[71] It consisted of the drafting of a Constitution and the Coalition

[65] B. Tamanaha, n 12 above, 1 (each rule of law project is unique).

[66] International Foundation for Election Systems (November 2005).

[67] ibid (September 2012).

[68] Business Anti-Corruption Portal, Ukraine Corruption Report (2019, 12); United States Department of State, 'Country Reports on Human Rights Practices' (2017), 2.

[69] Transparency International, 'Eastern Europe and Central Asia' (29 January 2019). The 2019 Trade Economics 'Perceptions of Corruption' ranked Ukraine 126 out of 180 countries, available at https://tinyurl.com/y684jxn8 (last visited 30 June 2021).

[70] S. Shumska, 'Shadow Economy in Ukraine: Methodology and Evaluation' 10(148) *Actual Problems of Economics*, 78, 74-83 (2013).

[71] The Law of Ukraine. On Ensuring the Right to a Fair Trial, (Zakon Ukrainy 'Pro zabezpechennya prava na spravedlyvyi sud') 18, 19-20 *Bulletin of the Verkhovna Rada* (VVR), 132, available at https://tinyurl.com/y6ogjjgx (last visited 30 June 2021) The Law of Ukraine. On the Judiciary and the Status of Judges (Zakon Ukrainy 'Pro sudoustrii ta status suddiv'), 41-42, 43, 44-45 *Bulletin of the Verkhovna Rada of Ukraine* (VVR), 529 available at https://tinyurl.com/n9c6fwt (last visited 30 June 2021). 'All translations from Ukrainian into English are by the author of the present work unless otherwise noted'.

Agreement of parliamentary fractions to move toward a 'European Ukraine'.

a) Draft Constitution of July 2014

Constitutional amendments,[72] proposed in 2014, included the abolition of supervisory powers of the Public Prosecutor's Office over the judiciary. The removal of the supervisory powers was particularly important since the Prosecutor's Office neglected its duty to fight corruption. However, the draft constitutional amendments shifted power from the Verkhovna Rada (Parliament) to the President. The President was granted the competence to appoint and dismiss key state officials, including Constitutional Court judges.

b) Presidential Decree and Coalition Agreement

President Poroshenko's 2015 Decree, entitled 'The Commission on Sustainable Development Strategy ('Ukraine – 2020'), noted that:

> An important basis for security (is) honest and impartial justice and the implementation of effective mechanisms for combating corruption.[73]

Poroshenko highlighted the low quality of the Ukrainian judicial system and the Ukrainian people's lack of trust or confidence in its workings.[74] Art 3 on 'Judicial Reform' stated that:

> (The goal was to) reform the judiciary and related legal institutions for practical implementation of the rule of law and ensuring everyone the right to a fair hearing by an independent and impartial court.[75]

The President deferred judicial reform to the Ukrainian Parliament.

On 21 November 2014, the leaders of the five major political parties initiated a draft Coalition Agreement[76] that set an ambitious reform agenda committing the country to reform numerous sectors of the government including, Constitutional Reform; Anti-corruption Reform; Justice Reform; and Law Enforcement Reform.

[72] Draft Law on Amendments to the Constitution of Ukraine on Decentralization of Power (Proekt Zakonu pro vnesennya zmin do Konstytutsii Ukrainy (shchodo detsentralizatsii vlady), no 2217a (1 July 2015), available at https://tinyurl.com/ohz5y5b (last visited 30 June 2021).

[73] Decree of the President of Ukraine no 5/2015 'The Commission on Sustainable Development Strategy' ('Ukraine–2020') (Ukaz Prezydenta Ukrainy Pro strategiu stalogo rozvytku 'Ukraina – 2020'), available at https://tinyurl.com/yy9g6y5u (last visited 30 June 2021), Art 2.

[74] President Petro Poroshenko, Address to Verkhovna Rada, 'On the internal and external situation of Ukraine', available at https://tinyurl.com/yymkckqv (last visited 30 June 2021).

[75] ibid Art 3.

[76] 'Coalition Agreement of Deputy Fractions European Ukraine' (2014) Eighth convocation of the Parliament of Ukraine (*Verkhovna Rada Ukrainy vos'mogo sklykannya Ugoda pro koalitsiu deputatskyh fraktsii 'Evropeiska Ukraina'*) (2014), available at https://tinyurl.com/y2eq7wro (last visited 30 June 2021).

Unfortunately, the follow-up implementation program[77] only established a general framework for defense and anti-corruption policies. No meaningful laws in the area of judicial reform were enacted.

c) Laws on 'Constitutional Reform' and 'Ensuring the Right to Fair Trial'

As a result of pressure from civil society organizations[78] international donors,[79] and the Venice Commission,[80] President Poroshenko introduced draft legislation at the end of 2015, which was widely criticized for its selection procedure for judges,[81] jurisdiction of the courts, and delays in implementing existing legislation.[82] In response, the Parliament voted against the 'Law on the Constitutional Court',[83] since it failed to create the independence of the Court under the Constitution.

The 2015 law 'On Ensuring the Right to Fair Trial'[84] was Parliament's response to public and international demands for judicial reform. The law's purpose was to improve the independence of the judiciary and ensure citizens' right to a fair trial. The framework of the law sought to improve judicial competence; reduce political interference; create an efficient structure; ensure financial independence; improve procedural law; ensure enforcement of judgments; and improve the quality of legal aid.

The strengthening of the role of the Supreme Court in unifying jurisprudence, substantive evaluation of judges and verification of compliance with anti-corruption laws, role of self-governance bodies in the appointment and of judges (HQCJ and HCJ), enhanced process for disciplining judges, and adoption of competitive procedures for judicial appointments were positive improvements. However, the law received mixed reviews from Ukrainian civil society and the Venice Commission.[85] The Venice Commission noted the law's numerous deficiencies did not ensure the independence of the judiciary. This is primarily due to the Constitution, which placed the power over judicial

[77] Adopted on 11 December 2014.

[78] Reanimation Package of Reforms, Anti-Corruption Action Center, Transparency International, and so forth.

[79] EU Delegation in Ukraine, US Embassy, EUAM, EUACI, COE, USAID, OECD, and so forth.

[80] Venice Commission, Opinion no 801/2015 (23 March 2015).

[81] Draft Law on High Anti-Corruption Court (Proekt Zakonu pro Vishchyi Antykoruptsiinyi sud), no 7440 (22 December 2017), available at https://tinyurl.com/yxmm78y8 (last visited 30 June 2021).

[82] R. van Rooden, 'Letter to Ihor Rainin, Head of Presidential Administration of Ukraine' (11 January 2018), available at https://tinyurl.com/y3t7y26k (last visited 30 June 2021).

[83] See Verkhovna Rada voted against the 'Law on the Constitutional Court', available at https://tinyurl.com/y4rufeaw (last visited 30 June 2021).

[84] Law on Ensuring the Right to a Fair Trial, n 71 above.

[85] Venice Commission, 'Opinion No. 801/2015' ('On the Law on the Judicial System and Status of Judges and Amendments to the Law on the High Council of Justice of Ukraine'), (23 March 2015), available at https://tinyurl.com/y38ldrxs (last visited 30 June 2021).

appointments with the President and Parliament. The Constitution needed to be amended in order to transfer the power of the President to appoint judges (initial five-year terms) and Parliament to appoint judges to permanent terms to the self-governance bodies. The Venice Commission also recommended that the new qualification assessment process be codified. Finally, the reform law has been criticized by the judiciary and civil society[86] because the process of its drafting and adoption lacked transparency.[87]

The core innovation of the law 'On Ensuring the Right to Fair Trial' was to depoliticize the judicial appointment process through the creation of new self-governance bodies – High Council of Justice (HCJ)[88] and High Qualification Commission of Judges (HQCJ).[89] Unfortunately, they were not made immediately operational because the Constitutional amendments to authorize their creation had been rejected. The lesson here is that failure to amend the Constitution first to place judicial reforms on strong legal footing led to problems in the implementation of reforms. Despite the underlying constitutional law problem, the judicial reform movement went forward with the establishment of the HCJ and HQCJ. After much delay, other initiatives were also implemented, most importantly, the establishment of the High Anti-Corruption Court to be discussed below.[90]

d) Trial and Error, Mostly Error in the Process of Selecting Judges to Supreme Court

The Ukrainian parliament enacted the law 'On the Judiciary and Status of Judges'[91] for the review and selection of judges to the Supreme Court. Ukraine established a new Supreme Court, consisting of a Grand Chamber and four specialized cassation courts.[92] Although, an issue arose as to the constitutionality of the creation of the new Supreme Court before the liquidation of the old Supreme Court, which will be discussed below.[93] The Judiciary Law also provides for two

[86] O. Ostrovska, 'Shadow' Side of Judicial Reform' (18 January 2019), available at https://tinyurl.com/yyfrvm4w (last visited 30 June 2021); O. Ovcharenko, 'Problems of Ensuring the Right to a Fair Trial' (Problemy zabezpechennya prava na spravedlyvyi sud), available at https://tinyurl.com/yygb5q72 (last visited 30 June 2021); R. Kuybida, 'On Pros and Cons of the New Law on Ensuring the Right to a Fair Trial', available at https://tinyurl.com/y6m52v2x (last visited 30 June 2021).

[87] H. Rakhalska, 'Problems of Restoring Confidence in the Judiciary in Ukraine' *Scientific Journal of HQCJ*, available at https://tinyurl.com/yxdbs5wb (last visited 30 June 2021).

[88] Ukraine Constitution, Art 125 (1996), available at https://tinyurl.com/r (last visited 30 June 2021).

[89] See 'HCJ and HQCJ' infra IV.B.2.

[90] Law of Ukraine, On the High Anticorruption Court n 81 above.

[91] Law of Ukraine, On the Judiciary and the Status of Judges n 71 above.

[92] S. Shtogun, 'Judicial Reform or Face-Lift of System of Judicial Power' 2(14) *Chasopys of National University Ostrog Academy*, 1, 1-13 (2016).

[93] B. Poshva, 'Reforms Should Be Made in a Way, Not Be Ashamed in front of Philip Orlyk while Looking into His Eyes' ('Reformy maut' provodytysya takym chynom, abi ne soromno bulo

new specialized first-instance courts – the High Court for Intellectual Property Law (HCIP)[94] and the High Anti-Corruption Court (HACC).[95] The HCIP is intended to be a specialized court of first instance for IP-related cases, with an appeal chamber, but has yet to be established. The HACC's purpose is to defend society from corruption and related crimes and provide judicial oversight over pre-trial criminal investigations. In addition, the Constitutional Court is now required to hear constitutional complaints brought by the President or through a petition from ten percent of the Parliament.

Judges are required to obtain a minimal score of six hundred and seventy out of one thousand points on a test, based upon criteria developed by experts, covering areas of substantive and procedural law, as well as a practice component relating to judging.[96] The one thousand points is divided into five hundred points for competence, two hundred and fifty points for professional ethics, and two hundred and fifty points for integrity.[97] The applicants' overall qualifications are deduced from their test scores, along with psychological evaluations, examination of their judicial dossiers; and a final review of professional activities before the HQCJ. The psychological testing of judicial candidates is unusual but was thought to be needed to ensure the appointment of ethically minded judges. Professional activities requirement includes annual submissions of asset declarations, as well as information provided by investigations of the National Anti-Corruption Bureau and the State Security Service.

On 30 September 2016, the main laws aimed at reforming Ukraine's judiciary came into force.[98] The initial step in a multi-stage process required the selection of new Supreme Court judges, introduction of judicial qualifications at different levels, formation of an anti-corruption court, application of electronic tools in the judicial system, and improvement of the legal framework of the court system. The HQCJ conducted an open competition for new justices to the Supreme Court, which attracted one thousand four hundred thirty six applications.[99] Judges from all levels of the court system, advocates, and

dyvytys' u bronzovi ochi Pylypa Orlyka'), Center of Judicial Studies (2018), available at https://tinyurl.com/y5ybw3mq (last visited 30 June 2021).

94 Order of President of Ukraine on Establishing an Intellectual Property High Court (*Ukaz Prezydenta Ukrainy 'Pro utvorennya Vyshchogo sudu z pytan intelektualnoi vlasnosti'*) (19 September 2017).

95 Law of Ukraine On High Anti-Corruption Court, n 81 above, Arts 1 and 3.

96 See High Qualification Commission of Judges, available at https://tinyurl.com/y4tywf7h (last visited 30 June 2021).

97 Legal Newspaper Online (2018).

98 J. Kirichenko, 'Changes to Constitution that launches judicial reform. How it will work? Infographics' *Ukrainian Pravda* ('Zminy do Konstytutsii, shcho zapuskaut' sudovu reformu. Yak tse pratsuvatyme? Infografika' Ukrainska Pravda) (30 September 2016), available at https://tinyurl.com/yxdlz3qj (last visited 30 June 2021).

99 New Supreme Court: Was the system reloaded? Center of Civic Monitoring and Control ('Novyi Verkhovnyi Sud: udalos li perezagruzit sistemy?' Tsentr gromadskogo monitoryngu ta kontrolu) (3 August 2017), available at https://tinyurl.com/y6rv739h (last visited 30 June 2021).

academicians applied for the positions. After its review, the HQCJ selected one hundred and eleven candidates for presidential approval.[100] In a parallel process, the Public Integrity Council (PIC) alleged that twenty-five of the candidates had previously engaged in politically motivated decisions, including the support of bans on public assembly, violations of human rights, or had not fulfilled their income-declaration requirement with sufficient transparency. The PIC, consisting of representatives of human rights communities, academic lawyers, advocates, and journalists, was established to assist the HQCJ in determining the professional ethics and integrity criteria of candidates. The PIC provides opinions on non-conformity of a judge.

Unfortunately, the process was marred by tensions between the PIC and the HQCJ.[101] In a number of cases, the PIC provided important information, which the HQCJ used in making its final selections.[102] However, the PIC process lacked transparency as to the procedures and methodology used for producing its opinions. HQCJ noted that at times the PIC did not follow its own procedures. Finally, some members did not participate in PIC decisions.[103]

The vetting of judicial qualifications for different positions continued throughout 2017. The overall number of sitting judges declined dramatically with more than three thousand judges resigning.[104] About one thousand of the resignations were due to judges failing to comply with the new transparency requirements that required judges to file income declarations.[105] Another one hundred and seventy two judges were discharged due to disciplinary actions.[106] The shortage of judges during the selection process reduced the efficiency and effectiveness of the judicial system, which threatened anti-corruption reform. Newly created bodies, National Anti-corruption Bureau of Ukraine (NABU) and the Special Anti-corruption Prosecutor's Office, faced significant impediments

[100] Poroshenko received one hundred and eleven candidates for the Supreme Court nomination, UNIAN (29 September 2017), available at https://tinyurl.com/y5mjv4e5 (last visited 30 June 2021).

[101] B. Poshva, n 93 above. See also, Center of Judicial Studies, 'The HQCJ Offered the PIC Reconciliation Outside the Court', (2018), available at https://tinyurl.com/yy835sbc (last visited 30 June 2021); Dejure Foundation, 'Formation of the New Supreme Court: Key Lessons' (Dejure Foundation, 'Formuvannya Novogo Verkhovnogo Sudu: Kluchovi Uroky') (January 2018), available at https://tinyurl.com/yy4lxb8n (last visited 30 June 2021).

[102] G. Mykhailiuk, 'Current Challenges for the Implementation of Constitutional Reform on Judiciary in Ukraine on its way towards European Integration' 14 *Journal Contemporary European Research*, 44, 40-46, (2018).

[103] Despite the lack of transparency, the PIC's decisions were deemed to be impartial. B. Poshva, n 92 above. See also, High Qualification Commission of Judges in Ukraine, available at https://tinyurl.com/y4tywf7h (last visited 30 June 2021).

[104] O. Zhukovska, 'Reform of courts without people', Ukrainian Pravda (17 August 2017), available at https://tinyurl.com/y399t666 (last visited 30 June 2021).

[105] M. Zhernakov, 'Independent anti-corruption courts in Ukraine: the missing link in anti-corruption chain', available at https://tinyurl.com/y5vpheua (last visited 30 June 2021).

[106] Espresso TV (19 March 2018), available at https://tinyurl.com/y5769fqx (last visited 30 June 2021).

in bringing cases to court.[107]

On 24 July 2018, the HQCJ announced the beginning of the registration of candidates for additional selections to the Supreme Court[108] and High Anti-Corruption Court (HACC). In 2017, one hundred and twenty out of two hundred Supreme Court positions were filled, while about eighty judges were selected in 2018-2019. The announcements were made in the presence of the media to increase transparency. The establishment of the HACC continued to be delayed.

e) Evaluations of Judges for Appellate and District Courts

The initial part of the selection process consisted of the evaluation of existing judges, begun in 2018, with those passing the evaluation process receiving higher salaries.[109] The next stage involved the evaluation and selection of appellate court judges. The plan was to promote the more qualified first instance court judges. But this strategy became problematic due to the high degree of attrition (retirements and resignations[110]) at the lower court level.[111] Some local courts did not have a single judge in place.[112] To cope with this situation, the HQCJ transferred eighty-nine trial judges to areas with shortages of judges.[113] The temporary secondment of judges was only a stopgap measure for six months. Further, about two thousand eight hundred judges from the pre-existing court system had their five-year terms expire before the implementation of lifetime terms under the new, ongoing appointment process.

Due to the above crisis, the process of appointing new judges was expedited. The idea was to duplicate the process used to select the first batch of Supreme Court judges. However, the time schedule proved to be overly ambitious, as the processing and appointment continued well into 2019. The comprehensiveness

[107] O. Zhernakov, n 105 above.

[108] In 2017, one hundred and twenty out of two hundred Supreme Court positions were filled, while eighty judges were selected in 2018-2019.

[109] R. Maselko, Realities of Judicial Reform: 99% of Old Judges and Significant Increase of Their Salaries ('Realii sudovoi reformy: 99% staryh suddiv ta znachne zbilwennya ihnoi zarplaty), available at tinyurl.com/dycgucp5 (last visited 30 June 2021).

[110] Resignations of nearly 1,500 judges occurred after the passage of the 2016 Constitutional Amendments, most did not want to participate in the new evaluation process.

[111] V. Gaponchuk, 'Judicial Reform: Pluses and Minuses of Transitional Period', *Legal Visnyk of Ukraine* (Victor Gaponchuk, 'Sudova Reforma: Plusy I Minusy Perekhidnogo Periodu', Jurydychnyi Visnyk Ukrainy) (27 February 2018), available at tinyurl.com/ydjl5n38 (last visited 30 June 2021). See also, O. Balanda, 'Ukraine and Judicial Reform: Results of the Last Year', *Ukrainian Law* (Oksana Balanda, 'Ukraina ta sudova reforma: pidsumky roku, shcho mynuv', Ukrainske pravo) (31 December 2017), available at tinyurl.com/38ygca2w (last visited 30 June 2021).

[112] Council of Judges of Ukraine, available at http://www.rsu.gov.ua/ua/pro-rsu (last visited 30 June 2021). See also, V. Pryhid, 'Orphaned Courts: How Hundreds of Thousands of Ukrainians Live without Justice' ('Sudy-syroty: yak sotni tysyach ukraintsiv zhyvut bez pravosuddya'), (10 September 2018), available at https://tinyurl.com/4y6ebdxm (last visited 30 June 2021).

[113] Fourteen judges seconded to courts without judges and another seventy-five to manage judicial caseload.

of the qualification assessment process was undermined by it being too hastily performed, resulting in the PIC resignation from the review process.[114] Regis Brillat, Special Adviser of the Secretary General of the Council of Europe for Ukraine stated that: 'There are not many examples in Europe where the entire judiciary has been reshuffled at the same time'.[115] In response, the HQCJ stated its confidence in the qualification assessment of judges due to the structure put in place before the actual assessments were performed including, benchmarks that de-politicized the process; establishment of criteria related to ensuring judicial independence and responsibility; and establishing transparent competitive procedures to prevent corruption of the process.[116]

Another rule of law issue related to the assessment process that included the changing of rules about the calculation of scores during the appointments process. Retroactive changes in assessment criteria are a technical violation of due process. However, the procedural and subsequent changes helped to improve the process to ensure the appointment of quality judges. For example, psychological testing was re-designed with the help of international donors, based on American and Western European standards that differ greatly from Eastern European standards. Unfortunately, the test questions were simple word-to-word translations from English, not adapted to the nuances of the Ukrainian language. Also, the transliteration of foreign words, such as abdication (*абдикація*), defamation (*дифамація*), procrastination (*прокрастинація*) was used without any translation into the Ukrainian language. Thus, the fairness and objectivity of the tests were less than optimal.

The compensation of judges and age discrimination were other issues that were confronted during the appointments process. The law 'On the Judiciary and the Status of Judges' did not provide guidance for judges nearing retirement age who fail new assessment tests. The Head of the State Judicial Administration recommended that if judges pass the tests, then their salaries would be increased immediately; secondly, if applicant-judges pass the qualification assessment and are older than sixty-two point five years old (sixty-five being the retirement age), then they should receive the higher pensions being provided to younger judges who had passed the assessment.[117] Unfortunately, the Ukrainian

[114] Center of Judicial Studies, n 93 above, 194.

[115] The Establishment of the New Judicial System in Ukraine is in Line with the Principles of the Council of Europe: Mr. Regis Brillat, Special Advisor to the Council of Europe Secretary General, available at tinyurl.com/1uzz6e20 (last visited 30 June 2021).

[116] Y. Shemshuchenko, 'Judicial Power in Ukraine: Current Doctrine, Mechanisms and Perspectives of Implementation' 2 *Visnyk of National Academy of Science in Ukraine* ('Sudova vlada v Ukraini: suchasna doktryna, mekhanismy ta perspektyvy realisatsii, Visnyk Natsionalnoi Akademii Nauk Ukrainy'), 37-47 (2017). See also, O. Scherbanuik, 'Competitive Selection of Judges: Problems of Constitutional Implementation' 3 *Law of Ukraine* ('Konkursnyi dobir suddiv: problem konstytutsijnoi realisatsii, Pravo Ukrainy 3'), 95, 92-109 (2018).

[117] N. Mamchenko, 'Discrimination in Judges' Remuneration: The State Judicial Administration does not Believe that the Problem will be Solved before the Elections' ('*Diskriminatsia v sudejskom*

Parliament rejected the proposals of the State Judicial Administration. Therefore, the discrimination within the retirement compensation packages for long-serving judges remains unsettled.

Although the above discussion demonstrates that new laws and processes have been adopted with good intentions. The selection and appointment process unfortunately were plagued with problems including, an overly ambitious implementation schedule and qualifications for judicial positions was set too low.

2. Present: Transitioning to Rule of Law System

This section focuses on the evolving nature of the HQCJ, including the 2019 amendments on self-governance bodies, as well as, the seating of the HACC and the alarming state of judicial salaries and budgets.

a) HCJ and HQCJ

The HCJ consists of twenty members with three members each appointed by the Parliament, President, Congress of Judges, Congress of Advocates, Congress of Representatives of Higher Legal Educational Establishments and Research Institutions and two members appointed by the Conference of Employees of Public Prosecution. The jurisdiction of the HCJ includes making proposals to the president for the appointment and dismissal of judges; executions of disciplinary proceedings against judges of the Supreme Court and the high specialised courts, and consideration of complaints regarding decisions of courts of appeal, local courts, and misconduct of prosecutors.

The HQCJ is entrusted with the assessment and selection of judges. Previously the HQCJ was loyal to the president. The new HQCJ has been largely detached from the executive and legislative bodies and has been given a broad area of competences including, organizing the selection of candidates, verification of judicial candidates' compliance with the requirements set forth by law, recommend judges for appointment, conducts disciplinary proceedings of local and appellate court judges, and monitor the lifestyles of judges. The HQCJ is divided into two chambers – one for the qualification of judges and the other for the disciplining of judges.

The reform law adopted international best practices by changing the HQCJ membership to include a majority of judges and members of the legal profession, with few appointments from the executive branch. The members are composed of eight members selected by the Congress of Judges,[118] two selected by the

voznagrazhdenii: v GSA ne veryat, chto problema reshitsa'), (2018), available at tinyurl.com/bqmn8j5i (last visited 30 June 2021).

[118] Law of Ukraine on Judiciary and the Status of Judges, n 71 above, Arts 123 and 127 (responsible for the enforcement of decisions of the Congress).

Congress of Representatives[119] of law schools and research institutions, two members by the Congress of Advocates,[120] and one each appointed by the Government Ombudsman[121] and the State Judicial Administration.[122] The Congress of Judges is the highest body of judicial self-government. The Congress of Representatives of law schools and research institutions is made up of educational and research institutions certified by the National Academy of Sciences. The Congress of Advocates is the supreme body of advocates selected by the Bar Council of Ukraine. The State Judicial Administration is a state body in the justice system that provides organizational and financial support to the judiciary and is accountable to the HCJ. Only the last two are political appointments. The importance of the amended reform laws is discussed in the next section.

b) Law 'On Amendments to Laws on Activities of Judicial Self-Governance Bodies'

The Law 'On Amendments to Some Laws of Ukraine on Activities of Judicial Self-Governance Bodies' (AJSB)[123] was enacted on 4 November 2019. The law is important in a number of ways. The AJSB requires the use of international experts, as noted above, in the judicial selection process and greater ethical investigation of candidates. However, this is another case of taking the good with the bad, as the 'devil is in the details'. The AJSB also mandates a reduction in the number of Supreme Court judges from two hundred to one hundred but fails to specify how the reduction should be implemented. Again, this is the recurring problem of good intentions hastily enacted without a deliberative process to make the law comprehensive. The result is vague mandates, such as the reduction of judges, but no legally approved process for achieving the goal. For example, an EU Delegation to Ukraine and the Canadian Embassy's 'Joint Letter' advised against the reduction of judges on the Supreme Court, arguing that

> any reduction should be based on a thorough analysis of its current structure, workload and jurisdiction.[124]

[119] Law of Ukraine on High Council of Justice (Zakon Ukrainy 'Pro Vyshchu Radu Justytscii'), Art 17.

[120] Law of Ukraine on the Bar of Ukraine (Zakon Ukrainy 'Pro advokaturu Ukrainy'), Art 54, available at tinyurl.com/3j8fdgpj (last visited 30 June 2021).

[121] Ombudsman – Ukrainian Parliament Commissioner for Human Rights, available at https://tinyurl.com/7f6nhhew (last visited 30 June 2021).

[122] See https://tinyurl.com/2hdzr7vd (last visited 30 June 2021).

[123] Law of Ukraine, On Amendments to Some Legislative Acts of Ukraine Regarding Judiciary (Zakon Ukrainy 'Pro vnesennya zmin do deyakyh zakonodavchyh aktiv Ukrainy shchodo pravosuddya') (2019).

[124] 'Joint Letter of the EU Delegation to Ukraine and Canadian Embassy to the Parliament of Ukraine' (11 September 2019).

The law changes the membership of the HQCJ and the way its members are appointed. Future candidates to the HQCJ will be assessed by a special selection panel partially composed of international experts, taken from the Public Council of International Experts (PCIE).[125] More important, a candidate must receive the unanimous support of all international experts.[126] The application process for positions on the HQCJ started in January 2020 and documents are under consideration at the present. The interview process has not started and has been postponed due to the Covid-19 pandemic. As a result, the HQCJ relaunch has been delayed for an indefinite time. Until the HQCJ is re-established, judicial appointments and the implementation of the amended reform laws are on hold.

The core mechanism of the legal reform movement, discussed in the previous sections, is the role of the HQCJ as the means to developing a competent, qualified, and independent judiciary. The relationship between the HQCJ and the creation of an independent court system is symbiotic—the higher the independence and competency of the HQCJ the greater the likelihood of creating an independent judiciary staffed by competent judges. The formula for appointing Commission members is promising, as noted above, only two of the fourteen members are government appointees. The HQCJ, along with qualifying and selecting new judges, is entrusted with drafting a judicial code of conduct, monitoring the 'lifestyles' of judges, and conducting judicial disciplinary proceedings. In sum, the HQCJ is the pivotal player in the current attempt to transform the court system from one anchored in the past, characterized by corruption and incompetency, to one that will act as a vanguard for the rule of law.

c) Appointments to High Anti-Corruption Court

Law on the HACC, enacted in 2018, includes procedural provisions dealing with the selection and training of its' judges, and how the Court conducts business.[127] Art 8 states that candidates apply and submit to the process of the HQCJ. Art 12(6) provides an ordering of criteria in the appointment of judges to the HACC: preference is given to the participant who has received a greater score for the practical part of the qualification exam; if the score is identical, the participant who has more judicial experience is given preference; the same experience, the participant who holds a scholarly degree (PhD) is preferred.

Instead of a role for the Public Integrity Council, Art 12 of the law requires the HACC to establish a Public Council of International Experts (PCIE). The PCIE is empowered to challenge the qualifications of a judicial candidate. In such cases, it meets with the HQCJ and a vote of fifty per cent of the PCIE is

[125] O. Halushka and H. Chyzhyk, 'Is Ukraine's New Judicial Reform a Step Forward?' *Atlantic Council* (24 October 2019), available at https://tinyurl.com/1kwpforc (last visited 30 June 2021).

[126] ibid.

[127] Law of Ukraine, On the High Anticorruption Court, n 81 above.

needed to advance the candidate. Given the nature of this specialized court, Art 12 requires HACC judges to continue to acquire specialized training by expanding their knowledge of professional competence, such as any new international anti-corruption standards and best practices in the fighting of corruption.

Art 11 requires the monitoring of the lifestyles of HACC judges and their families. Investigations can be pursued at the request of the HQCJ, HCJ, and PIC as well on information received from individuals and legal entities, from media and other open information sources containing data on the incongruence between the lifestyle of the judges and their declared incomes. On 20 March 2019, the HCJ sent the list of successful candidates for positions on the HACC to the President. On April 11, 2019 the President appointed the judges,[128] with the HACC becoming operational on 5 September 2019.

d) Judicial Salaries and Budgets

The law 'On the State Budget of Ukraine for 2019' suspended the planned increase in the salaries of judges of local, appellate and high specialized courts.[129] Currently, the monthly salaries of judges are less than a salary of a junior lawyer. The initial Budget of Ukraine increased salaries for new Supreme Court judges and members of the HQCJ to nine thousand two hundred euros per month. Unfortunately, law 'On the State Budget of Ukraine for 2019' suspended the planned increase in the salaries of judges, as was the case in past years.[130] To add salt to the wound, judicial salaries were cut to relocate funds to fight against Covid-19. On 1 April 2020, all judges' salaries are set at one thousand three hundred euros per month. Judicial salaries of judges and the budget for the court system remain anemic. The unreasonably low salaries are unlikely to attract the best and the brightest of legal practitioners. The financial resources needed to implement judicial reform and to create a robust independent judiciary are still lacking in Ukraine.

V. Future: Staying the Course

The judicial reform movement in Ukraine is both admirable and necessary, but its ultimate success will depend on a complex set of factors. The rule of law

> requires attention to myriad deficits such as, lack of technical capacity,

[128] Presidential Decree no 128/2019 'On Appointment of Judges to the High Anti-Corruption Court' (Ukaz Prezydenta Ukrainy 'Pro pryznachennya suddiv Vyshchogo Antykoruptsijnogo sudu') (2019), available at tinyurl.com/7l6exlov (last visited 30 June 2021).

[129] Law of Ukraine. On State Budget of Ukraine 2019 ('Zakon Ukrainy Pro Derzhavnii Budget Ukrainy na 2019 rik'), available at https://tinyurl.com/emhtvhld (last visited 30 June 2021).

[130] 'Draft State Budget for 2019 Suspends Planned Increase in Salaries of Judges,' Ukrainian News, available at tinyurl.com/1dpyemph (last visited 30 June 2021).

> lack of material and financial resources, and a lack of public confidence in government.[131]

The immensity of the task of creating an independent judiciary is captured in the following statement:

> Training judges accomplishes little by itself. A sizeable group of trained legal practitioners are needed to handle cases and help develop legal practices and shared legal knowledge. Competent clerks with adequate office space and equipment are necessary to process cases and record proceedings. Judicial compensation must be set to attract qualified individuals and ... government officials must abide by judicial rulings.[132]

Ukrainian judicial reform aims to create a competent, independent, and incorruptible judiciary. This will be a long-term project – numerous new laws are still needed, training sufficient numbers of quality judges will be extremely challenging, and continued vigilance by civil society will be required. The next sections will take a longer-term perspective, examining the importance of the right to a fair trial, oversight and accountability of judges, and the need to create a rule of law culture.

1. False but Important Start

Despite the numerous shortcomings, the judicial reform movement has been firmly established. The 2018 selection of Supreme Court judges marked the first time in history that an open competition was held for judicial positions. The pool of candidates was expanded to include current judges, advocates and academics.[133] The Law 'On the High Council of Justice' was established and provides the rules for reviewing and nominating of judges by the HQCJ. Political influence of the HJC was addressed by a composition that includes as majority of judges and non-political appointments. The HQCJ improved the level of transparency[134] and developed qualification criteria based on international standards. Transparency was also improved by advertising positions, the establishment of information channels, such as opening a Facebook account[135] and creating a YouTube channel,[136] resulting in nine thousand viewers watching the first day of applicant interviews for the Supreme Court. However, such transparency practices need to be institutionalized in law.

[131] S/2004/616, n 13 above, 3.

[132] B. Tamanaha, n 12 above, 3.

[133] Law of Ukraine on Judiciary and Status of Judges, n 71 above, Art 38.

[134] G. Mykhailiuk, n 102 above, 42-43.

[135] Facebook Page of the High Qualification Commission of Judges in Ukraine (2018) (Facebook HQCJ), available at https://www.facebook.com/vkksu (last visited 30 June 2021).

[136] YouTube Channel of High Qualification Commission of Judges in Ukraine (2018) (YouTube HQCJ), available at https://www.youtube.com/watch?v=pQh7shQ-ZrA (last visited 30 June 2021).

The goal at the core of the competition was to create a more diverse Supreme Court. First by expanding the pool of applicants to include practicing lawyers and academics. Secondly, to improve gender diversity.[137] This first goal of broadening qualifications to increase the quality and background diversity was mixed:

- Cassation Civil Court: eleven PhD degrees; twenty-five existing judges, three academicians, and two advocates;
- Cassation Commercial Court: eight PhD degrees, nineteen existing judges, four academicians, six advocates, and one varied;
- Cassation Administrative Court: nine PhD degrees, twenty-three existing judges, four academicians, and two varied; and
- Cassation Criminal Court: five PhD degrees, twenty-four existing judges, four academicians, one advocate, and one varied.[138]

On the positive side, 28.5% appointees had earned an advanced graduate law degree. On the negative side, 23.5% of appointees had no or minimal judicial experience.

In the area of gender diversity, the competition resulted in a more diverse pool of judges: (1) Cassation Civil Court with sixteen female and fourteen male judges; (2) Cassation Commercial Court with eleven females and nineteen males; (3) Cassation Administrative Court with sixteen females and fourteen males; and (4) Cassation Criminal Court with twelve female and eighteen males. In sum, forty-four percent of the judges appointed to the Supreme Court were women.

2. Judicial Accountability

In the rush to create an independent judiciary, the importance of judicial accountability is often neglected: established rule of law systems took a long time to develop the proper balance between judicial independence and judicial accountability.[139] With independence must come accountability; without accountability the seeds of corruption remain in place.

Ukrainian civil society has exposed the lifestyles of public servants at variance with their official salaries. Unfortunately, the role of civil society groups has largely been neglected in the judicial reform process.[140] Civil society is not represented on the HQCJ. Olena Halushka concludes that civil society's inability to apply to the HQCJ to open disciplinary proceedings or to appeal a

[137] R. Kuybida, 'Judicial Reform: Seven the Most Awaited Events of 2017' ('Sudova Reforma: sim najbilsh ochikuvanyh podij 2017 roku'), Center for Policy and Legal Reform (13 January 2017), available at tinyurl.com/1ui6as0h (last visited 30 June 2021).

[138] See, https://tinyurl.com/tynuuf8u (last visited 30 June 2021).

[139] J. Moliterno et al, n 29 above, 515-516.

[140] O. Halushka, 'What's next for judicial reform in Ukraine?' *Kyiv Post Ukraine Digest* (2016), available at tinyurl.com/15273fxe (last visited 30 June 2021). See also, R. Kuybida, 'Judicial Reform: Cognitive Dissonance with a Hope for Advancement' *Center for Policy and Legal Reform* (2016), available at tinyurl.com/57xfn6rn (last visited 30 June 2021).

decision needs to be remedied.[141] Another shortcoming is that PIC opinions that certain candidates fail to meet the criteria of professional ethics and integrity can be rejected by the HQCJ by a vote of eleven of sixteen members.

On the positive side, the Constitution was belatedly amended. At the same time, the new law on the 'Judiciary and the Status of Judges' was enacted to bring the pre-existing reform laws on the judicial system into conformity with the Constitutional amendments. The key elements of the amendments include: (1) removal of the power of Parliament to appoint judges and the President to dismiss judges; (2) abolishing of probationary periods for junior judges; and (3) transferred authority to discipline judges to the HCJ, the majority of whose members are required to be judges.[142] The amendments give Parliament the responsibility for establishing and dissolving courts under a procedure approved by the Venice Commission.[143]

The Venice Commission has recommended changing the four-level judicial system to a three-level one with the specialized courts within the Supreme Court.[144] The amendments authorize the creation of the High Court for Intellectual Property and High Anti-Corruption Court. The use of judicial councils to insulate the judicial selection process from improper influences has become a common feature in European countries.[145] They are seen as the best mechanism to ensure a merit-based selection of competent judges.

3. Creating a Rule of Law Culture

Brian Tamanaha notes that

> functioning legal systems require a host of secondary supportive conditions, involving a confluence of social, economic, cultural, and political factors.[146]

The transition for a country from a non-rule of law, authoritarian government to a fully democratic, rule of law country with an independent judiciary is a 'long and winding road'.[147] Creating an independent judiciary is only the first

[141] O. Halushka, n 140 above.

[142] Ukraine Constitution (2016).

[143] Venice Commission, 'Opinion 803/2015' ('Preliminary Opinion on the Proposed Constitutional Amendments Regarding the Judiciary of Ukraine'), (24 July 2015), available at tinyurl.com/yq2p79w4 (last visited 30 June 2021).

[144] S. Shtogun, n 92 above, 183.

[145] Judiciary commissions are found in Andorra, Belgium, Cyprus, Denmark, Estonia, Finland, France, Iceland, Ireland, Italy, Lithuania, Moldova, the Netherlands, Norway, Poland, Romania, Russia, Slovakia, Slovenia, the Former Yugoslav Northern Macedonia and Turkey. See CCJE, Opinion no 1 on standards concerning the independence of the judiciary and the removability of judges, available at tinyurl.com/22jtkjwq (last visited 30 June 2021).

[146] B. Tamanaha, n 12 above, 3.

[147] Beatles (P. McCartney and J. Lennon), 'The Long and Winding Road,' Let It Be Album, Apple Records (released 11 May 1970).

step in creating a rule of law system.

Historian Lawrence Friedman defines legal culture as 'the attitudes and expectations of the public with regard to law'.[148] Unquestioned obedience and devotion to the sanctity of an independent judiciary only comes with many years of fidelity to it as a bedrock principle of democratic societies. Western societies have had hundreds of years to create a culture among citizens, judges, and lawyers that holds judicial independence as sacrosanct. Currently, Ukraine is at the very beginning of implementing judicial reform. The current judicial reform movement needs to be placed in the context of an evolutionary process of creating a rule of law culture.

The Constitutional Court has been slow to hear disputes over the implementation of judiciary reform laws.[149] An active and independent Constitutional Court is needed to ensure a stable environment for judicial reform. Pressure by citizens and civil society groups must continue to ensure the government continues the reform process.[150] The United Nations has acknowledged the importance of civil society in such reform movements in order to increase the confidence of people, international monetary organizations, and foreign investors in a country's government and court system:

> Civil society organizations, national legal associations, human rights groups and advocates of victims and the vulnerable must all be given a voice in these processes.[151]

These voices are needed to ensure that judicial reforms are properly implemented. The reform movement will require reforming the Constitutional Court, improve the functioning of the PIC, and increasing the salaries and resources of judges. The likelihood of success will depend on institutional effort, political will, and guidance of the international community.[152]

4. Judicial Independence: Just a Piece of the Puzzle

Judicial independence and competency are essential elements in rule of law systems, but in and of themselves do not ensure the creation of a rule of law country. The complexity of a rule of law system is captured in the following description:

[148] S. Macaulay, L. Friedman and J. Stookey, *Law & Society: Readings on the Social Study of Law* (New York: W.W. Norton & Co., 1st ed, 1995), 71-77.

[149] G. Borkowski and O. Sovgyria, 'Current Judicial Reform in Ukraine and in Poland: Constitutional and European Legal Aspect in the Context of Independent Judiciary' 2(3) *Access Justice Eastern European*, 28, 5-35 (2019).

[150] United Nations, S/2004/616 (2004, 10), n 13 above ('Restoring the capacity and legitimacy of national institutions is a long-term undertaking').

[151] ibid 7.

[152] G. Mykhailiuk, n 102 above, 40, 44.

> ... well-established legal systems (are) highly differentiated (legislatures, police, prosecutors, judges), amply funded and have solidified legal institutions, well trained and disciplined legal officials, a well-educated legal profession, and a substantial body of legal knowledge ... (and) by and large the system operates effectively owing to the combination of broad voluntary compliance backed up by the threat of coercive sanctions imposed upon violators.[153]

The above description makes clear that judicial reform through the enactment of new laws is only a first step. The importance of an independent judiciary needs to be supported and accepted at a societal level, by politicians, business entities and civil society.[154] Civil society groups must remain diligent in monitoring the operations of the judiciary after the reform laws are implemented. It is only when trust and acceptance of the judiciary as an equal and independent branch of government, with unchallengeable power to review and strike unconstitutional acts can civil society be re-cultured. True constitutionalism and respect for the power of the courts need to be indoctrinated into civil society.[155] Until the public acknowledges the integrity of the judiciary as the protector of individual rights and as a safeguard against government corruption will the rule of law have a solid foundation in Ukraine. This perspective must be earned by the judiciary itself, with support from the government, over the coming decades.[156]

Government support of judicial reform was demonstrated by the enactment of the Law 'On Amendments to Some Laws of Ukraine on Activities of Judicial Self-Governance Bodies'[157] (ASL) on November 4, 2019. The ASL paused the process of judicial appointments to allow a greater role for international experts in the judicial selection and ethical oversight processes, and in appointments to the HQCJ. One essential reform requires future appointments to the HQCJ will be determined by a special selection panel partially composed of international experts.[158]

Despite good intentions, the new reform law repeats the errors of previous attempts. The EU Delegation to Ukraine and Canadian Embassy in a 'Joint Letter' asserted that any new reform law should come only through a deliberative and informed analysis:

[153] B. Tamanaha, n 12 above, 2.

[154] See H. Porsdam, Legally Speaking: Contemporary American Culture and the Law (Ambers: University of Massachusetts Press, 1999), 17 ('courts have generally been perceived to have a special social responsibility as arbiters, even legitimators, of change').

[155] ibid 16.

[156] G. Wood, 'The Origins of Judicial Review' 22 *Suffolk University Law Review*, 1293, 1301, 1293-1307 (1988).

[157] The Law of Ukraine on Amendments to Some Legislative Acts of Ukraine Regarding Judiciary (Zakon Ukrainy 'Pro vnesennya zmin do deyakyh zakonodavchyh aktiv Ukrainy shchodo pravosuddya') (2019).

[158] O. Halushka and H. Chyzhyk, n 125 above.

(a) speedy adoption of imperfect laws seriously undermine the reform effort, compromise the good intentions of the new Government, and lead to an inadvertent legacy.[159]

The new government has also focused on reforming the prosecution system including the Office of General Prosecutor and State Bureau of Investigations. The Office of President noted that corruption persecutions have not increased since the appointment of a new General Prosecutor in August of 2019. Consequently, on March 5, 2020, the Parliament dismissed the General Prosecutor and Head of State Bureau of Investigations. In sum, despite the creation of a committed reform movement, Ukraine remains in the grip of corruption and the judiciary has only marginally proved itself as a means of reducing corruption. The World Justice Project 'Rule of Law Index 2020' showed that the quality of the rule of law and Ukrainian judicial system remains at a low level.[160] In the area of corruption the rankings and score (one equal highest score) show widespread perceptions of corruption across all branches of government, the judiciary is seen as less corrupt, but still low (Corruption in Executive Branch, .32, Corruption in Judiciary, .49, Corruption in Legislature, .09). The overall rule of law score for Ukraine improved slightly from the previous year but still placed seventy-two of one hundred twenty-eight countries (score, .51/1).

Unfortunately, key metrics for judicial and rule of law related issues were substantially lower than the country's overall score. Regarding the government, the scores and rankings show restraints on government power, .46 (ninety of one hundred twenty-eight); absence of corruption, .33 (one hundred ten of one hundred twenty-eight); regulatory enforcement, .43 (one hundred of one hundred twenty-eight); and government limited by judiciary .32. The low scores on restraint of power and judicial oversight of government indicate that the judicial branch remains weak. The scores relating to the operations of the judiciary show corruption scores are especially low in the criminal law system – criminal justice: effective investigations, .26 and civil justice: no corruption, .41. Related parameters also are disappointing: no improper government influence, .37; effective and timely adjudication, .38; and due process .44. The scores were higher in two areas: accessibility and affordability of court proceedings, .62 and effective and impartial ADR, .63.

5. Constitutionality of Judicial Reforms

The more fundamental critique of the implementation of the early judicial reform laws – selection and nomination of judges – was based on rule of law rationales. However, Article 126 of the Constitution does not provide grounds

[159] Joint Letter, n 124 above.

[160] World Justice Project, 'Rule of Law Index 2020', available at tinyurl.com/4qfl5lor (last visited 30 June 2021).

for changes in the Supreme Court and the new appointment process. A number of existing judges challenged the constitutionality of the selection process. This caused a great deal of uncertainty over whether the work conducted by the HQCJ in the vetting and appointment of judges under the new scheme would be invalidated. The Constitutional Court delayed ruling on the matter for an unduly amount of time. This delay in rendering a decision on such an important constitutional issue is further evidence that Ukraine is a long distance away from a functioning rule of law system and a rule of law culture.

Finally, on 18 February 2020 the Constitutional Court ruled on the legality of the new selection process and judicial appointments.[161] The Court upheld the work of the HQCJ and the legality of the new appointments proceeding from the presumption that the:

> amendment of the Constitution of Ukraine must ensure the principle of institutional continuity which means that the bodies of state power established by the Basic Law of Ukraine continue to function in the interests of Ukrainian people and exercise their powers, fulfill their tasks and functions defined in the Constitution of Ukraine, regardless of these amendments, unless such amendments provide for a substantial (fundamental) change in their constitutional status, including their liquidation.

Moreover, the Venice Commission in its opinion stated that when adopting a new Constitution, its transitional provisions should not be used as means of suspending the powers of persons elected or appointed under the previous Constitution. The dismissal of all judges, apart from exceptional cases, does not comply with European standards and the rule of law; it is not possible to replace all judges without prejudice to the continuity of justice.[162]

The Court reasoned that mass dismissals were not permitted under the Constitution in existence at the time of the new selection process put in place by the HQCJ. However, it validated the selection and appointment process of new judges. As a result the existing judges will remain in office, as well as those appointed through the new process implemented by the HQCJ. The end result is that the rationale can be seen that the court was simply expanded with new appointments, while most existing judges retained their positions. As a result, the 2019 law authorizing the reduction in the size of the court has been put on hold. This result incidentally is in line with the Venice Commission's recommendation that any downsizing of the Supreme Court should be done slowly and after careful deliberation.[163]

[161] Constitutional Court of Ukraine Ruling no 2-p/2020 (18 February 2020), available at https://tinyurl.com/d6sbfh5e (last visited 30 June 2021).

[162] Venice Commission, 'Opinion on the Legal Status and Remuneration of Judges', (CDL-REF (2012)006), available at tinyurl.com/1lvymm2i (last visited 30 June 2021) (2012, para 111).

[163] nn 127 and 128 above and accompanying text.

6. Summary: Untangling the Chaos

Ukrainian legal reforms have been enacted in a piecemeal fashion and at a haphazard pace. Going forward the government and civil society will need to visually construct a more comprehensive framework for a functional rule of law court system. Constitutional amendments and the establishment of the HACC is evidence that the government recognizes the gaps in existing laws and the need to amend existing judicial reform laws. Figure 1 summarizes the status of the judicial reform movement in Ukraine based on the key elements of an independent and competency judiciary.

Fig. 1 - 'Status Report: Rule of Law Elements (Independent Judiciary) in Ukraine'

Element	Status	Shortcomings	Improvements Needed
Constitutional Framework	Amended	Amended too late	Additional amendments
Non-Political Evaluators	Best Practices (non-political appointees & international experts)	Somewhat untested	To be Determined
Qualifications	Anti-Corruption standards (integrity) but low competency criteria	Shortage of qualified candidates	Long-term investment
Quality of Appointment Process	Framework in place	Hastily conducted	Deliberative process
Transparency	Video-taping available online	Uneven	Complete transparency
Judicial Salaries	Acknowledged	Not funded	Minimum: staged funding
Funding	Insufficient	Insufficient	Full funding
Judicial Training	Acknowledged	Pre-planning	Piggyback Training offered internationally
Legal education	Some quality law schools	Uneven	Require LLM to practice; curricular changes; English training
Independent and anti-corruption prosecutor	Transition (prosecutor recently discharged)	No established process to select independent prosecutor	Amendments to law needed; development of career prosecutors willing to investigate corruption
Ethics; judicial code	Recognized need	Transition	Training needed

The recent Constitutional amendments placed the HQCJ on more sound legal

footing, but more amendments will be needed as the reform movement continues to expand in its scope. The law creating the HQCJ adopted best practices fixing its composition to mostly non-political actors. The role of international experts in HQCJ and HACC processes is another best practice, but it is untested so, it is still to be determined whether the role of international exports will provide optimal input in selection decisions or whether their role will need to be fine-tuned. The major obstacle to the competency of judges is the low standards required for appointment. This is due to the shortage of quality candidates. This issue can only be resolved in long-term investments in legal education and professional training. The initial appointments process undertaken by the HQCJ was sound in substance but weak in implementation. In the future, the process will need to be undertaken at a reasonable time frame that allows for careful deliberation. Parts of the initial selection process was made transparent (placed online), while other parts where more secretive. Going forward complete transparency is imperative in order to gain the confidence of the public. Judicial salaries remain woefully inadequate. Despite budgetary constraints, the government needs to fully fund the judicial system in order to reduce corruption and attract foreign investment. A well-funded judiciary should be a high priority since it is a major building block in creating a more prosperous country. The government is aware of the need for improvements in legal education and the training of judges. It specifically mandates that members of the HACC continue their educations after appointment. However, the infrastructure for skill development has yet to be constructed. A stopgap measure would be the greater use of training programs in foreign countries.

As important as the improvement of the judiciary, a quality court system is dependent on the independence and competency of government prosecutors. In the past, prosecutors were closely aligned with government officials resulting in few corruption investigations. To the present, prosecutors have been reluctant to bring claims of corruption. This is clearly a major problem that needs to be addressed or the connection between an independent judiciary and anti-corruption efforts will be greatly diminished. Finally, accountability of judges requires the drafting and inculturation of judicial conduct and ethics codes, and continuous oversight. This again is a long-term project, but initial steps should be expedited.

VI. Conclusion

There are a number of binary relationships tied to an independent judiciary. First, the judiciary, as a core element of the rule of law, is vital to the creation of efficient markets. Second, an independent judiciary is needed to anchor anti-corruption programs. In the end, the future prosperity of Ukraine (and other transitioning countries) and its integration into the EU hinges on the creation of

a competent, independent court system and sufficient reductions in corruption.

The judicial reform movement in Ukraine has attracted much political, academic and civil society attention. However, it has been beset by uncertainty and miscues. The lessons learned from Ukraine's initial attempt at judicial reform include the need to make all necessary structural changes in the Constitution before enactment of judicial reform laws. Second, greater transparency in the vetting and selection of judges is of paramount importance. In sum, successful transitioning to a rule of law system is enhanced by a process of deliberative, careful drafting of judicial reform laws that comprehensively implement the many elements needed to create an independent and competent judiciary.

The countries of Western Europe developed independent court systems over the course of two hundred years or more. The countries previously under Soviet rule, countries of the former Yugoslavia, and others without a rule of law tradition will need to continue to fight to create independent court systems. The Venice Commission has welcomed changes that have taken place during the last few years of the Ukrainian judicial reform movement. Similarly, the Council of Europe has highlighted the significant achievements of Ukrainian judicial reform, one of which is the newly formed Supreme Court.[164]

The recent amending of the Constitution and the establishment of the High Anti-Corruption Court is a further signal that there is a political and civic will coalescing to continue judicial reforms. The features of the reform laws, such as the enforcement of anti-corruption laws and improving the education and expertise of judges, lay the basis for the creation of a rule of law culture. Unfortunately, Ukraine's hasty implementation of a country-wide evaluation and appointment process for judges at all levels of the court system failed in numerous ways. However, it set a threshold that every judge should be required to participate in a fair and open competition to ensure selection is based upon merit and not political connections. True success will be the product of a long-term, 'evolutionary' process, including bringing well-educated and experienced newcomers into the judicial system at the point of entry. The future remains uncertain, but there is reason to hope that Ukraine is on the road to a rule of law society.

[164] Center of Judicial Studies, 'Council of Europe Accesses the Judicial Reform in Ukraine' (Tsentr Suddivskyh Studij, 'Sovet Evropy Provodit Otsenku Sudebnoj Reformy v Ukraine'), available at https://tinyurl.com/56qmee7s (last visited 30 June 2021).

Contract Automation from Telematic Agreements to Smart Contracts

Alberto Maria Gambino and Andrea Stazi*

Abstract

Technology creates new opportunities for socio-economic relations, commercial exchange and to overcome national borders, allowing to conclude and execute agreements more quickly regardless of the distance between the parties. However, technology also tests the contractual institution as it requires to adapt it to immediate, transnational, automatic uses, and to the legal issues that consequently arise. This chapter aims to analyze the evolution of the relationship between technology and contract through the *fil rouge* of contract automation, with specific regard to the conclusion of telematic agreements and the next frontier for contract automation, ie 'smart contracts'.

I. Automation and the Contract

Over a century has passed since the German doctrine, primarily with Auwers,[1] and a few years later the Italian doctrine, with Cicu and Scialoja,[2] began the

* Alberto Maria Gambino is Prorector and Full Professor of Private Law at the European University of Rome. Andrea Stazi is Associate Professor of Comparative Law and New Technologies Law at the European University of Rome. This essay is intended for publication in the forthcoming book 'The transformation of private law', Liber Amicorum for Mads Andenas' 65th birthday. The content is the product of the joint effort of the authors. Nevertheless, paragraphs I and V are attributable to Alberto Maria Gambino, paragraphs II, III and IV to Andrea Stazi.

[1] W. Auwers, *Des Rechtsschutz der automatischen wage nach gemeinem Recht* (Göttingen: W. Fr. Kästner, 1891); Id, *Des Rechtsschutz der automatischen wage nach gemeinem Recht* (Norderstedt: Hansebooks, 2016); F. Günther, *Das Automatenrecht* (Göttingen: W. Fr. Kästner, 1892); Id, *Das Automatenrecht* (Whitefish: Kessinger, 2010); K. Schels, *Der strafrechtliche Schutz des Automaten* (Munich: Jur. Diss. Erlangen, 1897); F. Schiller, *Rechtsverhältnisse des Automaten* (Zurich: Zürcher Diss., 1898); P. Ertel, *Der Automatenmissbrauch und seine Charakterisierung als Delikt* (Berlin: Druck von Wilhelm Pilz, 1898); H. Neumond, *Der Automat. Ein Beitrag zur Lehre über die Vertragsofferte, Archiv für die civilistische Praxis* (Heidelberg: Mohr Siebeck, 1899), 166.

[2] A. Cicu, *Gli automi nel diritto privato* (Milano: Società Editrice Libraria, 1901), 8; Id, *Scritti minori*, II (Milano: Giuffré, 1965); A. Scialoja, *L'offerta a persona indeterminata ed il contratto concluso mediante automatico* (Città di Castello: Tipografia S. Lapi, 1902). More recently, see among others: A.M. Gambino, *L'accordo telematico* (Milano: Giuffrè, 1997); F. Delfini, *Contratto telematico e commercio elettronico* (Milano: Giuffré, 2002); S. Sica and P. Stanzione eds, *Commercio elettronico e categorie civilistiche* (Milano: Giuffré, 2002). In recent European doctrine, see: R. Schulze and D. Staudenmayer, 'Digital Revolution – Challenges for Contract Law', in R. Schulze and D. Staudenmayer eds, *Digital Revolution: Challenges for Contract Law in Practice* (Baden-Baden: Hart-Nomos, 2016), 19; S. Grundmann and P. Hacker, 'The Digital

exploration of the then futuristic relationship between automatic devices, private relations and contractual stipulation.

Such authors paved the way to the analysis of the impact of the so-called automation on the contractual features, dynamics and remedies. A long and detailed analytical path, then, contributed to the taxonomy and evolution of the institution of contract, including its elements and related events.

The development of digital technology and telematics[3] has led to the emergence of new contractual typologies based on economic behaviors that go beyond evaluating the convenience of the terms, to satisfy needs through ever faster, often immediate, and effective exchanges.[4]

The process of depersonalization of relationships and the consequent objectification of the contract which had already matured with mass bargaining have been fulfilled with telematic negotiation. This is even more apparent when it operates through electronic agents, that is automatic programs which conclude contracts between machines on the basis of preventive instructions without individual control.[5]

Those forms of bargaining led to the evolution of the model of progressive development of contractual consent, where a reduction in transaction costs and

Dimension as a Challenge to European Contract Law. The Architecture', in S. Grundmann ed, *European Contract Law in the Digital Age* (Cambridge: Intersentia, 2018), 3.

[3] The term telematics derives from the Greek adverb 'tele-' which means distant, and from the suffix '-ema' which means functional element that gives shape to something. Thélème was also the imaginary abbey with which Gargantua, a character conceived by Francois Rabelais, a French humanist of the 16th century, foreshadowed a world of complete freedom. Unlike others, Thélème was an abbey without walls and external barriers: everyone could enter it and was well received, some could get lost. Therefore, the concept of telematics, indicates a set of IT services that are offered and used in real time through a telecommunication network, which may act as communication tools between the parties. On this subject, one may also see: A.M. Gambino, A. Stazi and D. Mula, *Diritto dell'informatica e della comunicazione* (Torino: Giappichelli, 3rd ed, 2019).

[4] In this regard, see, *ex multis*: K. Kryczka, 'Electronic Contracts and the Harmonization of Contract Laws in Europe – An Action Required, a Mission Impossible?' 13 *European Review of Private Law*, 149-170 (2005); P. Sammarco, 'I nuovi contratti dell'informatica. Sistema e prassi', in F. Galgano ed, *Trattato di diritto commerciale e diritto pubblico dell'economia* (Padova: CEDAM, 2006).

[5] The use of an electronic agent introduces an alternative path beyond the party's control in the traditional production process and manifestation of the will to negotiate. Indeed, complex processing mechanisms lead to the determination of an artificial, predetermined negotiating will which is potentially increasingly different from that of the user due to technological evolution. In this case, the results of the bargaining are not always foreseeable and it cannot be excluded that the electronic agent will at least in part complete contracts upon unwanted assumptions or beyond the program user's expectations. Regarding these multiple related issues, see among others: P. Perlingieri, 'Relazione conclusiva', in Id, S. Giova and I. Prisco eds, *Il trattamento algoritmico dei dati tra etica, diritto ed economia* (Napoli: Edizioni Scientifiche Italiane, 2020), 379; G. Teubner, *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi* (Napoli: Edizioni Scientifiche Italiane, 2020); G. Sartor, 'Cognitive Automata and the Law: Electronic Contracting and the Intentionality of Software Agents' 17 *Artificial Intelligence and Law*, 253-290 (2009); G. Finocchiaro, 'The conclusion of the electronic contract through "software agents" A false legal problem? Brief considerations' 19 *Computer Law & Security Review*, 20-24 (2003).

a greater possibility of information regarding the subject of the exchange have overcome the perplexities around fewer reflections on purchasing certain goods or services.[6]

II. Types of Telematic Contract

With reference to telematic contracts, which are characterized by the use of electronic means in order to put distant parts in contact, subcategories have been identified.

A first general distinction is between telematic contracts in the broad sense, characterized by the provision of a service electronically, and those in the strict sense, in which the bargain is formed thanks to the electronic impulses exchanged between the terminals connected to distance.[7]

A key classification between telematic contracts is based on the subjective profile, which differentiates the business to business contracts related to the negotiations between professional operators, the business to consumer contracts involving relationships between professional operators and consumers, and consumer-to-consumer relationships between private entities outside their professional activities.[8]

However, this classification appears to be linked to statutory schemes that the new commercial techniques have overcome. Indeed, the provisions aimed at protecting the 'weak part' in the regulation of electronic commercial relations are not anchored merely to the subjective condition of the party itself, whether consumer or professional, but they are based on the objective conditions in which the parties place themselves in such relationships.[9]

Another reconstruction of French origin proposes to subdivide telematic contracts into three heterogeneous classes. The first one includes agreements concluded outside the system and executed through the terminals. The second

[6] See eg: A.M. Gambino, *L'accordo telematico* n 2 above; Id, 'Il contratto telematico', in W. Bigiavi, *Giurisprudenza sistematica di diritto civile e commerciale*, 2 (Torino: UTET, 1999).

[7] French doctrine lists electronic contracts and 'conclus et exécutés par la télématique' (par exemple, procédures de réservation électronique) contracts; - soit conclus par la télématique mais exécutés en dehors de cette technique (par exemple procédures de commande par terminal); - soit conclus en dehors de la télématique mais exécutés par elle (par exemple contrats d'accès aux banques de données)': X. Linant De Bellefonds and A. Holland, *Contrats informatiques et télématiques* (Parigi: Delmas, 1988), 161.

[8] In a comparative perspective, see eg: C.W. Pappas, 'Comparative U.S. & (and) EU Approaches to E-Commerce Regulation: Jurisdiction, Electronic Contracts, Electronic Signatures and Taxation' 31 *Denver Journal of International Law & Policy*, 325-348 (2020); Z.S. Tang, *Electronic Consumer Contracts in the Conflict of Laws* (Londra: Bloomsbury, 2015); F.F. Wang, *Law of Electronic Commercial Transactions: Contemporary Issues in the EU, US and China* (London: Routledge, 2014).

[9] One may see also: A. Stazi, 'Digital copyright and consumer/user protection: Moving toward a new framework?' 2 *Queen Mary Journal of Intellectual Property*, 158-174 (2012); Id and D. Mula, n 3 above.

class identifies agreements concluded through the IT medium and executed outside the telematic network. In the third case, both the contract conclusion and implementation take place online, eg for the circulation of rights relating to intangible assets and IT services.[10]

With regard to the procedures to conclude electronic contracts that are functional to electronic commerce, two main options for expressing consent are identified: a) contracts where consent is expressed with a 'click', the so-called 'point and click' on an offer contained in a website – or more recently in an app on mobile devices;[11] b) contracts in which consent is expressed by email.[12]

In the context of contracts concluded via access to a website or app, according to a part of the doctrine the completion of the agreement and therefore the *Idealtypus* of the electronic contract consists in completing a form including the typing of the card numbers with a subsequent acknowledgement of receipt by the offeror.[13]

On the other hand, in the contract concluded by email the principle of receptivity is followed but tempered by the principle of effective knowledge: the agreement is completed through an effective dialogue with mutual communication.[14]

III. The Telematic Agreement

Agreement and dialogue do not constitute a monad. The Principles of European Contract Law, or PECL reaffirm the centrality of the agreement even in the absence of dialogue. The PECL identify the contract with sufficient agreement, thus marking a clear break with the principle of completeness of

[10] P. Le Tourneau, *Contrats informatiques et électroniques* (Parigi: Dalloz-Sirey, 4th ed, 2006).

[11] The term 'app' is an abbreviated form of 'application', which in practice is used especially with regard to mobile apps for mobile phones, tablets, etc. The majority of the applications are found in real virtual stores called app stores. The contracts concluded through the app appear similar to the hypothesis of the contract concluded through access to the site, as also this case is a form of communication one to many and not one to one as for the contracts via email.

[12] These typologies can be framed in the inter-absent relationships. However, they have at least an unusual aspect characterizing them, in that the parties do not follow the normal logical-chronological sequence between the moment of processing the communication and that of sending the reply, or at least this sequence is strongly compressed. So, while in the contact *de visu* the assignment that follows an announcement can be easily corrected according to canons of reasonableness, in telematic contracts the screen of the program does not allow to easily identify neither the professional quality of the offeror nor the legal binding nature of the commitment undertaken.

[13] In this perspective, the credit card spending manifests the willingness to legally bind the purchaser and has real efficacy involving the conclusion of the contract for the beginning of execution, according to a unilateral contract scheme. See A.M. Gambino, *L'accordo telematico* n 2 above, 138.

[14] Provided eg in the Italian legal system at Art 1335 of the Civil Code and in the common law systems in the so-called mailbox rule; in this regard, see *amplius*, below in the following paragraph.

consent on all elements of the contract.[15]

On the one hand, technology has been considered a solution to the problem of the formation of contracts with a view to facilitating the exchange of promises, reducing the time delay due to distance, limiting communication risks, automating responses and reducing transaction costs.[16]

Still, the trade-off between the use of technology and consent is increasingly evident, the bottom line being that their very presence is inversely related. In fact, consensual contracts have been more and more replaced by formality-based agreements.[17]

Therefore, the interpreter must assess the meaning of the tenderer's communications on a case-by-case basis, starting from the moment the purchase is solicited, and inspired by the usual criteria of good faith and correctness.

Furthermore, in a system specially based from the beginning on spontaneous adherence to certain rules of good conduct, the so-called 'netiquette',[18] it has been found that such 'needed' courtesy generates constraints, which are still socially penalized although not legally punishable.[19]

Different legal systems have responded to the problem of forming contracts between distant parties in different ways, and the two main solutions – the mailbox rule, and the reception rule – have given rise to an ongoing debate.[20]

A number of technical solutions have appeared in the history of the contract regarding the formation of agreements when the parties are not in the

[15] See: Arts 2: 101 and 2: 103 PECL. In the same sense, Art 2: 204 PECL states that: 'any form of declaration or behavior of the oblate that indicates acceptance of the proposal constitutes acceptance', and Art 2:211 PECL states that: 'even when the contract conclusion procedure is not structured in proposal and acceptance, the rules of this section apply equally with the appropriate adaptations'.

[16] See: M. Granieri, 'Technological contracts', in P.G. Monateri ed, *Comparative Contract Law* (Cheltenham-Northampton: Edward Elgar, 2017), 408; J.M. Moringiello and W.L. Reynolds, 'From Lord Coke to Internet Privacy: The Past, Present and Future of the Law of Electronic Contracting' 72 *Maryland Law Review*, 452-500 (2013); J. Savirimuthu, 'Online Contract Formation: Taking Technological Infrastructure Seriously' 2 *University of Ottawa Law & Technology Journal*, 105-144 (2005).

[17] In this sense, see: R.T. Nimmer, 'Electronic Contracting: Legal Issues' 14 *Marshall Journal of Computer & Information Law*, 211-246 (1996); C. Reed, *Internet Law: Texts and Materials* (Londra: Butterworths, 2000), 175; J.K. Winn and B.H. Bix, 'Symposium: Cyberpersons, Propertization, and Contract in the Information Culture: Diverging Perspectives on Electronic Contracting in the U.S. and EU' 54 *Cleveland State Law Review*, 175-189 (2006); J.M. Moringiello and W.L. Reynolds, 'From Lord Coke to Internet Privacy: The Past, Present and Future of the Law of Electronic Contracting' 72 *Maryland Law Review*, 452-500 (2013).

[18] Term composed of *net* (network) + *etiquette* ('label').

[19] See: I.T. Hardy, 'The Proper Legal Regime for Cyberspace' 55 *University of Pittsburgh Law Review*, 993-1056 (1993); M.R. Burnstein 'Conflicts of the Net: Choice of Law in Transnational Cyberspace' 29 *Vanderbilt Journal of Transnational Law*, 75-116 (1996).

[20] Among others, see: R.B. Schlesinger and P.G. Bonassies, *Formation of Contracts: A Study of the Common Core of Legal Systems* (New York: Dobbs Ferry, 1968); I.R. Macneil, 'Time of Acceptance: Too Many Problems for a Single Rule' 122 *University of Pennsylvania Law Review*, 947-979 (1964).

same place at the same time. Some of these solutions have been completely replaced over time, while others have been recently added or are under development. Whenever new technology emerged, the question arose as to whether pre-existing contractual law rules could meet trade needs and ensure an adequate level of certainty in commercial practice.[21]

The mailbox rule represents the solution adopted in the common law system, according to which the contract is intended to be perfected at the time the oblate sends their acceptance.[22]

Civil law countries have preferred the reception rule, according to which a contract is formed when the offeror receives the recipient's acceptance, following a similar logic to the contextual bargaining in person.[23]

[21] See: S. Holmes, 'Stevens v. Publicis: The Rise of "No E-Mail Modification" Clauses?' 6 *Washington Journal of Law, Technology & Arts*, 67-68 (2009); A. Rawls, 'Contract Formation in an Internet Age' 10 *Columbia Science & Technology Law Review*, 200-231 (2009).

[22] The rule was established in Adams v. Lindsell (1818) 106 ER 250, and later accepted in the United States: Mactier's Adm'rs v. Frith, 6 Wend. 103 (NY 1830). In doctrine, see: R. LeRoy Miller and G.A. Jentz, *Business Law Today* (Boston: Cengage Learning, 2010); R.T. Nimmer, n 17 above, 222; A. Rawls, n 21 above, 205; E.A. Farnsworth, 'Comparative Contract Law', in M. Reimann and R. Zimmermann eds, *The Oxford Handbook of Comparative Law* (Oxford: Oxford University Press, 2016), 916; A.F.M. Maniruzzaman, 'Formation of International Sales of Contracts: a Comparative Perspective' 29 *International Business Law Journal*, 487 (2001). The leading case on the matter is a decision of the British High Court on the case *Mondial Shipping & Chartering BV* v *Astarte Shipping Ltd*., (1995) CLC 1011. The case assessed the rituality of the declarations of legal relevance made by e-mail communications, as well as the operation of the same where the withdrawal from its contractual obligation was manifested by sending an electronic communication within the deadline. The Court opted for the applicability of the rules relating to inter absent contracts, based on the so-called shipping principle, or mailbox rule, which identifies the moment of consent with the act of sending the declaration by the oblate (ie the accepting subject). Having to establish the time to which the withdrawal dated, considering that it did not have to be activated before a certain term and that the relative declaration had been sent a few minutes before the same term, yet coinciding with the non-working weekend, the Court ended up stating the full operation of the withdrawal declaration in light of the fact that it would have become known only on the first following business day.

[23] See Art 11 of Directive 2000/31/EC and Artt 1326-1335 of the Italian Civil Code, while in France the Civil Code does not provide for a solution and the French courts have generally decided these questions on a case by case basis. J. Bell, S. Boyron and S. Whittaker, *Principles of French Law* (Oxford: Oxford University Press, 1998), 312, however, note that there seems to be a preference among the French courts for acceptance at the time and place of dispatch. In the Italian legal system, Art 1335 of the Civil Code establishes a presumption of knowledge *iuris tantum* with respect to the declaration sent to the recipient's address, therefore at the time of the knowledge or knowability of the communication by the latter. This presumption can thus be won by proof against the recipient who could not receive the news of the communication, without fault. In Germany, the contract is concluded when the acceptance reaches the offeror. This rule can be inferred from § 130(1) BGB, according to which any declaration of intention (being it an offer, a revocation of an offer, an acceptance or another declaration) directed to an absent person becomes effective when it reaches that person. See H. Kotz, *Vertragsrecht*v (Heidelberg: Mohr Siebeck, 2012), 99; W. Flume, 'Allgemeiner Teil des bürgerlichen Rechts', in *Das Rechtsgeschäft*, II (Berlino: Springer, 1979), 657; H. Beale, B. Fauvarque-Cosson, J. Rutgers and S. Vogenauer, *Cases, Materials and Text on Contract law* (Oxford: Hart, 2019), 257; K. Zweigert and H. Kötz, *An introduction to comparative law* (Oxford: Clarendon Press, 1998), 362, state that every declaration of will is effective as soon as it

International trade legislation has shown a preference for the reception rule,[24] while over time the mailbox rule has lost relevance, also as a consequence of the spread of digital technologies.[25]

Comparative studies have dealt extensively with the issue of evaluating one option over the other, and the ability of each solution to adapt to online bargaining.[26]

In some cases, in the telematic context, such as sending emails or instant messages, the interval between sending and delivery is so short that the offer and acceptance are separated by a negligible fraction of time, thus making the revocation almost impossible.

This explains the European Union legislator's rationale behind extending the application of the *jus poenitendi* already foreseen in the distance contracts directive to electronic bargaining.[27] If there is no time to weigh an agreement, some time must be given to change your mind and dissolve the agreement.

Also in the United States, the Restatement of Contracts subjects 'substantially instantaneous' bidirectional communications to the same principle applicable to acceptances where both parties are present.[28]

The mailbox and reception rules are only different solutions to allocate the communication risk between the parties. The common law stated that a rule based on the tenderer's actual receipt of the acceptance would have given rise to the

comes into the 'sphere of influence' of the addressee. They take the old school example of a bird-lover that chooses not to empty the letter-box in his garden for fear of affrighting the tomtits within; in that case, the declaration is treated as having arrived. The concept of reaching (*zugehen*) is well explained in the *Delivery to a housemaid* case, even though it concerns an offer. In that case, the Reichsgericht stated that an offer becomes effective when the letter or the telegram containing it has been delivered at the offeree's house, regardless of whether the offeree has been informed of the offer. See RG, 25 October 1917 RGZ 91, 60 (delivery to a housemaid).

[24] Regarding the complex interaction between the provisions adopted in the Convention on contracts for the international sale of goods and the Convention on the use of electronic communications in international contracts, see: C.H. Martin, 'The Electronic Contracts Convention, the CISG, and New Sources of E-Commerce Law' 16 *Tulane Journal of International & Comparative Law*, 467-504 (2008)

[25] A. Rawls, n 21 above, 207; E. Mik, 'The Effectiveness of Acceptances Communicated by Electronic Means, or – Does the Postal Acceptance Rule Apply to Email?' 26 *Journal of Contract Law*, 8 (2009); UCITA denies the application of the mailbox rule for electronic messages; see: S.M. Kierkegaard 'E-Contract Formation: US and EU Perspectives' 3 *Shlider Journal of Law Commerce & Technology*, 37 (2007); V. Watnick 'The Electronic Formation of Contracts and the Common Law "Mailbox Rule" ' 56 *Baylor Law Review*, 197 (2004), believes that since there is no clear default rule for electronically sent acceptance times, the mailbox rule should be maintained for electronic acceptances of contracts not covered by UCITA.

[26] J.M. Moringiello, 'Signals, Assent and Internet Contracting' 57 *Rutgers Law Review*, 1307-1350 (2005), highlighted how in practice consumers perceive transactions on paper as different from electronic transactions.

[27] See Arts 9-16 Directive 2011/83/EU, which has extended the provisions of Directive 97/7/EC.

[28] Restatement (Second) of Contracts § 64 (1981). In doctrine, see: A. Rawls, n 21 above, 210; P. Fasciano 'Internet Electronic Mail: A Last Bastion for the Mailbox Rule' 25 *Hofstra Law Review*, 971-1004 (1997); E. Mik, n 25 above, 16, according to whom the mailbox rule is still suitable for the use of email in the formation of the contract.

risk that a withdrawal of the acceptance could arrive before it was received.[29]

To regulate this situation, common law has evolved in such a way that all contractual communications – offers, revocations, refusals – are effective upon receipt, except acceptance.[30] The mailbox rule is also instrumental in favoring a faster conclusion of the contract, to the point that a contract is in any case concluded if the tenderer has not yet received the acceptance.[31]

Since the withdrawal of acceptance is less and less practicable due to technological evolution and it would not make sense to maintain a difference in treatment between the different ways of forming the contract, the reception rule has become prevalent in many legal systems.[32]

Within the European Union, Art 11 of the E-commerce Directive requires that the service provider acknowledges receipt of the recipient's order without undue delay and by electronic means.[33] The law considers the order and the acknowledgement of receipt received when the parties to whom they are addressed are able to access them; this rule also applies to the exchange of e-mail or equivalent individual communications. Therefore, the Directive is based on a sort of reinforced reception rule in the aim of establishing a harmonized solution at European level.[34]

Moreover, as an additional protection for users, the Directive introduces a differentiating element to overseas solutions, requiring that the service provider makes available appropriate, effective and accessible means[35] to the recipient,

[29] I.R. Macneil, n 20 above, 953. Regarding the time that passes from the offer and acceptance, Macneil notes that it is not surprising that the Anglo-American courts have kept it as short as possible by adopting the mailbox rule. Indeed, it can be said that one of the main functions of this rule is to reduce the duration of the offeror's right of revocation. Furthermore, not only does the rule itself shorten the revocation period, but it also removes an element of uncertainty from the contractual relationship. By comparing the risks, the bidder is already exposed to the possibility that his offer is never received by the counterparty; see: A. Rawls, n 21 above, 212.

[30] See: P. Fasciano, n 28 above, 222.

[31] In this regard, see: I.R. Macneil, n 20 above, 954; E. Mik, 'The Effectiveness of Acceptances Communicated' n 25 above, 9. 'Receiving' an acceptance does not necessarily correspond to the recipient's actual knowledge. In electronic bargaining, it has been argued that receipt of acceptance occurs when it has entered the information processing system designated for such messages by the tenderer; see: A. Rawls, n 21 above, 211.

[32] In this regard, see again: A. Rawls, n 21 above, 204.

[33] On this point, see: S.M. Kierkegaard, n 25 above, 28; C.H. Ramberg, 'The E-Commerce Directive and formation of contract in a comparative perspective' 26 *European Law Review*, 429-450 (2001).

[34] J.K. Winn and J. Haubold, 'Electronic Promises: Contract Law Reform and E-Commerce in a Comparative Perspective' 27 *European Law* Review, 575 (2002), recognize a possible interference with the national provisions, mentioning § 130.1 of the German BGB regarding the moment in which the contractual declaration is considered as received. Neither the UETA nor the E-Sign take a position on the applicability of the rule of the shipment or mailbox or of the reception for the formation of the contract. A. Rawls, n 21 above, 209, proposes the adoption of the reception rule for all contracts in the United States, however formed, in order to ensure consistency in the decisions on the subject at national and international level.

[35] For example, a home screen, or a pop-up window, or an intermediate review image.

to allow the identification and correction of insertion errors before placing the order.[36]

In common law systems, the binding nature of the offer has traditionally been opposed to the use of the invitation to treat, which rather is a mere promotional message.[37]

The Vienna Convention – also applicable to international sales – is in line with this interpretation, where in Art 14.2 it is expected that:

> (a) proposal other than one addressed to one or more specific persons is to be considered merely as an invitation to make offers, unless the contrary is clearly indicated by the person making the proposal.[38]

Thus, a 'click' on the digitized goods on the screen could mean a request for information on the pre-contractual phase, an invitation to offer, a purchase request, a request to send the goods or an acceptance of the offer. Through hypertext navigation, then, you can easily switch from an online contract advertisement to the actual offer.[39]

The common law courts method consisted in the case-by-case assessment between invitation to offer and offer binding, attaching strong relevance to the context in which the business takes place[40] and the existence of an *intention* to be obliged.[41]

Instead, the civil law system (such as the French system) grounds the contract analysis on the principle of completeness of the offer, which is traced

[36] This step is only procedural and does not change the solutions adopted by national laws for the formation of contracts; see: M. Granieri, n 16 above, 19.

[37] This, in the reductive perception that 'the merchant might find himself involved in any number of contractual obligations which he would be quite unable to carry out his stock being necessarily limited'; see: P. Owsia, *Formation of Contract. A Comparative Study Under English, French, Islamic and Iranian Law* (Londra: Kluwer, 1994); *contra*: P.S. Atiyah, *An Introduction to the Law of Contract*, (Oxford: Oxford University Press, 4th ed, 1989), 65.

[38] Therefore, the intention to bind to the offer must be explicitly expressed, since the presence of the essential elements of the contractual proposal is not considered a sufficient requirement. Also Art 2.2 of the UNIDROIT Principles of International Commercial Contracts of 1994 is in substantial compliance with the common law principle, according to which: 'A proposal for concluding a contract constitutes an offer if it is sufficiently defined and indicates the intention of the offeror to be bound in case of acceptance'.

[39] In this regard, see eg: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A comprehensive approach to stimulating cross-border e-Commerce for Europe's citizens and businesses; one may also see: A. Stazi, *La pubblicità commerciale online* (Milano: Giuffrè, 2004), 16.

[40] See US District Court, Southern District of New York, September 9, 1996; Memorandum, Attorney General, State of Minnesota, County of Ramsey, Second Judicial District, C6-95-7227, December 1996, 6.

[41] Which can be inferred through 'an inquiry whether the facts show some performance was promised in positive terms in return for something requested'. See S. Williston, *Contracts* (New York: Thomson, 1957), 65. See also, among others: E. Peel, *Treitel The Law of Contract* (London: Thomson, 14th ed, 2015), 2; P.S. Atiyah and S.A. Smith, *Atiyah's Introduction to the Law of Contract* (Oxford: Oxford University Press, 6th ed, 2006), 35.

back to the figure of the unilateral declaration of will (*déclaration unilatérale de voluntée*) and must include also the conditions of the contract in addition to the expression of the intention to contract.[42]

Thus, for example the sending of catalogs and price lists, and the display of the goods in the shop window with the relative price are generally considered to be offers made to the public, whose conditions cannot change once accepted by the customer.[43] The bidder can refuse to fulfill only because of a serious and legitimate reason, such as the exhaustion of the goods.[44]

In the Italian system, if one considers online electronic catalogs in which goods or services are offered for direct purchase, a type of offer to the public pursuant to Art 1336 of the Civil Code is configured on condition that the site: a) is open to any users or in any case to a number of users so vast as to make the sending of the offer independent of the individual person; b) contains all the essential elements of the contractual proposal.[45]

If there is a mere invitation to offer, on one hand the applicable rules will be those on advertising communications, with particular reference to the principles of non-deception, truth, correctness and completeness; on the other hand, the regulation of pre-contractual liability in Art 1337 of the Italian Civil Code applies.[46]

In the French system, the level of information required in the content of the offer changes with the *nature du contrat*.[47]

In the German system, a declaration is not qualified as an offer but rather as an invitation to make offers if the offeror prevents his or her offer from

[42] A. Weill and F. Terré, *Droit civil, Les obligations* (Paris: Dalloz, 4th ed, 1986), 142, define the offer as 'une declaration unilatérale de volunté adresée par une personne à une autre, et par laquelle l'offrant proposed à autrui la conclusion d'un contract'. According to J. Ghestin, *Traité de droit civil: Les obligations: Le contrat: Formation* (Paris: Librairie Générale de Droit et de Jurisprudence, 1988), 219: 'On peut a priori définir l'offre comme une manifestation de volunté unilatérale par laquelle une personne fait connaître son intention de contracter et les conditions essentielles du contrat'. The invitation, the French expression of the invitation to treat, distinguishes between the non-binding offer subsystems, the invitation to faire des offres and the offer avec réserves or sans engagement; see: M. Planiol and G. Ripert, *Traité pratique de droit civil français* (Paris: Librairie Générale de Droit et de Jurisprudence, 2nd ed, 1953), 145.

[43] In this regard, see eg: A. Weill and F. Terré, n 42 above, 220.

[44] The problem of the scarcity of the goods available to the seller is resolved here with the application of the 'first come, first served' principle, since '*est de la nature des choses que l'offre au public soit réservée aux premiers acceptants dans les limites des quantités offertes*'; see: J. Ghestin, *Traité de droit civil: La formation du contrat* (Paris: Librairie Générale de Droit et de Jurisprudence, 1993), 266.

[45] Examples of offers to the public can be found in the products of a supermarket that have exhibited the sale price, in the products offered in teleshopping, etc. According to Art 1336, para 2, the revocation of the offer to the public in the same forms as the offer – or in another equivalent – is effective also for those who have not heard of it. It is not necessary, therefore, that anyone who has heard of the offer must then know of the revocation in order for it to be effective also against her. It is sufficient that offer and revocation are carried out in the same forms.

[46] In this regard, one may also see: A. Stazi, *La pubblicità commercial*, n 39 above, 176.

[47] See eg Court of Cassation 27 June 1973 has ruled that in a hypothesis of *bail* (lease), the offer must mention '*la chose louée, le montant du loyer, et la date possible d'entrée en jouissance*'.

having binding force through the use of express phrases, such as '*freibleibend*' or '*ohne Obligo*' ('without engagement').[48] However, if the addressee agrees to the invitation and the proposer remains silent, the jurisprudence usually considers such declaration as an acceptance.[49]

Beyond the formal solutions individually adopted by different legal systems, the comparative perspective shows that the most interesting question concerns the role of consent as a mechanism to form the contract. Online, the relevance of the human factor is reduced to the point that there is a tendency to replace the complexity of the will with formality, and the 'humanistic model' of contractual behavior is a prerequisite in favor of prevailing needs for speed and efficiency of transactions.[50]

IV. The Next Frontier: Smart Contracts (?)

The development of distributed ledger technologies, first of all the Blockchain,[51] allows the creation of so-called 'smart contracts',[52] characterized

[48] According to § 145 BGB, the offeror is bound by his or her offer and cannot withdraw it.

[49] For example, in the *Aeroplane Charter* case (BGH, 8 March 1984, NJW 1984, 1885) the Federal Court of Justice has declared that the use of the words 'without engagement' does not necessarily prevent a communication from constituting an effective offer. Indeed, according to the principle of good faith, the proposer should have expressly rejected the offer. On the contrary, his or her silence should count as acceptance. See H. Beale, B. Fauvarque-Cosson, J. Rutgers, S. Vogenauer, *Cases, Materials and Text on Contract law* (Oxford: Hart, 3rd ed, 2019), 362.

[50] In this regard, see: M. Granieri, n 16 above, 19, who notes that in mass market transactions the practice of standard terms has provoked discussions on consensus which are now superseded by the so-called «Rolling contracts», which continue until someone decides to terminate them, rather than until a certain date (for example such contracts are generally known in the practice of insurance relationships); R.T. Nimmer, 'Electronic Contracting: Legal Issues' 14 *Marshall Journal of Computer & Information Law*, 212 (1996); M.J. Radin, 'Humans, Computers, and Binding Commitments' 75 *Indiana Law Journal*, 1125-1162 (2000), who distinguishes between "contract as consent" and 'contract as product'. For a discussion of the reconstruction of individual will in contract theory and the adoption of a subjective consensus theory linked to liberalism, see: E.M. Weitzenboek, 'Electronic Agents and the Formation of Contracts' 9 *International Journal of Law and Technology*, 218 (2001).

[51] A distributed ledger technology, such as the Blockchain, is a consensus mechanism for geographically distributed, shared and synchronized digital data without a central administrative authority or centralized data store. See eg: UK Government Scientific Adviser (2016) Distributed Ledger Technology: beyond Blockchain, available at urly.it/3d6tg (last visited 30 June 2021); M. Giuliano, 'The Blockchain and smart contracts in the innovation of law in the third millennium', available at urly.it/3d49n (last visited 30 June 2021); P. De Filippi and A. Wright, 'Decentralized Blockchain Technology and the Rise of Lex Cryptographia', available at https://tinyurl.com/5ex3j233 (last visited 30 June 2021).

[52] The idea of smart contracts was proposed by Nick Szabo in the nineties of the last century; see: N. Szabo, 'Formalizing and Securing Relationships on Public Networks', available at https://tinyurl.com/v2r5mv5m (last visited 30 June 2021); Id, 'Smart Contracts: Building Blocks for Digital Markets', available at https://tinyurl.com/9nbeeftt (last visited 30 June 2021); Id, 'Smart Contracts', available at https://tinyurl.com/r95ura8a (last visited 30 June 2021), who argued that the objectives of such contracts would be to fulfill contractual obligations such as payment terms,

by the self-execution of the contractual clauses without the need for human intervention, and generally excluding the possibility of interrupting such execution or modifying the content, with the exception of the options of multi-signature or self-destruct.[53]

In some cases, the smart contracts represent the implementation of a previous contractual agreement in the legal sense, whose clauses of are formalized in the computer source code.[54] Therefore, the contracting parties have the advantage of structuring their relations and services in a more efficient and self-executing way, regardless of the ambiguity of natural language.[55]

In other cases, smart contracts introduce new coded relationships that are both defined and automatically applied by the computer code, but are not linked to any underlying contractual rights or obligations.[56]

On the other hand, regardless of the technical necessity, there may be a legal need to draw up a smart contract in writing in order to make its clauses legally binding and applicable at the judicial level.[57]

privileges, confidentiality and even enforcement, and to minimize both harmful and accidental exceptions and the need for trusted intermediaries. With regard to the definition of smart contracts, see: R. Pardolesi and A. Davola, 'What Is Wrong in the Debate About Smart Contracts', available at urly.it/3d49a (last visited 30 June 2021); R. De Caria, 'The Legal Meaning of Smart Contracts' 26 *European Review of Private Law*, 731-751 (2019); R. Herian, 'Legal Recognition of Blockchain Registries and Smart Contracts', available at https://tinyurl.com/2j3tf32y (last visited 30 June 2021); L.W. Cong and Z. He, 'Blockchain Disruption and Smart Contracts', available at https://tinyurl.com/4697v22b (last visited 30 June 2021); one may also see: A. Stazi, *Automazione contrattuale e "contratti intelligenti". Gli smart contracts nel diritto comparato* (Torino: Giappichelli, 2019), 105.

[53] Multi-signature, or 'multisig', verification technology allows an individual to stop running a smart contract until several parties have signed the transaction with their private keys. These can include not only the parts of the smart contract, but also an external third party, a so-called referee. See: K.D. Werbach and N. Cornell, 'Contracts Ex Machina' 67 *Duke Law Journal*, 345 (2017). Furthermore, the code of most smart contracts contains a so-called kill switch. Solidity, the language used to write smart contracts on the Ethereum Blockchain, allows an operation called self-destruction, which removes the smart contract code from the Blockchain; see: H. Eenmaa-Dimitrieva and M.J. Schmidt-Kessen, 'Creating markets in no-trust environments: The law and economics of smart contracts' 35 *Computer Law & Security Review*, 84 (2019).

[54] The source code, in computer science, is the text of an algorithm of a program written in a programming language by a programmer during programming. It therefore defines the flow of execution of the program itself. See: Wikipedia, 'Source code', available at https://it.wikipedia.org/wiki/Codice_sorgente (last visited 30 June 2021).

[55] Thus, for example, a smart contract was created that simulates the mechanism for a public funding campaign, the so-called crowdfunding, with fifty-six lines of computer code (see http://www.mintchalk.com/c/68f3e). The creation of smart contract models, in practice, could lead to a reduction of the role of lawyers in the moment of contract formation, especially with respect to those that can be easily modeled; on this point, see: M. Corrales at al eds, *Legal Tech, Smart Contracts and Blockchain. Perspectives in Law, Business and Innovation* (Berlin: Springer, 2019).

[56] In this regard, see among others: Chamber of Digital Commerce – Smart Contracts Alliance (2018), 'Smart Contracts: Is the Law Ready?', available at https://tinyurl.com/w39ndcnx (last visited 30 June 2021).

[57] See: P. De Filippi and A. Wright, 'Decentralized Blockchain Technology' n 51 above, 11, who found that, while at the beginning smart contracts were mainly developed to automatically execute

However, a smart contract is not always immutable. First, the Blockchain could be 'forked' by the majority of users. Second, the computer code of smart contracts can contain several functions that allow for a certain range of flexibility.[58]

Again, registering smart contracts on a Blockchain platform usually comes at a cost.[59] This effectively excludes the inclusion of detailed and complex clauses in smart contracts, as vice versa happens in legal practice especially in common law but increasingly also at an international level.[60]

In a technical sense, it is possible to define smart contracts as computer protocols that execute themselves by applying the lines of the computer code for which they were programmed, which are stored on a distributed register.[61]

Eventually, they allow the drafting and possible automation of the agreements between the parties – as a truly 'contractual' case in the legal sense – according to an 'if/then' logic.[62] This happens if an economic function is pursued through the computer protocols and is recognized by the legal system in which they are intended to carry out their effects.

A smart contract program is executed by a network of so-called miners who, once consensus has been reached on the outcome of the execution, update the status of the contract on the Blockchain accordingly. In this way, users can send or receive money, data, etc. through a contract.[63]

Therefore, based on the capabilities of the Blockchain, the smart contracts operate autonomously in a transparent, anti-tampering and tendentially immutable way.[64]

derivatives, options, futures and swaps, later they began to be used to facilitate the sale of goods on the network between unrelated persons without the need for a centralized organization. The authors cite in this sense the example of OpenBazaar, an open source service aimed at creating a decentralized global market in which people can buy and sell products directly, without intermediation costs or centralized control (see: https://tinyurl.com/t54xdax5, last visited 30 June 2021).

[58] Like the multi-signature or self-destruct assumptions mentioned above, but also functions like 'call' (which accepts an arbitrary number of arguments of any type), 'enums' (a way to create a user-defined type), 'self-destruct', and also variable functions that allow the smart contract to process inputs external; in this regard, see: A. Juels and B. Marino, 'Setting Standards for Altering and Undoing Smart Contracts', in J.J. Alferes et al eds, *Rule Technologies. Research, Tools, and Applications* (Cham: Springer, 2016), 151.

[59] Called for example 'gas' on the Ethereum platform; see: https://tinyurl.com/vtj2ex2v (last visited 30 June 2021).

[60] On this point, see: G.O.B. Jaccard, 'Smart Contracts and the Role of Law', available at https://tinyurl.com/ys8dj3m9 (last visited 30 June 2021).

[61] P. De Filippi and A. Wright, *Blockchain and the Law: The Rule of Code* (Cambridge: Harvard University Press, 2018), 33; V. Buterin, 'Ethereum White Paper: A Next-Generation Smart Contract and Decentralized Application Platform', available at https://tinyurl.com/y2bfutse (last visited 30 June 2021).

[62] See eg: F. Idelberger et al, 'Evaluation of Logic-Based Smart Contracts for Blockchain Systems', in J.J. Alferes et al eds, *Rule Technologies* n 58 above.

[63] See: G.O.B. Jaccard, n 60 above, 5; A. Juels et al, 'The Ring of Gyges: Investigating the Future of Criminal Smart Contract', in E. Weippl ed, *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (New York: ACM, 2016), 283.

[64] In this regard, see: P. De Filippi and A. Wright, *Blockchain and the Law*, n 61 above, 72; D.

These features give the contracting parties several significant advantages over traditional contracts: they can rely on contractual promises that are stored in the smart contract, ie the transaction protocol automatically executed without recourse to judicial intervention, and do not have to trust the counterparty.

This allows them to take calculated risks even in areas where the parties are not directly opposed to each other, but which are often characterized by anonymity and application risks, as is usually the case in electronic commerce and international contracts.[65]

Consumers/users could particularly benefit from these advantages in a relevant way, since they usually face difficulties and costs for which they neglect to assert their rights in court.[66]

Furthermore, smart contracts involve the possibility of reducing transaction costs by performing some functions currently performed by intermediaries such as eBay, Amazon, PayPal, etc.[67] Smart contracts allow the parties to incorporate the commercial practice in their agreement, bypassing the need for explicit but[68] redundant negotiation.

Automatic application or compensation has the potential to reduce the amount of disputes, so increasing certainty and limiting performance monitoring costs.[69] Therefore, smart contracts give generally rise to a further reduction of human intervention and formalization of the contract.[70]

Once more, smart contracts increase the speed with which it is possible to execute contractual relationships if compared to traditional contracts. Since they are not dependent on paper and related procedural steps, and can be performed in real time, they simultaneously enable more cost savings and faster execution than paper contracts.[71]

Finally, smart contracts offer an alternative to one of the most characteristic aspects of contractual drafting: the intrinsic ambiguity of natural language,[72]

Linardatos, 'Smart Contracts: Some Clarifying Remarks From a German Legal Point of View', available at https://tinyurl.com/urrms8an (last visited 30 June 2021).

[65] See: P. Ryan, 'Smart Contract Relations in e-Commerce: Legal Implications of Exchanges Conducted on the Blockchain' 7 *Technology Innovation Management Review*, 14-21.

[66] In this regard, see: O. Borgogno, 'Usefulness and Dangers of Smart Contracts in Consumer and Commercial Transactions', in L.A. Di Matteo et al eds, *The Cambridge Handbook of Smart Contracts, Blockchain Technology and Digital Platforms* (Cambridge: Cambridge University Press, 2019), 8.

[67] ibid 13; M. Sokolov, 'Smart Legal Contract as a Future of Contracts Enforcement', available at https://tinyurl.com/26sm33p3 (last visited 30 June 2021); E. Mik, 'Smart contracts: terminology, technical limitations and real world complexity' 9 *Law, Innovation and Technology*, 277 (2017).

[68] On this point, see: J.M. Sklaroff, 'Smart Contracts and the Cost of Inflexibility' 166 *University of Pennsylvania Law Review*, 282 (2017).

[69] K.D. Werbach and N. Cornell, n 53 above, 318, 352.

[70] A. Savelyev, 'Contract law 2.0: Smart contracts as the beginning of the end of classic contract law' 26 *Information & Communications Technology Law*, 120, (2017).

[71] In this regard, see: P. De Filippi and A. Wright, 'Decentralized Blockchain Technology' n 51 above, 25.

[72] See, among others: M. Raskin 'The Law and Legality of Smart Contracts' 1 *Georgetown*

with the relative flexibility in terms of contractual performance.[73]

The ambiguity and editorial shortcomings can also be used by the parties who intend to free themselves from contractual conditions that they no longer want to honor.[74] Compared to this phenomenon, smart contracts provide a different binding option by incorporating legal provisions into the computer code.[75]

On the other hand, smart contracts also present a number of new issues and challenges for trade law and practice.

A first question that can arise is that of the identification of the other contracting party, considering that the Blockchain allows anonymous, or rather pseudonymous transactions,[76] such as in the case of transactions that are registered by referring to an IP address or cryptocurrency wallet.[77]

The codification of the clauses in computer language could lead to a limitation of the possible contents of smart contracts, linked with the possibilities of automation of the contractual prose according to the if / then logic.[78]

Connected to this is the risk that the code, drawn up by IT technicians, incorrectly reports the provisions of the contractual agreement between the parties, or that it may operate differently from what was planned, with the related issue of attributing liability.

Furthermore, there is the question of the risk of parties and legal operators misunderstanding of smart contracts.[79]

Moreover, in practice the connection between the text in computer code and a contractual text drawn up in natural language is increasingly widespread; the texts may have the same content, the so-called split contracting, which is being the specification and/or execution of the other, the so-called hybrid agreement.[80]

Law Technology Review, 324 (2017); E.A. Farnsworth "'Meaning" in the Law of Contracts' 76 *Yale Law Journal*, 939-965 (1967).

[73] In this regard, see: M.P. Gergen, 'The Use of Open Terms in Contract' 92 *Columbia Law Review*, 1006 (1992); G.K. Hadfield 'Judicial Competence and the Interpretation of Incomplete Contracts' 23 *Journal of Legal Studies*, 159-184 (1984).

[74] In this sense, see: S.J. Burnham et al, 'Transactional Skills Training: Contract Drafting-Beyond the Basics' *Transactions: The Tennessee Journal of Business Law*, 253-296 (2009).

[75] Thus: P. De Filippi and A. Wright, 'Decentralized Blockchain Technology' n 51 above, 25.

[76] By pseudonymity we mean the possibility that, although a person is not identifiable with his real name, such identification can still take place through the acquisition of further information about him, such as a pseudonym, an IP address, a current account, etc; on the subject, see: Article 29 Data Protection Working Party (2014) Opinion 05/2014 on Anonymisation Techniques, WP 216, available at https://tinyurl.com/yb8rz48p (last visited 30 June 2021).

[77] On this point, see: European Bank For Reconstruction and Development, Clifford Chance (2018) 'Smart Contracts: Legal Framework and Proposed Guidelines for Lawmakers', available at https://tinyurl.com/tac5274 (last visited 30 June 2021).

[78] Cardozo Blockchain Project (2018) Research Report 2: 'Smart Contracts' & Legal Enforceability, available at https://tinyurl.com/9dnauruj (last visited 30 June 2021).

[79] Regarding these profiles, see: M. Giancaspro, 'Is a Smart Contract Really a Smart Idea?' 33 *Computer Law and Security Review*, 830 (2017); E. Mik, 'Smart contracts', n 67 above, 281.

[80] On this point, see: European Bank For Reconstruction and Development, Clifford Chance

Contrary to traditional contracts where the parties can decide whether or not to fulfill their obligations, the smart contract cannot be violated from the point of view of execution.[81]

In a system governed by self-imposed smart contracts and other technical agreements, there would be less need for judicial intervention, since the computer code through which the rules were defined is the same tool through which they are applied.[82] This raises the question of what is legally or technically binding.

Although the implementation of basic contractual guarantees and consumer protection regulations in smart contracts is theoretically possible, in practice it can prove to be complex, given the formalized and deterministic nature of the computer code.[83]

Furthermore, although most of the data comes from the Blockchain or other databases connected to it, some smart contracts may have to acquire data from outside the Blockchain to be executed. This creates the need to make use of reliable external sources, the so-called oracles, which represent interfaces between contracts and the outside world.[84]

Reliable oracles that support and can satisfy a wide range of data requests are of paramount importance to many smart contracts.[85] On the other hand, this phenomenon requires the guarantee that the oracle is reliable and actually a third party, and that there is no interference or security threats during the acquisition of data from the same.[86]

Another problematic issue concerns the need to intervene on a smart contract in the event that an injunction issued by the judicial authority must be executed.

In general, given the impossibility of interrupting the execution of a smart contract – excluding the exceptions mentioned above – this result may be realized in the hypothesis of using a private Blockchain that provides mechanisms for

(2018), n 77 above; P. De Filippi and A. Wright, *Blockchain and the Law*, n 61 above, 76; J.P. Allen, 'Wrapped and Stacked: Smart Contracts and the Interaction of Natural and Formal Languages' 14 *European Review of Contract Law*, 307-343 (2018).

[81] Unless of course the parties could terminate the contract if they decide they do not want to remain tied to it.

[82] From the merger of law and code it follows, therefore, that the only way to violate the law is to effectively break the code.

[83] T. Cutts, 'Smart Contracts and Consumers, LSE Legal Studies Working Paper No. 1/2019, available at https://tinyurl.com/4yp7tkn7 (last visited 30 June 2021); P. De Filippi and A. Wright, 'Decentralized Blockchain Technology' n 51 above, 26.

[84] A case of smart contracts activated by external inputs is, for example, that of the insurance policies proposed by AXA and Etherisc, ie insurance companies that offer policies that compensate travelers who suffer flight delays or cancellations. Flight information is acquired automatically and in real time by an oracle company indicated in the contract and the compensation is paid automatically.

[85] In this sense, see: M. Sokolov, n 67 above, 10.

[86] E. Mik, 'Smart contracts', n 67 above, 292.

blocking the execution under the responsibility of certain nodes.[87]

The fields of application of smart contracts are numerous. They can be used, at least in theory, in all cases in which economic activities are correlated to the Internet and some events can be digitally verified.[88]

In addition to the financial and insurance sectors where digital bargaining already plays a central role, the use of smart contracts is developing in sectors such as agri-food, energy, entertainment, etc.[89]

Contracts that concern access to digital content and are thus easily translatable into software represent privileged use cases of smart contracts.[90] By virtue of the growing interconnection of devices, sensors, etc. through the Internet of Things this phenomenon affects ever wider areas.[91]

Devices and other material properties can be registered on a Blockchain and, by using smart contracts, transformed into 'smart properties', thus allowing the control of material properties on the network, even through other machines.

A Blockchain can store the relationship between Internet-enabled machines at any time and smart contracts can allocate the corresponding rights and obligations of connected devices.

Different relationships and credentials can also be encoded in the Blockchain regarding certain cryptographically activated resources, such as key blocks or smartphones, to ensure that only certain subjects or nodes can have access to the functionality of the property.[92]

V. Conclusion

This chapter investigated the evolution of the use of technology for the

[87] On this point, see: M. Giuliano, n 51 above.

[88] In this regard, see: D. Linardatos, n 64 above, 9; G. Governatori et al, 'On legal contracts, imperative and declarative smart contracts, and blockchain systems' 26 *Artificial Intelligence and Law*, 377-409 (2018).

[89] In this regard, see: Chamber of Digital Commerce – Smart Contracts Alliance (2018) n 56 above; R. Unsworth, 'Smart Contract This! An Assessment of the Contractual Landscape and the Herculean Challenges it Currently Presents for "Self-executing" Contracts', in M. Corrales et al, n 55 above, 17.

[90] This would allow the generalized implementation of a so-called metered Internet, where stocks are tied to micro-payments through related smart contracts. Since cryptocurrencies and smart contracts greatly reduce transaction costs, in particular, they allow artists, musicians, authors, etc to automatically collect royalties inherent to the copyrights on their works every time they are viewed or used. In this regard, see: P. De Filippi and A. Wright, 'Decentralized Blockchain Technology' n 51 above, 29; D.A. Wallach, 'Bitcoin for Rockstars: How Cryptocurrency Can Revolutionize the Music Industry', available at https://tinyurl.com/9nvdbsxm (last visited 30 June 2021).

[91] C. Wendehorst, 'Consumer Contracts and the Internet of Things', in R. Schulze and D. Staudenmayer eds, *Digital Revolution: Challenges for Contract Law in Practice* (Baden-Baden: Hart-Nomos, 2016), 189.

[92] On this point, see again: P. De Filippi and A. Wright, 'Decentralized Blockchain Technology' n 51 above, 14.

conclusion and execution of contracts. In this context, machines can considerably support contracting parties because they facilitate, accelerate, and make less onerous distant relationships.

Starting from the so-called automatic contracts, there has been an increase in the capacity of these means to replace human behaviours, thus becoming more and more autonomous. As a matter of fact, while telematic contracts only put in contact distant parties, smart contracts can even guarantee the self-execution of contracts without the need for human intervention.

Technological development has led to a new way of conceiving contracts. In particular, as stated above a process of depersonalisation of relationships has developed. Telematics and informatics have favoured the mass market and instant contracts to the detriment of the bargaining phase, which was the real essence of the agreement.

To face these changes, legal systems have tried to interpret old rules and adapt them to best fit the characteristics of online contracting. Think, for example, to the preference for the reception rule instead of the mailbox rule.

Additional rules have also arisen to accompany traditional contract law. They especially aim to ensure the protection of the weakest part of the contract, for instance with the application of the *jus poenitendi* to electronic bargaining, given that there is usually no time to weigh an agreement.

As already underlined, these means can bring many solutions but also some risks. Besides the advantages, smart contracts present also different challenges, as illustrated in the preceding section.

Smart contracts are the next frontier of contract automation. Recently, legal debates around smart contracts have increased because of the emergence of the Blockchain.

When smart contracts are based on a blockchain, they benefit from its characteristics such as decentralisation and immutability. It is considered that the features of blockchain technology better ensure the automatic performance of contracts, so having the potential to reduce disputes and overcoming the problem of lack of trust in electronic commerce and international contracts.

However, as with every new phenomenon, there is a need of evaluating the impact of smart contracts on contract law.

At the various levels, different initiatives have flourished to analyse smart contracts. To name a few, in 2019 the United Nations Commission on International Trade Law (UNCITRAL) and the International Institute for the Unification of Private Law (UNIDROIT) organised a joint workshop on this topic.[93]

In 2020, the European Commission published a report on blockchains that also focuses on the legal issues concerning smart contracts and contract law.[94]

[93] The workshop took place on 6-7 May 2019 in Rome. To see the summary of the discussions and conclusions of the workshop: urly.it/3d49w (last visited 30 June 2021).

[94] European Commission (2018).

The Commission responded to the invitation of the European Parliament, that with the Resolution of 3 October 2018 had stressed the need for the European Commission to undertake an in-depth assessment of the legal implications of smart contracts.[95]

At the national level, some countries have issued dedicated legislation. For example, Italy introduced a legislation specifically relating to technologies based on distributed ledgers and smart contracts,[96] unlike other countries which limited themselves to dictate provisions on specific[97] aspects.

According to the Italian law, smart contracts can meet the requirement of the written form after computer identification of the interested parties, through a process that shall meet the requirements set by the Agency for Digital Italy.[98]

In the same sense, the storage of a computer document through the use of technologies based on distributed ledgers produces the legal effects of the electronic time validation referred to in Art 41 of the EU e-IDAS regulation.[99]

All these actions have in common the search for appropriate legal answers to a continuously changing environment, with the ultimate goal of fostering economic development by encouraging the spread of digital technology without renouncing legal certainty.

[95] European Parliament resolution of 3 October 2018 on distributed ledger technologies and blockchains: building trust with disintermediation (2017/2772(RSP)) P8_TA-PROV(2018)0373. In particular, see paras from 36 to 38 of the Resolution. See also the European Parliament Resolution of 16 February 2017 on civil law rules on robotics (2015/2103 8INL)) P8_TA-PROV(20170051).

[96] Legge 11 February 2019 no 12, Art 8-*ter*, entitled 'Technologies based on distributed registers and smart contracts', converting decreto legge 14 December 2018 no 135. See eg: G. Finocchiaro and C. Bomprezzi, 'A legal analysis of the use of blockchain technology for the formation of smart legal contracts' *MediaLaws*, 111-135 (2020); one may also see: A. Stazi, *Automazione contrattuale e "contratti intelligenti"*, n 52 above, 129.

[97] In France, the Ordonnance no 2016-520 du 28 avril 2016 relative aux bons de caisse and the Ordnance no 2017-1674 du 8 décembre 2017 relative à l'utilisation d'un dispositif d'enregistrement électronique partagé pour la représentation et la transmission de titres financiers, allowed the use of the Blockchain for the registration and transfer of unlisted financial securities as an alternative to the traditional registration in accounting and corporate books. In this regard, see: R3, Norton Rose Fullbright 'Can smart contracts be legally binding contracts?', available at urly.it/3d49z (last visited 30 June 2021); L.D. Muka Tshibende, 'Contract Law and Smart Contracts: Property and Security Rights Issues' 26 *European Review of Private Law*, 874 (2019).

[98] Through guidelines to be adopted within ninety days from the date of entry into force of the law, but which so far have not been adopted. The provision is in line with Art 20, c. 1-*bis*, of the Digital Administration Code, contained in the decreto legislativo no 82/2005, which establishes the conditions for which an electronic document is suitable to satisfy the requirement of the written form.

[99] Regulation (EU) no 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC, OJ L 257 of 28 August 2014. In the abbreviation «e-IDAS», where «e» stands for «electronic» «ID» for «identification», «A» for «authentication» and «S» for «signature». See eg: J. Dumortier, 'Regulation (EU) No 910/2014 on Electronic Identification and Trust Services for Electronic Transactions in the Internal Market (eIDAS Regulation)', available at urly.it/3d4b9 (last visited 30 June 2021).

Essays

The Trust Experience in San Marino Between *Ius Commune* and International Models

Daniele Mantucci and Luca E. Perriello*

Abstract

San Marino trust law is embedded in a consolidated civil law tradition stretching back to the *ius commune* system of fiduciary instruments, thereby making it possible to trace, to a large extent, an itinerary related to common law trusts, and to challenge unwarranted allegations (now, fortunately, fading away) that trusts cannot be transplanted into civil law countries. San Marino has not confined itself to imitating other offshore legislation, but has drawn up a unique trust system reconciling the typical features of international models – thus embarking on the race to attract the trust business within its borders – with the peculiarities of its own system of sources. Hovering between the principles of confidence and patrimonial separation, the international models and *ius commune*, San Marino trust law proves to be the perfect combination of innovation and tradition, and longs to become a benchmark for the regulation of trusts established in civil law jurisdictions.

I. San Marino Trust Law and Its Interaction with the Hague Trust Convention

It has been fifteen years since San Marino adopted a written law on trusts (enforced by legge 17 March 2005 no 37, amended by legge March 2010 no 42). Just a few months earlier, San Marino had ratified the Hague Convention of 1 July 1985 on the law applicable to trusts and their recognition (hereinafter referred to as the Hague Convention),[1] that speeded up the process of internationalization of the Republic and contributed to dismissing its misrepresentation as 'an area escaping innovation due to a marginalization tantamount to isolation'.[2] Notwithstanding the enforcement of a domestic law, San Marino has not opted out of the Convention, as Art 3 of legge no 42 of 2010 overtly refers to it for purposes of identifying the governing law and recognizing foreign trusts. San Marino

* Daniele Mantucci is Full Professor of Private Law at Marche Polytechnic University. Luca E. Perriello is Assistant Professor of Private Law at Marche Polytechnic University. This paper is the result of a joint effort of the two authors. However, paragraph I is to be attributed to Daniele Mantucci, while paragraphs II-VI are to be attributed to Luca E. Perriello.

[1] Ratified with decreto consiliare 20 September 2004 no 119. In Italy, the Hague Trust Convention was ratified with legge 16 October 1989 no 364, which came into force on 1 January 1992.

[2] This way, R. Sacco, 'Introduzione al diritto comparato', in R. Sacco ed, *Trattato di diritto comparato* (Torino: UTET, 1997), 149. Unless specified otherwise, all texts in Italian have been translated by the authors.

lawmakers were astute enough to avoid restricting party autonomy to the extent of forcing its own trust model upon parties and preventing them from relying on structures governed by foreign laws. Accordingly, San Marino will continue recognizing trusts regulated, for example, by English or Jersey law.

However, the merits of San Marino legislation do not boil down to matters of choice of law or the interaction with the Hague Convention. The legislation of the tiny state offers much food for thought to scholars (especially those with a historical-comparative background) and professionals alike. San Marino trust law is embedded in a consolidated civilian tradition stretching back to the *ius commune* system of fiduciary instruments, thereby making it possible to trace, to a large extent, an itinerary related to common law trusts, and to challenge unwarranted allegations (now, fortunately fading away) that trusts cannot be transplanted into civil law countries. On its way to regulating trusts, San Marino has also accommodated many elements of the 'international' trust model, launched by the Jersey (Trusts) Law 1984 and adopted by several offshore jurisdictions.[3] Unlike other systems, San Marino has not duplicated Jersey law, but has drawn up a unique trust system reconciling the typical features of international models – thus embarking on a race to attract the trust business within its borders –[4] with the peculiarities of its own system of sources. San Marino is so aware of its uniqueness, that new legislation concerning a different institution, the fiduciary trusteeship (*affidamento fiduciario*), was passed with legge 1 March 2010 no 43, based on witty, albeit contested, scholarship.[5]

With San Marino being a country geographically and culturally very close to Italy, its trusts legislation may also provide Italian professionals with a valuable benchmark. Being the first law in the world on trusts written in Italian, legge no 42 of 2010 may facilitate the understanding of a number of trust issues, and be designated by Italian settlors as the law governing their trusts. There is no denying that the settlor's autonomy to choose the applicable law (Art 6 Hague Convention) may be directed at San Marino law, whose trust model is in line with the features laid down in Art 2 Hague Convention, amounting to a

[3] Jersey codified its own trust law, which is quite distinct from the traditional English model, not only because Jersey law is based on Norman-French customs and lacks an equity system comparable to the English one, but also because the codifiers sought to boost the offshore trust business already rooted in the island: P. Matthews, 'La legge sul trust a San Marino e il modello di trust internazionale' *Contratto e impresa*, 251 (2007). Many offshore jurisdictions drew inspiration from the Jersey (Trusts) Law 1984, such as Anguilla, Belize, Dubai, Grenada, Guernsey, Labuan, Malta, Mauritius, Nevis, Niue, Seychelles, St Kitts & Nevis, Turks & Caicos.

[4] R. Pardolesi, 'Destinazioni patrimoniali e trust "internazionale" ' *Rivista critica del diritto privato*, 215, 221 (2008) argues that trusts have triggered competition between legal systems, that do not communicate but compete to attract trust business within their boundaries.

[5] The leading theorist on the fiduciary trusteeship contract is M. Lupoi, *Il contratto di affidamento fiduciario* (Milano: Giuffrè, 2014). For strongly critical remarks, see A. Vicari, 'L'affidamento fiduciario quale contratto nominato: un'analisi realistica' *Contratti*, 357, 362 (2018), claiming that Lupoi's theory appears to be decontextualized from civil law categories, using them in a rhetorical and provocatively heretical way.

legal relationship created – *inter vivos* or on death – by a person, the settlor, when assets have been placed under the control of a trustee for the benefit of a beneficiary or for a specified purpose.

II. The Centrality of *Ius Commune* in San Marino Legal System

It is impossible to understand San Marino trust law fully without framing it within the sources of law of the Republic. The articulation of San Marino sources of law is complex[6] but harmonious, and makes it possible for its law effectively to adapt to social changes. Formally, at the apex of the pyramid lies the *ius proprium*, ie the Statute and the *reformationes*, followed by local customs and *ius commune*.[7] This is true, however, from a hierarchical point of view, whereas, from the standpoint of the application of law, *ius commune* takes priority in regulating private relations, giving rise to a principle-based law, which is much more 'plastic' than a narrow legislation consisting of detailed rules.[8] San Marino has not codified private or commercial law, which, on the one hand, may emphasize judicial discretion and challenge legal certainty through discrepancies in judgments, but on the other, may ward off the danger of redundant legislation, which, on the contrary, is a distinguishing mark of the nearby Italian system.[9] It is true that every codification entails breaking with the past and laying down a self-standing text whose loopholes may not be filled through resort to other systems, not even scholarly opinions,[10] but the lack of a codification in San Marino has shown the merit of facilitating the assimilation of trusts, unlike what has happened in Italy.

Ius commune is the main source of San Marino private law. Once that *ius commune* ceased to be effective in Germany with the entry into force of the

[6] Contending that the San Marino legal system is based on a 'multiple regulatory competence': S. Caprioli, 'Satura lanx 30. Linee sammarinesi per lo studio del diritto comune' *Studi in onore di Pietro Rescigno* (Milano: Giuffrè, 1998), I, 221.

[7] Art 3 *bis* legge 8 July 1974 no 59, added by Art 4 legge costituzionale 26 February 2002 no 36, para 6, reads that 'customs and common law constitute an integrative source in the absence of legislative provisions'. Considering that the constitutional recognition of *ius commune* as a source of law is an 'epochal change': S. Caprioli, 'Per una lettura della Costituzione sammarinese riformata' *Giurisprudenza italiana*, 914, 917 (2004).

[8] L. Di Bona, 'Trust e affidamenti fiduciari nel confronto tra «modello» sammarinese e italiano' 5 *Cultura giuridica e diritto vivente*, 1, 6 (2018).

[9] ibid 9. Giudice delle Appellazioni, G. Astuti, 30 July 1963, *Giurisprudenza sammarinese*, I, 46 (1965), remarks that '*ius novum*, consisting of statutory law and subsequent local legislation, does not amount to a complete codification of private, civil, commercial and civil procedural law'. See also S. Caprioli, 'Il diritto comune nelle esperienze di San Marino' *Rivista internazionale di diritto comune*, 90 (1994), arguing that 'San Marino *ius commune* cannot be understood without referring to its contrary, ie the civil code'. On the failed attempt to codify San Marino private law, undertaken at the end of the nineteenth century, see C. Pecorella, 'Un codice mancato' 3(2) *Archivi per la storia*, 113 (1990).

[10] P. Peruzzi, *Appunti per le lezioni del corso di diritto Sammarinese* (Urbino: Quattroventi, 1998), 96.

BGB in 1900, San Marino remained the only state in Europe to preserve its force of law.[11] San Marino *ius commune* does not coincide with Justinian's law but with the law that in the Middle Ages developed across the most evolved systems of continental Europe, Italy in particular, under the influence of Roman law, canon law and customs, and that can be found 'in the writings of the most authoritative jurists and the decisions of the most renowned courts'.[12] San Marino lawyers are invited to consult the 'writers (...) of the XVI, XVII and XVIII centuries: that is, the time when *ius commune* was in its greatest splendor', prioritizing 'practical writers' over 'connoisseurs or theoreticians'.[13] *Ius commune* does not distinguish between practical and theoretical works.[14] The problem, however, is the absence of a modern academic thought and an authentic San Marino school of *ius commune*, which may interpret, as well as innovate the rules handed down by tradition.[15]

If the *lex posterior* criterion were to govern the succession of laws in time, the new legislation would abrogate the ancient *ius commune*, yet San Marino applies a different criterion, that is, the new legislation derogates from *ius commune*; put differently, it does not abrogate the previous rule but simply makes it inapplicable to the case at issue. When the new legislation is no longer applicable, *ius commune* becomes applicable again. The derogating rule does not determine abrogation but 'quiescence' of the derogated rule, which is intended to revive as soon as the former ceases to apply. The abrogation of *ius commune* can only result from an explicit provision in the new legislation.[16]

The foregoing may thus further the bold idea that European *ius commune* has not died out as a result of the coming into force of the codes, for these replaced *ius commune* only in matters explicitly regulated, whereas, in all the others, *ius commune* has survived the age of codification, insofar as it is in accordance with

[11] V. Scialoja, 'Nota a App. Roma 1 dicembre 1906' *Rivista di diritto internazionale*, 154 (1907), in a dispute concerning citizenship.

[12] Giudice delle Appellazioni, V. Scialoja, 12 August 1924, *Giurisprudenza sammarinese*, 18 (1924).

[13] G. Ramoino, *Le fonti del diritto privato Sammarinese* (San Marino: Arti Grafiche F. Della Balda, 1928), today in *Le fonti del diritto privato Sammarinese* (San Marino: Banca Agricola Commerciale, 2000), 19.

[14] P. Peruzzi, n 10 above, 149.

[15] On this matter see S. Caprioli, *La legislazione societaria sammarinese* (Rimini: Maggioli Editore, 1990), 13 and 28, claiming that 'in the dialogue between citizens, lawmakers and courts, the voice of scholarship resounds sporadically; its polyphony, which was one the key features of the historical common law, has died out'. See also V. Crescenzi, 'La rilevanza dell'opinione dei giuristi negli attuali ordinamenti di diritto comune: Andorra e San Marino' *Rivista di diritto civile*, 129, 148 (1995), contending that San Marino case law is now exclusively the one decided by the courts, as the other source, the academic one, which may well perform a practical, humble, and vital function of maintenance of the system, has failed.

[16] This happened, for instance, within family law, with the reform enacted with legge 26 April 1986 no 49, which expressly repealed 'all the rules (...) including the *ius commune* ones' at variance with the new legislation.

the general principles of the system.[17] Rules of different sources intertwine in the dense web of San Marino diachronic, multi-secular and stratified legal system, which academics and practitioners may find hard to disentangle,[18] all the more so if they come from a codified system such as the Italian one, in which lawyers seldom look at history and, when they do, do so without any perspective of normative value.

Ius commune feeds on historical memory and shared traditions. Yet, it is not obsolete law, nor a re-edition of Roman law modernized or common law handed down intact from the *Ancien Régime*.[19] That San Marino *ius commune* builds on collective conscience implies that

> deciding a case today as it would have been decided in the seventeenth century would distort its spirit and take away its value, and thinking that the relations between sovereign power, citizens and foreigners have stayed motionless, so discretion may be exercised in ways that in other countries would be arbitrary, would fail San Marino's historical conscience and its tradition of freedom.[20]

Drawing on that historical conscience, legal interpretation in San Marino becomes *interpretatio*, and it is no coincidence that this was widely practiced in the classical age of *ius commune* but later dismissed in the modern age of codification. *Interpretatio* is not merely cognizant of enunciations in their meaning (this was *exposition*) but is determinative 'of rules, given other rules'.[21] In this attitude may be found an extraordinary affinity of *ius commune* with English equity, which has always been understood as a jurisdiction of conscience, capable of transforming the core values of social coexistence into certain rules and

[17] This way M. Lupoi, *I trust nel diritto civile* (Torino: UTET, 2004), 197, arguing that 'throwing away a complex of wisdom and centuries of experience cannot be beneficial, not even when one wants to give an unequivocal and even forced signal of rupture with the previous age'. Critical comments by F. Treggiari, 'Trust e diritto comune a San Marino', in F. Treggiari et al, *Il trust nella nuova legislazione di San Marino* (Santarcangelo di Romagna: Maggioli Editore, 2005), 47, fn 28: 'if we agree that "matters" are neither the general areas of codified private law, nor the single institutions that are outlined there, but rather all the objects that can be linked to the titles of the books of the code (...), the room for a *directly positive* common law of trusts – and therefore for a *by-pass* common law of trusts – is drastically reduced'.

[18] S. Caprioli, n 15 above, 24. See also M. Simoncini, 'Abrogazione ed altre vicende delle norme nello stile della legislazione sammarinese' *Miscellanea dell'Istituto Giuridico Sammarinese*, 121 (1993); and, as regards the reform of company law, U. Santarelli, 'Cinque lezioni sul diritto comune delle società' *Miscellanea dell'Istituto Giuridico Sammarinese*, 36 (1991).

[19] A. Landi, *Note a margine di un recente convegno sul diritto comune vigente*, available at tinyurl.com/jyc9r474 (last visited 30 June 2021).

[20] Giudice delle Appellazioni, V. Scialoja, 12 August 1924, n 12 above.

[21] S. Caprioli, *Lineamenti dell'interpretazione* (San Marino: Banca agricola commerciale, 2008), 31. See also A. Landi, n 19 above, describing *ius commune* as 'a legal experience which, by continuous judicial interpretation, with its own sensitivity and that common conscience of which Jemolo spoke, still uses with profit the worthwhile normative product of a centuries-old system'.

adequate remedies, though constantly evolving and adapting to ever new situations, ultimately turning 'right' into 'legal'.[22]

It is precisely by relating to shared values that San Marino *ius commune*, kept up-to-date through the *interpretatio* determinative of rules, preserves the vibrancy that codifications, linked to a precise historical moment, have lost.[23] The absence in San Marino of a codification of private and commercial law, and of the relating straits of *analogia legis*, along with the possibility of drawing from the pool of *ius commune* principles, makes it possible for courts to decide cases with a due sense of proportionality and reasonableness, take account of the parties' interests and pursue substantial justice rather than formal legality.[24]

If these premises are correct, *ius commune* cannot be understood as an appendix or a compendium of the Statute or the *reformationes*, as subsidiary or supplementary law which should be applied only under exceptional circumstances, that is, where there is no legislation or the existing legislation is deficient. It is exactly the other way around; *ius commune* is the rule, while legislation is the exception, firstly, from a quantitative point of view, because of the greater number of *ius commune* rules, and secondly, because of *ius commune* being supplementary, as well as innovative of legislation in the regulation of private law institutions.[25] Even in the Italian republics of the Middle Ages, *ius commune* co-existed with the statutes, as a source that was formally subsidiary, but in reality very broadly regulative of anything that had been overlooked by the statutes, and at the same time innovative of the system.[26]

Ius commune does not only make up for deficient legislative texts; nor does it step in only when legislation neglects a case or dictates obscure provisions to be interpreted.[27] Even when legislation does cover a case, *ius commune* may operate concurrently with 'new' legislation in those areas that might be

[22] In this respect see M. Lupoi, 'English "Equity" and the Civil Law – A Tale of Two Worlds' *Trusts & Trustees*, 176, 180 (2020), remarking that the equity court was originally known as 'court of conscience', as the Chancellor 'purported to come to the aid of justice and did so in the manner that was the most becoming for a shepherd of souls: calling upon the *conscience* of the affected parties'.

[23] V. Pierfelici, 'I rapporti fiduciari in San Marino nella pratica notarile e giudiziaria' *Trusts*, 537, 544-545 (2015).

[24] L. Di Bona, n 8 above, 9. See also V. Pierfelici, 'La Corte per il trust a San Marino' *Trusts*, 5, 9 (2016), arguing that the San Marino Court for Trusts 'should not only be the guardian of the compliance of a given solution with law, but also implement the parties' intention through a just solution'.

[25] Giudice delle Appellazioni, G. Astuti, 30 July 1963, n 9 above. See also G. Guidi, *Le fonti scritte nella Repubblica di San Marino* (Torino: Giappichelli, 2004), 159, claiming that the relationship between *ius commune* and statutory law is based on competition and subsidiarity.

[26] V. Scialoja, 'Nota a App. Roma 1 dicembre 1906' n 11 above.

[27] For the application of *ius commune* on tort liability in the absence of legislative provisions on industrial property and unfair competition, see Giudice delle Appellazioni, G. Astuti, 20 September 1965, *Giurisprudenza sammarinese*, 150 (1964-1969); and, in matters of joint ownership, Commissario della legge G. Ramoino, 3 May 1965, *Giurisprudenza sammarinese*, 268 (1964-1969).

regulated by alternative institutions.[28]

III. Historical-Comparative Report of Trust and *Fiducia*, and the Impact of International Models

One of these areas is trusts, which, in San Marino, co-exist with the *ius commune* fiduciary institutions. The reasons underlying this co-existence are rooted in *ius commune*, across which one may trace an itinerary largely shared between the common law trust and the civil law *fiducia*.[29]

In the fourteenth century, while the English courts of equity recognized the legal value of the obligations undertaken by the trustee, continental European commentators developed a solid model of testamentary *fiducia*, whereby the fiduciary heir was instructed by the testator to pass on to the beneficiary what was bequeathed through succession. The *fiducia* reflected a genuine legal obligation, in keeping with the idea that the fiduciary, while being the owner of the assets, received them only to implement the program outlined by the testator. The fiduciary was understood as a '*nudus minister a commodo sed non a titulo*',[30] in that he obtained the title to the property but could not receive any advantage therefrom, and undertook the obligations of custody and retransfer, enforceable by the law. This mechanism made sure that the fiduciary's creditors could not have recourse against the assets transferred to him, thereby producing the ring-fencing effect which distinguishes the modern trust from fiduciary relationships of merely obligatory nature. In England too, trusts were subject to Roman influence, for most jurists sitting in the Court of Chancery studied Roman and canon law in continental universities.[31]

The divorce between trust and *fiducia* took place later, when the Pandectist school, dusting off the classical Roman fiduciary law, shamefully overlooked the *ius commune* contribution,[32] eventually handing over a construction of *fiducia*

[28] As highlighted by F. Treggiari, n 17 above, 44.

[29] See M. Graziadei, 'The Development of *Fiducia* in Italian and French Law from the 14th Century to the End of the *Ancien Régime*', in R. Helmholz and R. Zimmermann eds, *Itinera Fiduciae: Trust and Treuhand in Historical Perspective* (Berlin: Duncker und Humblot, 1998), 327; Id, 'La fiducia nella tarda età moderna. Le "confidenze" tra vincolo di coscienza e disciplina politica dei soggetti e dei beni', in P. Prodi ed, *La fiducia secondo i linguaggi del potere* (Bologna: il Mulino, 2008), 235; Id, 'Trust, confidenza, fiducia', in R.H. Helmholz and V. Piergiovanni eds, *Relations Between the* Ius Commune *and English Law* (Roma: Rubbettino, 2009), 223. Arguing that *trust* is 'part of a pan-European tradition': P.H. Glenn, 'The Historical Origins of the Trust', in A.M. Rabello ed, *Aequitas and Equity: Equity in Civil Law and Mixed Jurisdictions* (Jerusalem: Hebrew University, 1997), 775.

[30] Baldo degli Ubaldi, *Commentaria in secundam Digesti veteris partem*, Venice, 1572, f. 192va-rb, in D. 22, 1, 3, 3; Id, *In sextum Codicis librum commentaria*, Venice, 1599, f. 146ra, in C. 6, 42, 12, n. 7.

[31] M. Lupoi, n 22 above, 178.

[32] M. Graziadei, 'Fiducia e trust in Italia', in M.L. Biccari ed, *Fiducia, Trusts, Affidamenti. Un percorso storico-comparatistico* (Urbino: Università degli Studi di Urbino, 2015), 362.

as an obligatory relationship between a settlor and a fiduciary.[33] This reflected in the weak notion of testamentary *fiducia* laid down in Art 627 Italian Civil Code (whereby the beneficiary is not given action to establish that the disposition of property upon death was actually made to his own advantage) and the common misconception that sees the fiduciary as a figurehead of the settlor; in a nutshell, a 'distrust in the trust'. On the contrary, the settlor of a trust definitively foregoes the ownership of his assets and transfers them to the trustee, and no obligatory relationship arises between them. As a consequence, the trustee is not an agent acting in the name or on behalf of the settlor, is liable only to the beneficiaries and, should he have doubts about the exercise of his powers, he could only turn to the judiciary. This explains why trusts are not, as a rule, revocable (not so the mandate: Art 1723 Italian Civil Code), and why trusts do not expire upon the death of the settlor or the trustee (not so the mandate: Art 1722 Italian Civil Code). These decisive remarks, on a historical-comparative and legal level, militate against equating trust with *fiducia* (which, according to leading authority, is a development of the mandate).[34]

The fiduciary element of the trust may not be dismissed, provided it is understood as 'confidence' (*affidamento*), which does not amount to the confidence placed by the settlor in the trustee, but to the 'commission of a right to the trustee so that he can advance certain interests or purposes either through or as a consequence of this right'.[35] Such confidence justifies the loss of ownership on the part of the settlor and the destination of the trustee's title to the beneficiaries' interests or a given purpose. Confidence gives normative content to the limitations of the trustee's proprietary position and makes him directly liable to the beneficiaries (and not to the settlor).[36]

[33] The fiduciary agreement is intended as 'a manifestation of intention through which one transfers to others a right of ownership in one's own name but in the interest, or also in the interest, of the transferor or a third party. The attribution of the assignee is full, but he undertakes a mandatory obligation in order to the destination or use of the asset transferred': C. Grassetti, 'Del negozio fiduciario e della sua ammissibilità nel nostro ordinamento giuridico' *Rivista del diritto commerciale*, 345, 363 (1936).

[34] M. Graziadei, n 32 above, 353, argues that the *fiducia* theory in Italy developed in parallel to the mandate theory. For some critical remarks on the conflation of the *fiducia* with the mandate, see L.E. Perriello, 'Unitarietà causale, proprietà confermata e tutela reale: verso una lettura rafforzata della fiducia' *Rassegna di diritto civile*, 421, 431 (2019), arguing that *fiducia* has a programmatic attitude, that is, 'the fiduciary is not a simple agent-manager, a mere executive appendage of the settlor, but the direct interpreter and implementer of the program, holding powers that are not exhaustively predetermined, but proportionate to the actual circumstances'. See also F. Alcaro, 'Il programma contrattuale: l'attività dell'affidatario fiduciario e i rapporti fra le parti', in F. Alcaro et al, *Contratti di convivenza e contratti di affidamento fiduciario quali espressioni di un diritto civile postmoderno* (Milano: I Quaderni della Fondazione Italiana del Notariato, 2017), 163.

[35] M. Lupoi, *Trusts* (Milano: Giuffrè, 2nd ed, 2001), 307. Similarly see Corte di Cassazione 25 February 2015 no 3886, *Vita notarile*, 386 (2015).

[36] The trustee's liability to beneficiaries is an essential element of trusts: see D. Waters, 'The Concept Called "the Trust" ' *Bulletin for International Fiscal Documentation*, 118, 124 (1999),

While the separation of trust and *fiducia* occurred in codified civil law systems, it did not take place in San Marino. The historical background of fiduciary instruments under *ius commune* has facilitated the introduction of trusts in San Marino much more than elsewhere. San Marino trust co-exists with the hereditary *fideicommissum*, which presents the fiduciary characteristics outlined above. This institution, although obsolete,[37] is still in force. When, in 1902, the Regent Captains asked Vittorio Scialoja for an opinion on the existing *fideicommissa*, the renowned scholar advised against their abolition, and the Council followed suit.[38] There seems little doubt that, in commercial practice, trusts attract larger fortunes compared to *ius commune* hereditary *fideicommissa*, because investors prefer relying on the extensive body of codified rules and remedies under legge no 42 of 2010. Yet, *ius commune* maintains its peculiar hermeneutic function even when parties enter into a trust agreement, for it may facilitate the interpretation of trust legislation, at least when it comes to general civil law notions, whereas the specific trust notions can only be drawn from the Hague Convention and the trust models in use in common law or mixed jurisdictions.[39]

As a consequence, the influence of the international models on San Marino trust law is considerable, including the express definition of what a trust is, rules on its duration, the trustee's powers and duties, and the guardian, the possibility of settling trusts without beneficiaries, thus making San Marino compete with other jurisdictions in the race to attract foreign capitals and investments. However, not all of the international models has been transplanted. For example, San Marino has decided to set up a trust register,[40] which is unusual in common law jurisdictions, because, unlike companies, the trust is not a legal person and publicity is not seen as an instrument for the protection of third parties but as 'a disgrace, a violation of privacy'.[41] The civil law imprint of the San Marino legal system has required many other adjustments in order to

arguing that '(f)rom the moment of the creation of the trust there must be an ability of the beneficiary to secure an accounting or, to put it another way, a power in the beneficiary to enforce the discharge of his duties'.

[37] V. Pierfelici, n 23 above, 538, remarks that recent practice knows no example of *fideicommissa*. 'The reasons are probably to be found in the progressive detachment of the practice from the categories of *ius commune* and concurrent imitation of Italian models, which are uncritically assumed to be identical to those of San Marino'.

[38] This episode is reported by F. Treggiari, n 17 above, 69, challenging G.B. Curti-Pasini and E. Ranza, *Principi elementari del diritto privato della Repubblica di S. Marino* (Bollate: Zappa, 1939), 58, who claim that the *fideicommissum* has not been abrogated.

[39] See F. Treggiari, n 17 above, 49 and M. Graziadei, 'Prima lettura delle disposizioni civilistiche contenute nella legge di San Marino sul trust', in F. Treggiari et al, *Il trust nella nuova legislazione di San Marino* n 17 above, 17.

[40] The register collects all the trust instruments governed by San Marino law and foreign trust having their seat in San Marino. See E. Montanari, 'La trasparenza dei titolari effettivi dei trust' *Trusts*, 310 (2015).

[41] P. Matthews, n 3 above, 254.

reconcile the trust with the system of fiduciary instruments handed down by *ius commune* and with the principle of patrimonial separation. The international model is not unique, but fragmented in the various identities of the legal systems that are inspired by it.

IV. The Confidence Principle in San Marino Trust Law: The Trustee's Powers and Duties, Self-Declared Trusts, Reserved Powers and the Guardian

The confidence principle, intended as commission of a given right to a fiduciary to advance a program beyond the settlor's control, inspires the notion of trust laid down in Art 2(1) of San Marino law, which focuses on the ownership 'of assets in the interest of one or more beneficiaries, or for a specific purpose'. Wisely, San Marino lawmakers have not replicated the controversial Art 2 Hague Convention, which understands the trust as the placement of assets under the 'control' of a trustee. By doing so, the Convention accepts the possibility that the settlor does not transfer rights but rather the 'control' of assets, thereby creating a mandatory relationship with the trustee, who would have to account for his actions to the settlor. Many 'fiduciary' relationships characterized by the direct protection of the settlor *vis-à-vis* the trustee, which have little to do with the traditional trust model,[42] in which the settlor's detachment is an essential element, are thus drawn into the scope of the

[42] M. Lupoi, n 35 above, 501. See also H. Kötz, 'The Hague Convention on the Law Applicable to Trusts and Their Recognition', in D. Hayton ed, *Modern International Developments in Trust Law* (London: Kluwer Law International, 1999), 37, 40, claiming that the conventional definition encompasses the relationships that in common law jurisdictions are known as 'trusts', but the Convention is also applicable to the institutions of many civil law countries where they are not known as trusts, despite performing similar functions. However, it is not sufficient for the institution at hand to be merely 'functionally analogous'; it must also be 'structurally similar', which requires that the assets constitute a distinct mass and are not part of the trustee's estate. These requirements are met, by way of example, by the *fideicomiso* of several Latin American legal systems and the Quebec *fiducie*. A.E. von Overbeck, 'Rapport explicatif/Explanatory Report' *Conference de La Haye de droit international privé – Hague Conference on Private International Law, Actes et documents de la Quinzième session – Proceedings of the Fifteenth Session*, II, § 26 (1985), mentions the analogous institutions of Egypt, Japan, Luxembourg and Poland. For critical comments see D. Hayton, 'International Recognition of Trusts', in D. Hayton ed, *The International Trust* (Bristol: Jordan, 3rd ed, 2011), 165; and Id, 'Reflections on The Hague Trusts Convention after 30 Years' *Journal of Private International Law*, 1, 8 (2016), taking the view that only with superficiality Art 2 Hague Convention can be construed as extending beyond proprietary relationships (ownership-management of assets) including agency relationships (agency-management of assets). He makes the example of an owner giving 'control' of his assets to an agent in his own interest. The agency or mandate relationship is not covered by the Convention because Arts 2 and 11 clearly provide that the assets in trust are in the name of the trustee and constitute a mass distinct from the rest of his assets. When, however, the assets are placed under the control of an agent, they remain in the settlor's name. Where title is not transferred to the trustee but remains with the settlor on whose behalf the trustee administers the assets with powers of representation, the trust is not covered by the Convention.

Convention. San Marino has not made the same mistake as the Convention; indeed, it has not even expressly referred to a fiduciary element when outlining the characteristics of the relationship,[43] thus avoiding any misunderstanding in the assimilation of the trust with a fiduciary agreement.

Under San Marino law, the powers and duties of a trustee are laid out in such a way as not to turn him into a fiduciary. Accordingly, the trustee 'exercises over the trust property all the powers belonging to the right-holder, except for the limitations resulting from the trust register' (Art 31(1)). The rule does not even give way to the widespread representation of the trust as a *patrimoine d'affectation*, which was a ruse of the French private law scholarship in the 1930s[44] to facilitate the approval of the foreign institution by the adverse continental jurists. This was quite a misrepresentation, for the trustee is not a mere custodian of assets according to the destination established by the settlor, but a full and exclusive owner, and has the same powers that a *dominus* would have in his own interest.[45] The trust fund is not a collection of 'things', but 'wealth'. The trustee can use, replace, transfer any objects of the trust, having to confer to the beneficiaries not this or that asset but their value.[46] The trustee has the power to perform all acts of ordinary and extraordinary administration as full and absolute owner of the trust fund, as well as active and passive procedural capacity. The 'open-ended' formulation of the trustee's powers is typical of modern trust legislation. The limitation of the trustee's powers to those expressly provided for in the trust instrument belonged in the past and sought to ensure maximum protection of the beneficiaries' rights, when wealth was predominantly non-financial and, therefore, not easy to dispose of. Modern trust laws, on the contrary, afford the trustee the widest management powers, outweighed by tight 'fiduciary' duties,[47] such as good faith and diligence (Art 20), independence (Art 23), impartiality (Art 24), confidentiality (Art 25) and information (Art 27). To enforce the trustee's obligations, San Marino law lays

[43] M. Graziadei, n 39 above, 21.

[44] P. Lepaulle, *Traité théorique et pratique des trusts en droit interne, en droit fiscal et en droit international* (Paris: Rousseau & Cie, 1931), 39.

[45] M. Lupoi, 'Trust e vincoli di destinazione: qualcosa in comune?' *Trusts e attività fiduciarie*, 237, 240 (2019). Comparing the civil law representation of the trust as a destination to a program defined by the settlor with the common law representation of the trust as a gift to beneficiaries, see A. Vicari, 'La scelta della legge regolatrice dei trust: una questione di *Principia* beneficiari' *Trusts e attività fiduciarie*, 364, 369 (2011).

[46] Cf B. Rudden, 'Things as Thing and Things as Wealth' *Oxford Journal of Legal Studies*, 81 (1994). See also P. Matthews, 'The Compatibility of the Trust with the Civil Law Notion of Property', in L. Smith ed, *The Worlds of the Trust* (Cambridge: Cambridge University Press, 2013), 316, icastically noting that '(t)he trustee is (…) the *owner* of the trust assets in the most complete sense possible. The trustee is *not* an agent or a representative of the settlor or the beneficiaries. (…) (T)he trustee's ownership is no less than the ownership of a person who is not a trustee'.

[47] Cf J.H. Langbein, 'The Contractarian Basis of the Law of Trusts' *Yale Law Journal*, 625, 640 (1995), arguing that 'the substitution of fiduciary law for law restricting the powers of the trustee (is) a central event in the development of modern trust law'.

down a set of criminal sanctions (Art 57 ff), which may even include imprisonment.

However, San Marino trust law presents a few elements that appear to dilute the confidence principle. One of these is the possible coincidence of settlor and trustee (Art 2(2)). The so-called self-declared trust, in which the settlor acts as trustee, separating a subset of assets from his own general assets, invites suspicion from part of the Italian tax courts, because

> although it is called a trust, it does not have the same features; in fact, it lacks one of its typical features, ie the transfer by the settlor to third parties of the assets settled on trust.[48]

These conclusions appear to be hasty and ill-advised. The self-declared trust complies with the confidence principle, for the trustee undertakes fiduciary obligations towards third parties, who will then have the right to a diligent performance by the trustee. The unilateral declaration of trust falls within the scope of the Hague Convention, which does not discriminate as to the way in which the trust is settled and, once the self-declared trust is validly created, the issues concerning its governing law and its suitability for recognition are the same as those of an ordinary trust.[49]

Actually, what may hamper the confidence principle is the possibility for the settlor of reserving rights or powers to himself (Art 2(2)), which San Marino law does not restrict, thereby leaving the settlor a large amount of leeway. Another rule provides for the reservation of the power to revoke the trust (Art 14), which may be justified by the settlor's reluctance to lose control of the trust property permanently. Revocable trusts are unusual in England, while the offshore operators of the international trust tend to suggest to replace the power to revoke the trust with the power to appoint beneficiaries, including the settlor,

[48] Corte di Cassazione 24 February 2015 no 3735, *Notariato*, 207 (2015); Corte di Cassazione 24 February 2015 no 3737, *Foro italiano*, 1215 (2015); Corte di Cassazione 25 February 2015 no 3886, *Vita notarile*, 386 (2015). By contrast, contending that self-declared trusts may be fully recognized: Corte di Cassazione 26 October 2016 no 21614, *Trusts*, 66 (2017); recently, see Corte di Cassazione 7 June 2019 no 15456 and Corte di Cassazione 21 June 2019 no 16700, available at dejure.it.

[49] J. Harris, *The Hague Trusts Convention. Scope, Application and Preliminary Issues* (Oxford-Portland: Hart, 2002), 106; L. Tucker, N. Le Poidevin and J. Brightwell, *Lewin on Trusts* (London: Sweet & Maxwell, 19th ed, 2014), 528, claiming that it is 'inconceivable (…) that such trusts, which once created are no different from those created by transfer to a trustee, were intended to be excluded from the Convention'. In the Italian literature see M. Lupoi, *Istituzioni del diritto dei trust negli ordinamenti di origine e in Italia* (Padova: Kluwer, 3rd ed, 2016), 240-241, noting that not only does Art 2 Hague Convention not specify by whom the property must be placed under the control of the trustee, but by not requiring transfer to a trustee, Art 2 supports the inclusion of the self-declared trust within the scope of the Convention. Lupoi articulated a different, but doubtful, opinion in *Trusts* n 35 above, 504, arguing that the Convention requires that settlor and trustee be different parties, and the reference to the settlor in Art 2 was inserted at the request of the civil law delegates, whereas the common law ones took it for granted that trusts in which settlor and trustee coincide are covered by the Convention.

because financial administrations may re-qualify the trust and disown the transfer of property to the trustee, while trusts with the power to appoint beneficiaries are generally irrevocable, making it possible to preserve the integrity of the trust without depriving the settlor of the power to add beneficiaries.[50] Moreover, the power of the settlor to revoke the trust is a matter of concern in many common law jurisdictions because it conflicts with the Norman customary maxim '*donner et retenir ne vaut*'.[51]

The provision of reserved powers does not affect the existence or validity of a San Marino trust; it does not turn the settlor into a trustee, nor does it entail that a trustee acting in accordance with the reservation commits a breach of trust. The reservation is also permitted by the Hague Convention, whose Art 2(3) reads,

> the reservation by the settlor of certain rights and powers, and the fact that the trustee may himself have rights as a beneficiary, are not necessarily inconsistent with the existence of a trust.[52]

The Convention does not provide any criteria for narrowing the boundaries of the reservation, that is, for establishing how far the settlor may go in determining his powers without clashing with the conventional notion of trust, being the legal relationship whereby assets are 'placed under the control of a trustee for the benefit of a beneficiary or for a specified purpose' (Art 2). Nevertheless, Italian tax authorities insist on declaring the fiscal non-existence of trusts in which the settlor has reserved 'control over the trust assets to himself in such a way as to preclude the trustee from fully exercising his disposition powers' and in which 'the trustee's management and disposition powers, as established by the trust instrument or by law, are in some way limited or even simply conditioned by the settlor's intentions'.[53]

Italian courts too are very wary of trusts with wide-ranging reservations of powers and tend to deny their recognition (most cases concerned asset-protection

[50] D. Harris, 'No Such Thing as a Sham Trust' *Private Client Business*, 95, 98 (2004). Neglecting these arguments altogether and apodictically deeming trusts subject to later appointment of beneficiaries to be 'radically void or, at most, under development, in which no separation of assets takes effect until beneficiaries are named': D. Muritano, 'Il nuovo art. 2929-*bis* c.c.: quale futuro per la protezione del patrimonio familiare?' *Rivista di diritto bancario*, 1, 18 (2015).

[51] A. Dejardins, 'Recherche sur l'origine de la règle «donner et retenir ne vaut»' *Revue critique*, 207 (1868). Jersey law expressly reads that 'the rule *donner et retenir ne vaut* shall not apply to any question concerning the validity, effect or administration of a trust, or a transfer or other disposition of property to a trust' (Art 9(5)), thus providing for its abrogation.

[52] D. Waters, 'The Concept Called "the Trust" ' *Bulletin of International Fiscal Documentation*, 118, 120 (1999), contends that 'the Convention appears as reluctant as the present writer to say that a reserved power of control or disposition of the trust property, automatically makes the trust a "sham" '.

[53] Agenzia delle Entrate, Circolare no 61/E of 27 December 2010. Defining this document 'devoid of legal basis and indeed visibly (at odds) with the applicable rules': M. Lupoi, 'Il "controllo" in materia di trust, auto-dichiarato e non' *Trusts e attività fiduciarie*, 121, 127 (2020).

trusts settled to the detriment of creditors).[54] This way, however, they seem to turn a blind eye to the clear provisions of the Convention and the common law jurisdictions, where trusts are not invalidated simply because they contemplate retention of rights or powers or because the settlor has drawn up a letter of wishes, for courts focus on the conduct of the trustee, who could meet the settlor's requests based on a correct and independent determination. If the settlor has made a request and the trustee, in good faith and in the exercise of independent judgment, after considering all the circumstances of the case, has decided to grant that request, the trust is not sham.[55] This was also the view taken by San Marino Trust Court, which, after correctly refusing to narrow the concept of sham down to that of simulation,[56] held that

> the trustee must take into account the intention of the settlor, declared in the trust instrument or inferable from the provisions of the trust instrument or even subsequently expressed, obviously considering it not as a binding instruction, but as a manifestation of a desire, the fulfillment of which remains with the prudent appreciation of the trustee.[57]

Accordingly, the distinction does not lie in the amount of powers reserved,[58] but in whether the trustee attaches importance to the independence of the exercise of his powers, or if he routinely ignores the trust instrument considering himself to be a front man. When it comes to the reservation of the power to revoke the trust, a sham allegation may be raised where the trustee, knowing that the settlor could terminate the trust at any moment, considers himself to be a puppet of the settlor, slavishly following his instructions even if they are contrary to the interests of the beneficiaries or to the purpose of the trust.[59]

Cast in these terms, the reservation of powers may be reconciled with the confidence principle. In civil law systems, the risk of a court denying the recognition of a trust with reserved powers remains, but could be averted by providing for an additional office, the guardian of the trust, whose appointment,

[54] For an analysis of the Italian case-law see L.E. Perriello, *Lo* sham *trust nell'ordinamento giuridico italiano. Meritevolezza degli interessi e tecniche di tutela* (Napoli: Edizioni Scientifiche Italiane, 2017), 207.

[55] *Grupo Torras SA* v *Al Sabah* [2003] JRC 092; *Shalson* v *Russo* [2003] EWHC 1637 (Ch); *Charman* v *Charman* [2005] EWCA Civ 1606; *Kan Lai Kwan* v *Poon Lok To Otto* (2014) 7 HKCFAR 414. See also G. Thomas and A. Hudson, *The Law of Trusts* (Oxford: Oxford University Press, 2nd ed, 2010), 65.

[56] See L.E. Perriello, n 54 above, 212.

[57] Corte per il trust (Repubblica di San Marino) 5 December 2017, *Foro italiano*, 163 (2018).

[58] This way, instead, A. Braun, 'The Risk of "Misusing" Trusts: Some Lessons from the Italian Experience' *European Review of Private Law*, 1119 (2016), remarking that 'the fact that (the settlor) maintained some control over the assets transferred to the trustee, for instance, in the form of a right to live in the trust property, is not conclusive' (at 1132).

[59] G. Thomas and D. Hayton, 'Shams, Revocable Trusts and Retention of Control', in D. Hayton ed, *The International Trust* n 42 above, 604-605.

under San Marino law, is mandatory in purpose trusts and voluntary in trusts with beneficiaries (Art 52). The 'guardian' does not have a single legal meaning in the common law galaxy[60] and is differently referred to in the legislation as 'enforcer', 'protector', 'nominator', 'committee', etc. He may be chosen either from among the members of the settlor's family or his friends, so as to be the ideal interpreter of his and the beneficiaries' intentions, or from among persons with particular professional skills (which is more desirable because of the possible conflicts that could arise between the settlor and persons close to him and lead to a fall-out).[61] The guardian may be the settlor himself (even more so if the settlor can also act as trustee).[62] However, the office of guardian is incompatible with that of trustee. The controller must be differentiated from the controlled. Indeed, a guardian with too invasive powers over the trust management may be re-qualified in court as trustee, and take on the relating responsibilities.[63]

The guardian must oversee the proper administration of the trust fund. He is entitled to take action against the trustee in the event of default (Art 52(1)). Other powers attributable to the guardian include the appointment or revocation of trustees, beneficiaries or other guardians; the amendment of the law governing the trust; the veto on certain acts of the trustee (Art 52(5)). The guardian's consent affects the trustee's standing, removing an obstacle to the exercise of a given power, without prejudice to the trustee being sovereign in deciding whether or not to perform an act.[64] Indeed, Art 52(6) provides that the exercise of the above powers does not confer on the guardian the office of trustee. Nor should his consent exempt the trustee from liability, otherwise 'the very notion of trusteeship would be undermined and the trustee would become a sort of manager, in joint ownership with the guardian'.[65]

Unless the trust instrument provides otherwise, the powers of the guardian are fiduciary and not personal (Art 52(3)), meaning that their exercise may not be waived, must necessarily be directed to the benefit of the beneficiaries of the trust and not to the holder of the power, and be subject to judicial review.[66] The

[60] See, for instance, the Bahamian case *Rawson Trust Co* v *Perlman* [1990] 1 Butterworths OCM 31, for Justice Smith asserting that 'the term protector is not a term of art and is not known as such to our law'.

[61] E. Campbell et al, 'Protectors', in D. Hayton ed, *The International Trust* n 42 above, 196.

[62] M. Lupoi, 'Il "controllo" in materia di trust, auto-dichiarato e non' *Trusts e attività fiduciarie*, 121, 123 (2020). See also E. Campbell et al, n 61 above, 196, fn 8, deeming this to be the best way to ensure that the settlor retains some form of control over the trustee.

[63] M. Lupoi, n 35 above, 404.

[64] A. Di Sapio, 'Riflessioni su un provvedimento genovese' *Trusts e attività fiduciarie*, 639, 645 (2019). D. Hayton, P. Matthews and C. Mitchell, *Underhill and Hayton, Law of Trusts and Trustees* (London: LexisNexis, 19th ed, 2016), 58, note that the trustee must refuse to comply with orders of the guardian if he believes that they amount to a breach of his fiduciary duties or a breach of trust on the part of the trustee.

[65] M. Lupoi, n 35 above, 402.

[66] E. Campbell et al, n 61 above, 199. See also *Re Bird Charitable Trust and Bird Purpose Trust* [2008] JLR 1, where the Jersey court asserted its 'very wide powers to supervise and control'

'fiduciary' guardian[67] must, from time to time, with independence and due information, assess whether or not to exercise the power granted to him in accordance with the interests of the beneficiaries, never indulging the whims of the settlor or pursuing personal advantages.[68] This implies, for example, that the guardian may not appoint himself, or persons related to him by family or friendship, as trustee or beneficiary,[69] or direct the trustee to sell him trust property, which may amount to conflict of interests.

V. San Marino Trust as a Separate Patrimony, the Trustee's Liability and the Lack of Legal Personality

The confidence principle, and its many applications, is carved into San Marino trust model, where it co-exists with the principle of patrimonial separation, a purely civilian doctrine. The representation of the trust as a separate patrimony is typical of civil or mixed law systems,[70] and is central to the definition of 'trust' contained in the Draft Common Frame of Reference,[71] but it is not shared by pure common law models, which are not even familiar with the concept of patrimony.[72] The property settled on a common law trust does not constitute a

the exercise of fiduciary powers. Another interesting case discussed by the Jersey court was *The M Settlement* [2009] JRC 140, concerning a request by the settlor and guardian of a trust to pay his massive personal debts with trust money. Quite rightly, the trustee had refused on the ground that the settlor was an alcoholic and that the trust fund was in any case insufficient to prevent the settlor's bankruptcy, and the settlor, disappointed, had decided to exercise his fiduciary power to replace the trustee with a long-time friend. The court suspended the settlor/guardian's power and ordered that the trust be dissolved and the remainder be allocated to his children.

[67] It is accepted that nothing prevents certain powers of the guardian from being qualified as 'personal': M. Conaglen and E. Weaver, 'Protectors as Fiduciaries: Theory and Practice' *Trusts & Trustees*, 17, 20 (2012).

[68] On this point see D. Hayton, P. Matthews and C. Mitchell, n 64 above, 51. Cf also the following decisions: *Re Hay's Settlement Trusts* [1981] 3 All ER 786; *Re Manisty's Settlement* [1974] Ch 17; *Schmidt* v *Rosewood Trust Ltd* [2003] AC 709; *Centre Trustees Ltd* v *Pabst* [2009] JRC 109.

[69] Advocating for the fiduciary nature of the power to appoint and remove trustees: *Re Skeats Settlement* (1889) 42 Ch D 522; *Mettoy Pension Trustees* v *Evans* [1990] 1 WLR 1587; *Simpson Curtis Pension Trustees Ltd* v *Readson Ltd* [1994] OPLR 231.

[70] The Scottish mixed model, for example, exemplifies the tendency to 'patrimonialize' the trust: K. Reid, 'National Report for Scotland', in D.J. Hayton, S.C.J.J. Kortmann and H.L.E. Verhagen eds, *Principles of European Trust Law* (Den Haag-Deventer: Kluwer, 1999), 67; Id, 'Patrimony Not Equity: The Trust in Scotland' *European Review of Private Law*, 427 (2000); G. Gretton, 'Trust and Patrimony', in H. MacQueen ed, *Scots Law into the 21st Century* (Edinburgh: Sweet & Maxwell, 1996), 182; Id, 'Trusts Without Equity' *International & Comparative Law Quarterly*, 599 (2000).

[71] X. – 1: 202: 'the trust fund is to be regarded as a patrimony distinct from the personal patrimony of the trustee and any other patrimonies vested in or managed by the trustee'. For a commentary see A. Braun, 'Trusts in the Draft Common Frame of Reference: The "Best Solution" for Europe?' *Cambridge Law Journal*, 327 (2011).

[72] P. Matthews, n 3 above, 254, points out that the exemption of trust assets from the action of the trustee's personal creditors is not a consequence of trust law, but of the law on the enforcement of judgments and bankruptcy. See also M. Lupoi, 'Si fa presto a dire "trust" ' *Trusts e attività*

patrimony as might be understood, but a fund, made only of assets. Trust assets are not part of the trustee's patrimony (and his personal creditors cannot have access thereto), unlike trust liabilities, for which the trustee is personally liable, ie with his own patrimony,[73] without prejudice to his right to seek indemnity from the trust fund.[74] The trust fund is also made immune to claims for liabilities incurred by the trustee for purposes relating to the trust. By separating assets and not liabilities, the common law trust amounts to an 'asset partitioning tool'.[75] It is true that the trustee may agree to an exclusion of liability, so that the trust liabilities may be enforced against the trust fund,[76] but this agreement entitles the trust creditor to subrogation to the rights of the trustee, that is, the trust creditor does not have action when the trustee has exceeded his powers or committed a breach of trust.[77]

However, the San Marino legal system differs considerably from the traditional model, for it conceptualizes the trust as a separate patrimony composed of assets *and* liabilities. The doctrine of patrimonial separation is so entrenched in San Marino *ius commune* on fiduciary instruments,[78] that Art 1(j) defines the trust fund as a 'complex of assets in trust and the legal relations inherent to them', thereby including assets as well as liabilities. Art 12 states that the trust fund is separate from the trustee's personal property; it cannot be subject to claims

fiduciarie, 585, 586 (2017), holding that patrimonial separation 'must not be considered fundamental by the English since no one except Underhill mentions it'.

73 L. Smith, 'Trust and Patrimony' *Estates, Trusts and Pensions Journal*, 332, 338 (2009); D.J. Hayton and C. Mitchell, *Hayton & Marshall Commentary and Cases on the Law of Trusts and Equitable Remedies* (London: Sweet & Maxwell, 12nd ed, 2005), 693; J. Penner, *The Law of Trusts* (Oxford: Oxford University Press, 2006), 22.

74 'A trustee (a) is entitled to be reimbursed from the trust funds or (b) may pay out of the trust funds, expenses properly incurred by him when acting on behalf of the trust': s. 31(1) Trustee Act 2000. Defining this mechanism as 'clumsy and formalistic ritual (which serves) no functional purpose': R.H. Sitkoff, 'Trust Law as Fiduciary Governance plus Asset Partitioning', in L. Smith ed, *The Worlds of the Trust* n 46 above, 436. In fact, a different model is gaining currency in the United States, which considers the trustee as a representative of the trust property, thus allowing creditors to have direct recourse against the trust fund. The requirement for the exclusion of the trustee's personal liability is that he has disclosed his status to the third party and has not violated the law or the trust instrument. This trend brings trusts closer to corporations and has been enshrined in the Restatement (Second) of Trusts, the Uniform Probate Code and the Uniform Trust Code. On this matter, cf J.D. Johnston Jr, 'Developments in Contract Liability of Trusts and Trustees' *New York University Law Review*, 483 (1966); G.G. Bogert and G.T. Bogert, *The Law of Trusts & Trustees* (St. Paul-Minneapolis: Thomson/West, 1980), § 712; J. Dukeminier and S.M. Johanson, *Wills, Trusts and Estates* (New York: Aspen, 2000), 975; A. Gallarati, *Il trust come organizzazione complessa* (Milano: Giuffrè, 2010), 174-175, for a list of US states where the traditional English model based on the trustee's personal liability has been dismissed.

75 Arguing this way and comparing the common law trust with the San Marino trust based on patrimonial separation: A. Vicari, 'Country Reports: San Marino' *Columbia Journal of European Law Online*, 81, 91 (2012).

76 *Muir* v *City of Glasgow Bank* (1879) 4 App Cas 337.

77 L. Smith, n 73 above, 340-341. See also *Re Johnson* (1880) 15 Ch D 548; *Ex p Garland* (1804) 10 Ves 110; *Re Frith* [1902] 1 Ch 342; *Re British Power Traction and Lighting Co* [1910] 2 Ch 470.

78 F. Treggiari, 'Trust e diritto comune a San Marino' n 17 above, 63-64.

from the trustee's personal creditors, nor does it form part of the trustee's succession, matrimonial property or insolvency proceedings. This is a minimum effect of the recognition of the trust (Art 11(2) Hague Convention),[79] that is, if it is not provided for in the applicable law, the trust cannot be recognized; conversely, if it is provided for in the applicable law, the separating effect cannot then be questioned by the authorities of the state in which the trust takes effect. Strictly speaking, patrimonial separation is not an essential element of the trust under the governing laws, but an essential effect of its recognition in non-trust countries.

In line with the representation of San Marino trust as a separate patrimony, encompassing any legal relationships relating to the trust, whether active or passive, is the rule requiring that creditors for obligations undertaken by the trustee in his capacity as trustee be satisfied solely out of the trust fund (Art 47(1)). Accordingly, not only are the trust assets separated from the trustee's, but so also are the relating liabilities (the trust appears to be also a 'liability partitioning tool'). The trustee is liable to third parties only with the assets settled on trust, not with his own assets. The reason behind this is that the trustee holds a private law office and to this end he is the owner of the trust assets; just as he cannot enrich himself (Art 23), neither can he impoverish himself.[80] If he discharges the trust liabilities personally, he will have recourse against the trust fund, before any other creditors (Art 47(2)). Unlike the traditional model, wherein the trustee is personally liable for the trust liabilities, the San Marino model affords greater protection to the trustee and encourages those who are afraid of exposing their assets to loss, to take on the trustee office, bringing the trust closer to the organizations with legal personality.[81]

The comparison with legal persons should not, however, be over-emphasized. Legal personality may help consider the trust assets as distinct from the trustee's personal assets, as is envisaged for the organizations endowed with corporate personality, so that creditors can never claim the trustee's liability for debts incurred by reason of his office, while his personal creditors have no reason to attack the trust fund. In fact, when continental private law scholars began to come to terms with the trust, they took the view that '*la solution la plus efficace et la plus simple est de doter le trust de la personne morale*'.[82] In several common law jurisdictions, trusts can be used to carry out non-profit or business activities

[79] J. Harris, n 49 above, 317.

[80] A. Vicari, n 45 above, 373.

[81] A. Gallarati, '*Fiducie* v trust. Spunti per una riflessione sull'adozione dei modelli fiduciarî in diritto italiano' *Trusts e attività fiduciarie*, 238, 249 (2010). Discussing modern US trust law, R.H. Sitkoff, n 74 above, 436, remarks that 'because modern law sharply separates the property of the trustee personally from the property of the trust, the contemporary American trust is in function (though not in juridical form) an entity. Reifying the trust in expression is an embrace of substantive function over technical form'.

[82] 'The most effective and simple solution is to endow the trust with legal personality': P. Lepaulle, 'Review of Roberto Pasqual's *La Propriété dans le* Trust' *Revue internationale de droit comparé*, 377, 378 (1952).

(the so-called Massachusetts trust) as an alternative to corporations, they can go bankrupt or have legal standing.[83] In Italy, in addition to registering a trust as a charity,[84] it is possible to open a bank account in the name of a trust, register the sales entered into by the trustee, report trust incomes, and/or treat a trust as a legal person for the purposes of anti-money laundering legislation.[85]

These are, however, forms of fictitious and strictly instrumental personality, for the trust, *per se*, is not a legal person. There is no such thing as a trust that contracts, commits crimes, takes legal action or pays taxes.[86] Personalizing the trust means thwarting its particularities and conflating it with already known entities, while history has taught us that if individuals use trusts, it is because their goals could not be equally pursued through the law of contracts[87] or legal persons.[88]

VI. The Worthiness of the Trust Program

Hovering between the principles of confidence and patrimonial separation, the international models and *ius commune*, San Marino trust proves to be the perfect combination of innovation and tradition. Legge no 42 of 2010 longs to become a benchmark for the regulation of trusts established in civil law jurisdictions. The choice of San Marino law cannot, however, pander to trusts contrary to mandatory rules, public policy or good morals (Art 10(1)(a)) or simulated (Art 10(1)(e)). The legislation, therefore, requires a review of the 'worthiness' of the transaction. On the contrary, the Italian Supreme Court – which, it is hoped, San Marino courts will not follow – held that the Hague Convention

> has given recognition in our legal system, if we can say so, (to the trust), so it is not necessary for courts to determine from time to time whether the single

[83] G. Gretton, 'Up there in the *Begriffshimmel*?', in L. Smith ed, *The Worlds of the Trust* n 46 above, 529.

[84] N.D. Latrofa, 'Dal trust charitable al trust ente del Terzo settore' *Trusts e attività fiduciarie*, 27 (2020).

[85] A. Vicari, n 45 above, 376.

[86] D. Hayton, P. Matthews and C. Mitchell, n 64 above, 16.

[87] On the autonomy of contract law from trust law, from a law & economics perspective, see H. Hansmann and U. Mattei, 'The Functions of Trust Law: A Comparative Legal and Economic Analysis' *New York University Law Review*, 434 (1998).

[88] L. Smith, n 73 above, 354, contends that 'the "entification" of the trust spells, in the long run, the end of the law of trusts by assimilation'. See also G. Gretton, n 83 above, 530, claiming that 'to turn trusts into persons is to abolish the trust, while at the same time adding an extra item to the list of species of the genus "juristic person" '. Concurring A. Zoppini, 'Fondazioni e trusts (spunti per un confronto)', in I. Beneventi ed, *I trusts in Italia oggi* (Milano: Giuffrè, 1996), 147, who, while comparing foundations and trusts, states that 'it does not really facilitate the understanding of the institution to cast the trust as a "surrogate" for legal personality'.

trust complies with the requirements laid down in Art 1322 Civil Code,[89]

under which

> the parties can also make contracts that are not of the types that are particularly regulated, provided that they are directed to the realization of interests worthy of protection according to the legal order.[90]

A commentator followed suit, arguing that 'the very concept of "interests worthy of protection" is alien, incomprehensible and (...) antithetical' to the law of trusts.[91]

The Italian Supreme Court is probably concerned that the concept of trust 'worthiness' will end up providing courts with an argument for continuing to deny the recognition of trusts or artificially re-qualifying them according to categories of the *forum* that have little to do with trusts.[92] However, this concern is not justified, as the arguments that, years ago, the 'pre-comprehension' doctrine[93] used to hinder the transplant of trusts in civil law countries, were not based on their being 'unworthy' of legal protection but on the alleged splitting of the right of ownership in defiance of the *numerus clausus* of real rights[94] and on the

[89] Corte di Cassazione 19 April 2018 no 9637, *Trusts e attività fiduciarie*, 504 (2018).

[90] Translation by S. Beltramo ed, *The Italian Civil Code and Complementary Legislation* translated in 1969 by M. Beltramo, G.E. Longo and J.H. Merryman (New York: Thomson Reuters, 2012).

[91] M. Lupoi, 'I trust, i flussi giuridici e le fonti di produzione del diritto' *Trusts e attività fiduciarie*, 5, 9 (2019). Lupoi seems to have reconsidered the position he took in Id, n 35 above, 549: 'a trust chooses the interest to protect among several conflicting interests and (...) this choice, consistent with the idea that the legal system has of what is worthy and what is not, results in a selection of interests which is forbidden by our traditional legal instruments. (...) A domestic trust removes a legal relationship from national legislation because only in this way can the protection mentioned above be obtained; a removal which is not undue, but worthy as worthy is, in each case, the interest to protect'.

[92] This is the questionable view embraced by G. Petrelli, 'Trust interno, art. 2645 ter c.c. e «trust italiano»' *Rivista di diritto civile*, 167 (2016), claiming that, after the introduction of Art 2645 *ter* in the Italian Civil Code, Italy has become a fully-fledged trust country. As a consequence, a domestic trust could not escape the application of the new provision; most importantly, it can only concern immoveable or registered movable property and must have determined beneficiaries.

[93] Defined so by M. Lupoi, 'Le ragioni della proposta dottrinale del contratto di affidamento fiduciario' *Contratto e impresa*, 734, 735 (2017). Elsewhere, Lupoi complains about the ignorance of the average lawyer on trusts 'because he did not study it at university, no longer reads legal journals and does not attend refresher courses on the subject; and this also applies to the average judge, no matter whether civil, criminal or tax, whether of merit or legitimacy. Our Universities do not offer refresher courses on trusts and it's been years since the High Council of the Judiciary held courses on trusts; the same (or even more) can be said for the National Council of Lawyers, the National Council of Notaries and the National Council of Chartered Accountants and Accounting Experts': Id, n 45 above, 237.

[94] Cf F. Weiser, *Trusts on the Continent of Europe* (London: Sweet & Maxwell, 1936), 7; H. Motulsky, 'De l'impossibilité juridique de constituer un trust anglo-saxon sous l'empire de la loi française' *Revue critique de droit international privé*, 451 (1948); H. Battifol, 'Trusts – The Trust Problem as Seen by a French Lawyer' *Journal of Comparative Legislation and International Law*, 18 (1951); P. Hefti, 'Trusts and Their Treatment in the Civil Law' *American Journal of Comparative*

close-ended formulation of the limitations to patrimonial liability.[95] However, the Hague Convention itself makes room for the 'worthiness' of the program underlying the trust, by permitting contracting states to refuse recognition of a trust, the significant elements of which are more closely connected with non-trust countries (Art 13). This provision is embedded in a complex system of checks that makes the judicial review of trusts particularly tight, preventing the recognition in non-trust countries of trusts running counter to 'provisions (that) cannot be derogated from by voluntary act' (Art 15), 'provisions of the law of the forum which must be applied even to international situations, irrespective of rules of conflict of laws' (Art 16), and public policy (Art 18).

The only way to avoid an *interpretatio abrogans* of Art 13 is to construe it as a 'wrap-up' provision in the Hague Convention, preventing the recognition of trusts that do not fall short of Arts 15, 16 and 18, but still have repugnant consequences in the legal system.[96] This way, Art 13 may sanction situations in which the use of trusts in non-trust countries has taken place without any 'reasonable and legitimate justification',[97] or

> not according to reasonableness and/or *bona fide* and/or the protection

Law, 553 (1956); A. Gambaro, 'Problemi in materia di riconoscimento degli effetti dei *trusts* nei paesi di *civil law*' *Rivista di diritto civile*, 93 (1984). Recently, see F. Fimmanò, 'La Cassazione "ripudia" il trust concorsuale' *Fallimento*, 1156, 1169 (2014), contending that 'our right of ownership is intended in such a way that the owner holds all powers of enjoyment, management and disposal of property. The trust, generating a doubling of the right (dual ownership), or rather a dissociation between ownership and control, should be considered a kind of atypical real right. Since real rights are predetermined and recognized by the Civil Code (*numerus clausus*), the free formation of new conventional situations is not allowed'.

[95] C. Castronovo, 'Il trust e "sostiene Lupoi" ' *Europa e diritto privato*, 441, 447 (1998); G. Broggini, 'Il trust nel diritto internazionale privato italiano', in I. Beneventi ed, *I trusts in Italia oggi* n 88 above, 11; F. Gazzoni, '(Lettera aperta a Maurizio Lupoi sul trust e su altre bagatelle)' *Rivista del notariato*, 1247, 1251 (2001).

[96] Art 13 Hague Convention is the extreme remedy 'offered when, notwithstanding Arts 18, 16 and 15, the modalities or the purposes of the trust are found by the court to be repugnant to a legal system (not necessarily the forum) which is not familiar with that particular form of trust, but in which, nevertheless, the trust has its main effects: Art 13 prevents the risk that a trust may succeed in producing repugnant effects despite all the conventional defenses': M. Lupoi, n 35 above, 545. This opinion was endorsed by the Tribunal of Bologna, 1 October 2003, *Trusts e attività fiduciarie*, 67 (2004), which held that 'since the "domestic" trust cannot be considered invalid *ex se* due to the lack of foreign elements (…), nor to its contrast with overriding mandatory rules or public policy (safeguarded by Arts 15, 16, 18, which, however, concern the effects of a trust already recognized), the only possible and reasonable hermeneutical solution (unless we want to give Art 13 an *interpretatio abrogans* of Arts 6 and 11) is to consider the provision as a "closing rule of the Convention" (comparable to Art 1344 of the Civil Code), aiming to grasp cases which escape rules of a specific nature: in other words, Art 13 amounts to an extreme and exceptional remedy provided for cases in which the modalities and purposes of a trust, whose effects escape the provisions of Arts 15, 16 and 18, are in any case considered by a court to be repugnant to a legal system which does not know that particular form of trust, but in which, nevertheless, the agreement actually carries out its effects'.

[97] R. Luzzatto, '«Legge applicabile» e «riconoscimento» di trusts secondo la Convenzione dell'Aja del 1° luglio 1985' *Rivista di diritto internazionale privato e processuale*, 5, 20 (1999).

> of lawful interests, but with the sole aim of abusively removing the situation in which the trust operates from the law that would be applicable thereto according to the ordinary application of the rules of private international law.[98]

A trust governed by San Marino law, which does not present any program worthy of protection[99] and pursues the sole aim of putting assets out of the reach of creditors,[100] cannot be recognized in Italy, especially now that the Republic of San Marino has undertaken to make its economic and financial system more transparent and abolish the legal institutions which may be used to perpetrate fictitious interpositions and conceal the ownership of assets.[101]

[98] S.M. Carbone, 'Autonomia privata, scelta della legge regolatrice del trust e riconoscimento dei suoi effetti nella Convenzione dell'Aja del 1985' *Rivista di diritto internazionale privato e processuale*, 773, 782-783 (1999).

[99] San Marino's most dated case-law shows that fiduciary institutions must attain interests worthy of legal protection. See the decision of the Commissario della Legge 4 September 1936 in the civil case no 33 of 1936, unpublished but cited by V. Pierfelici, n 23 above, 537, fn 1, in a case wherein the testator had addressed to the universal heir 'a special recommendation or rather obligation never to abandon but always to help and assist in the best way possible his sister Ester who, due to illness, is incapable of earning a living'.

[100] Claiming that today the trust 'is a favorite legal coding device among the wealthy who wish to protect their assets from tax authorities and other creditors', see K. Pistor, *The Code of Capital. How the Law Creates Wealth and Inequality* (Princeton and Oxford: Princeton University Press, 2019), 43. A purely asset-protection trust, which does not enunciate any program, may not be recognized in Italy.

[101] V. Pierfelici, n 23 above, 544, mentions significant reforms such as 'the abolition of bearer shares, savings accounts and bearer financial instruments, the revision of banking secrecy, the strengthening of control and vigilance instrument'. All these measures highlight how 'external confidentiality is protected and safeguarded, but is no longer functional to "hide", to "conceal", to create areas of opacity in which to shelter capitals of unlawful origin'. For a review of San Marino anti-money laundering legislation see E. Montanari, 'Antiriciclaggio nella legislazione della Repubblica di San Marino: adeguata verifica della clientela e identificazione del titolare effettivo' *Trusts e attività fiduciarie*, 164 (2015).

Innovation Partnerships and Italy's Participation in the European Space Economy Plan

Leonello Marraccini*

Abstract

The present study intends to analyse the development process of the Space Economy, firstly at EU level, so as to subsequently examine the characteristics, especially the legal ones, that characterise the Italian Plan for the Space Economy, for the implementation of which the Innovation Partnership was used, in particular for the enactment of the Mirror GovSatCom Programme.

I. Introduction: The Central Role of the Innovation Partnerships for the Italian Space Economy

A study of the evolution of the European Union shows that throughout its short history, economic aspects have regularly played a key role in the process of internal integration, supported by the promotion and testing of innovative ways of solving problems common to the Member States.

With this perspective, numerous projects of general interest have been launched at EU level in strategic sectors, aimed at boosting the competitiveness of the economic system as a whole, among which the experiences of international cooperation in the aerospace sector, including the space economy, are particularly important.

The Space Economy is to be interpreted as the production process which, from the outset, begins with research, development and implementation of space infrastructures, ie the 'upstream' sector, where the associated innovative products and services can be realised, ie the respective 'downstream' sector, which may include, for example, services for environmental monitoring and weather forecasting.

For the development of the Italian Space Economy, the innovation partnership procedure has played a key role.

The innovation partnership is an entirely new procedure compared to the range of procedures provided for by the previous rules, which allows the public authority to develop innovative products, services or works in collaboration with a private economic operator, and then to purchase the result, without the

* PhD in Comparative Law, University of Campania 'L. Vanvitelli'.

need to give rise to a new procedure.

This is a scheme of EU origin, provided for by Art 31 of Directive 2014/24/EU and Art 49 of Directive 2014/25/EU, transposed in Italy by Art 65 of decreto legislativo 18 April 2016 no 50.

However, as early as 2010, the European Commission described a new model of intervention in the areas of research and innovation in the Communication 'Europe 2020 Flagship Initiative: Innovation Union', promoting interventions based on partnerships of national and European regional players involved in the entire chain of research and development activities, through cooperation between public and private operators in order to

> support innovation in areas that represent challenges for European society, such as climate change, energy efficiency, food security, health and an ageing population.

These policies are the result of economic studies carried out in the United Kingdom since the 1970s, which gave impetus to a real innovation policy based on a different and revolutionary way of understanding public demand: the so-called PPI, Public Procurement for Innovation.[1]

The above-mentioned Anglo-Saxon economic studies have shown that it is not essential to intervene with public subsidies to stimulate innovation policies, but it is much more effective to stimulate the interaction between supply and demand through public procurement, which is the main instrument of innovation policy.[2]

Public authorities are able to take more risks related to the use of non-established technologies, high upfront and transition costs or market fluctuations. This applies to PPI aimed at implementing the services provided by the contracting authority as well as to innovation to satisfy collective interests, such as energy saving or environmental protection.[3]

The European Union could therefore be considered as the soil in which these studies and theories have been able to be developed in practice, through the implementation of wide-ranging interventions on issues of Community importance.

In particular, the promotion of the 'Common Agricultural Policy' reform strategy of 2012 has led to the launch of the 'Agricultural Productivity and Sustainability' European Innovation Partnership, which aims to promote competitive and sustainable agricultural systems capable of providing effective solutions to the problems identified as a result of the analyses carried out on the real needs of farmers and the level of innovation in the field.

[1] J. Edler and L. Georghiou, 'Public procurement and innovation – Resurrecting the demand side' *Research Policy*, 7 (2007).

[2] E. von Hippel, 'The dominant role of users in the scientific instrument innovation process' *Research Policy*, 3 (1976).

[3] C. Edquist and J.M. Zabala-Iturriagagoitia, 'Public Procurement for Innovation as mission-oriented innovation policy' *Research Policy*, 10 (2012); B. Aschoff and W. Sofka, 'Innovation on demand. Can public procurement drive market success of innovations?' *Research Policy*, 8 (2009).

In the last analysis, the European Innovation Partnership can be considered as the precursor of the partnership envisaged by the 2014 European Directive, although it is an instrument for implementing general policies through the funding of projects on broad and strategic sectors, where research and innovation are key to economic and social development, whereas the more recent Innovation Partnership for procurement has the same ambition of fostering research and innovation, but on a smaller scale, as it is applied on a contract-by-contract basis, through an award procedure that fosters innovation not through policies with a general impact, but through bespoke interventions.[4]

II. EU Aerospace Policies and the GovSatCom Programme

Historically, as mentioned above, economic aspects have always played a crucial role in the process of European integration, fostering the experimentation of innovative ways of solving problems common to the member states, and in this sense the establishment of the common market and the introduction of the single currency have been the most important illustration of this journey.

With the adoption of the TFEU, the debate at European level focused on European economic governance, in order to identify suitable ways to ensure a revival of the competitiveness of the European economy at international level, even though an attempt had already been made in March 2000 with the definition of the Lisbon Strategy to outline the most suitable policy to make the European Union the most competitive economy in the world, although it was soon realised that the objectives set were particularly difficult to achieve, and this encouraged the study of new and different solutions to revitalise the European economic system.

In the current globalised economic environment, measures to guide economic development need to be coordinated at EU level, avoiding the adoption of different and disharmonious policies by individual Member States.

Indeed, the aerospace sector is one of the privileged sectors in which public institutions have historically played an active role in supporting and directing the development of the sector, probably due to its link with national industry and national defence policy.

However, even the operations carried out in the first half of the 20th century by each country individually have important common features, and this has led to the emergence of single European strategies for carrying out ambitious and innovative projects since the 1960s.

Nevertheless, it is not easy to identify suitable instruments for international cooperation in the aerospace sector without a real European government,

[4] F. Gambardella, *Le regole del dialogo e la nuova disciplina dell'evidenza pubblica* (Torino: Giappichelli, 2016), 145-146.

although projects such as Concorde and Airbus, in civil aeronautics, or Galileo, in satellite positioning, have recently been implemented.[5]

The Airbus project, the result of a Franco-German initiative in 1967, was intended to support the development of the civil aeronautics sector in Europe in order to drive European economic modernisation in other sectors as well.[6]

The aeronautics industry had already become a sector of public interest at the beginning of the 20th century. This was due, firstly, to its importance for military purposes and, secondly, to its characteristic of being a 'research and innovation' intensive industry, which enabled its development for civil purposes.[7]

In the past, the civil aircraft market was mainly controlled by US companies such as Boeing, which meant that Airbus was at a particular disadvantage from the outset. It must be considered that, while Boeing was addressing a particularly large and imposing market such as the United States, Airbus could not rely on a single European market, which at that time was fragmented into several national markets, which were also regulated differently in political terms. The division into numerous small national markets in Europe did not allow the creation of a single internal market capable of absorbing production on economically advantageous terms, to the benefit of the US industry.[8]

Indeed, at the end of the 1960s, European aircraft manufacturers were on the fringes of the international market and essentially played the role of subcontractors for US industry.[9]

Moreover, Boeing had a certain synergy between civil and military production, whereas Airbus concentrated only on civil aeronautics, bearing in mind also that Boeing received federal public aid for research, whereas Airbus could not count on a European research policy but only on national aid.

And still, in spite of the many obstacles, both political and economic, Airbus succeeded in taking Boeing's world market leadership in 2004, the result of a series of courageous long-term choices, starting with the choice of legal status for the initiative, that of a *Groupement d'intérêt économique* (GIE), without

[5] P. Miller, 'Aerospace Companies and the State in Europe', in J. Hayward ed, *Industrial Enterprise and European Integration. From National to International Champions in Western Europe* (Oxford: Oxford University Press, 1995); K. Hayward, *European collaboration in civil aerospace* (London: Pinter, 1986).

[6] P. Drucker, Economia, politica e management: nuove tendenze nello sviluppo economico, imprenditoriale e sociale (Milano: Liguori, 1989); See also R. Koselleck, Futuro passato: per una semantica dei tempi storici (Genova: CLUEB, 1986); F. Mosconi, 'La politica industriale europea e la competitività italiana nei settori high-tech', in P. Guerrieri et al eds, Tornare a crescere (Roma: Arel, 2005).

[7] D. Velo, 'L'impresa europea di interesse generale', in G. Rossi ed, *L'Impresa europea di interesse generale* (Milano: Giuffrè, 2006).

[8] K. Hayward, Industrial Enterprise and European Integration, from National to International Champions in Western Europe (Oxford: Oxford University Press, 1995).

[9] A. Bonaccorsi, *Cambiamento tecnologico e competizione nell'industria aeronautica civile* (Milano: Guerini e associati, 1996); G. Raffaello, 'La grande impresa federale europea: il caso Airbus', in D. Velo ed, *L'Europa dei progetti* (Milano: Giuffrè, 2007).

share capital and with the economic responsibility of the founding members. This structure retained an important public character of a confederal nature, suited more to protecting the interests of the participants than to supporting the entrepreneurial project; this meant that the enterprise represented only an instrument subordinate to the interests of the participants, which presented itself in a unitary manner solely for marketing purposes.[10]

In a very concise way, it can be said that this operation represents the success of a political vision aimed at collaboration to support long-term political choices. Airbus certainly represents an important precedent, considering, moreover, that its success took place before the creation of the European Union.

The Galileo project, on the other hand, started in 1999 on the joint initiative of the European Union, the Commission and the European Space Agency (ESA) and represents the first example of collaboration between the Community institutions and the ESA, with the aim of building an autonomous satellite radio navigation system.[11]

Indeed, one could consider the Galileo project as the main reference point for the development of a European Union space policy, which was inaugurated at the time for three different reasons.[12]

Firstly, the European Union and the ESA wanted to develop information technology on the basis of the American experience; indeed, the greatest innovations in this field are obtained as a result of research in the space sector.

US companies operating in technology-intensive sectors still benefit from an advantageous position due to the US leadership in space exploration programmes. For this reason, it is worth mentioning the example of the first model of electronic calculator built in 1946 for the US Army and characterised by its considerable size and weight of thirty tonnes. Later, the need to use this instrument in the space sector meant that its weight and size had to be reduced, which favoured the development of microelectronics and more commercial applications such as personal computers.

Taking into consideration the US experience, the intention was therefore to provide Europe with the most advanced technological infrastructure in the satellite

[10] D. Hickie, 'Airbus Industrie: A Case Study in European High-Technology Cooperation', in U. Hilpert ed, *State Policies and Techno-Industrial Innovation* (London: Routledge, 1991); P. Drucker, *Innovation and Entrepreneurship* (New York: Harper&Row, 1985); S. Mcguire, *Airbus Industrie. Conflict and Cooperation in US-EC Trade Relations* (Oxford: St. Antony Series, 1997).

[11] F. Von der Dunk, 'Towards one Captain on the European Spaceship – Why the EU should Join ESA' *Space Policy*, XIX, 83-86 (2003); W. von Kries, 'Which Future for European Space Agencies?' *Space Policy*, XIX, 157-161 (2003).

[12] L. Bottinelli, 'L'impresa europea di interesse generale: il progetto Galileo', in G. Rossi ed, *L'Impresa europea di interesse generale* (Milano: Giuffrè, 2006), 148; G. Venturini, 'La sfida di Galileo, paper for the Conference 'Il futuro dell'Europa nelle tecnologie' ' Venice, 7 May 2004; European Space Agency, *Galileo – The European Programme for Global Navigation Services*, available at www.esa.int.

navigation sector, with positive spin-offs in other areas as well.[13]

The second reason is instrumental, since the policies pursued at the end of the 1990s by the European Union concern the sectors in which the implementation of the Galileo project could bring the greatest benefits, such as agriculture, the environment, transport and scientific research.

Finally, political considerations must be underlined, since the programme is intended as an alternative to the civil satellite radio navigation services provided by the American Global Position System (GPS). The creation of an autonomous satellite system underlines Europe's desire to become independent of US services. The creation of Galileo will therefore enable Europe to equip itself with an autonomous system consisting of thirty satellites in orbit and the ground facilities that will receive their signals.[14] Among the most recent programmes, the Governmental Satellite Communications ought to be highlighted, referred to as the GovSatCom programme. This programme was launched in 2013 by the European Council to prepare for the next generation of governmental satellite communications in 2025, through close cooperation between Member States, the European Commission, the European Space Agency (ESA) and the support of the European Defence Agency (EDA).

In particular, the goal of the GovSatCom programme is to ensure reliable, secure and cost-effective civil and military satellite communication services for EU and Member State public authorities managing critical security missions and operations.

In pursuing this goal, the GovSatCom programme also aims to strengthen European autonomy and overcome fragmentation of demand through the use of accessible and innovative solutions in synergy with industrial actors.

Indeed, it is very timely to note that satellite communications have become critical and essential elements for defence, security, humanitarian aid, emergency responses or diplomatic communications, given their crucial role in military missions and civil operations, especially those taking place in remote locations with little, if any, infrastructure available.

Decision-making processes cannot now ignore highly sensitive and timely information, which is why the importance of having secure connection and communication systems available cannot be underestimated.

Satellite communications in the GovSatCom category, in particular, provide secure access with high standards of protection, without, however, matching the levels of MILSATCOM communications, which are generally provided by sovereign military systems. The COMSATCOM category, which includes satellite communications purchased on the commercial market on the basis of need and availability, is even different.

However, the EDA's priorities for satellite communications are directed in

[13] W. Hutton, *Europa VS USA* (Roma: Fazi, 2003).
[14] L. Bottinelli, n 12 above, 149.

particular at the GOVSATCOM (as well as EUSATCOM Market) initiatives, which have developed following a sequential procedure.

From the outset, EDA priorities for satellite communications were identified by the Member States' Ministries of Defence as early as 2011, while the GovSatCom Programme was endorsed by the EDA Governing Council on 19 November 2013, thus providing for the establishment of a dedicated Government Satellite Communications Group (GOVSATCOM) composed of Germany, Spain, France, the United Kingdom and Italy in order to assess how their respective current and planned national capabilities could address future needs.[15]

The European Council's approval, on the other hand, dates back to 19 December 2013, and emphasises once again that the next generation of government satellite communications must be achieved through close cooperation between the member states, the European Commission and the European Space Agency.

Finally, in November 2014, the EDA Steering Committee approved the initial satellite communication requirements for European military actors involved in the conduct of national operations and Common Security and Defence Policy (CSDP), included in the Common Staff Target document. Subsequently, with the support of a feasibility study, the EDA developed the technical and mission requirements and evaluated various solutions to meet these needs.

The 'Common Staff Requirements' document and associated Business Case, which details the GOVSATCOM requirements and proposes the way forward to meet these requirements, were finally approved by the EDA Steering Committee in March 2017.

Following this approval, the EDA, together with contributing Member States, developed the EDA GOVSATCOM Pooling and Sharing demonstration project (GSC Demo) from June 2017. In January 2019, the GSC Demo entered the execution phase with the first meeting of the project organization management group.

This means that the project is now ready to provide GOVSATCOM services to meet the specific requirements of Member States and European CSDP actors through the pooled capabilities of contributing Member States.

This governmental pooled capability was created to provide satellite communication resources that cannot be obtained on the commercial market with a sufficient level of guaranteed access and security. The GSC demo responds to an existing need and is fully in line with the revised 2018 Capability Development Plan and the resulting EU capability development priorities.

The project should also be seen in the light of ongoing EU efforts to establish

[15] Conclusions of the European Council of 19-20 December 2013 on the Common Security and Defence Policy Communications of the Minister of Defence – 23 January 2014, Camera dei Deputati, Ufficio Rapporti con l'Unione europea, XVII legislatura, 15 /2014 – 22 January 2014, available at https://tinyurl.com/s9kv7r46 (last visited 30 June 2021).

a European GOVSATCOM in the framework of the next space programme and in the context of the European Space Agency's activities in the field of satellite communications security.

Furthermore, the GSC Demo project complements the EDA's EU Satcom Market project, which has been running since 2012, which provides commercially available SATCOM and CIS services in an efficient and effective way.

The GOVSATCOM initiative marks a new partnership, not only between military and civilian institutional actors, but also with industry, in order to better contribute to Europe's competitiveness.

Indeed, according to the European Council, satellite communications (SatCom) are critical elements for defence, security, humanitarian aid, emergency response or diplomatic communications. They are a key element for civilian missions and military operations, in particular in remote and austere environments with little or no infrastructure, which is why they have been defined as one of the four capability development programmes, together with Air-to-Air Refuelling, Remote Piloting of Air Transport Systems and Cyber Defence.

Considering these premises, it is easy to see that the sector has plenty of room for research and experimentation, where the innovation partnership has finally found fertile ground for its application.

At the national level, in fact, the contribution to the European GovSatCom initiative is represented by the Mirror GovSatCom Programme, designed to give Italy an important position in a strategic sector such as institutional telecommunications. In particular, the national objective is to build and activate Ital-GovSatCom, an innovative and competitive satellite system for the provision of telecommunications services with security, resilience and reliability features that allow it to be used for institutional purposes in various fields of application, such as civil protection, security, defence, humanitarian aid, telemedicine and maritime surveillance.

III. The Innovation Partnership Applied Within the Italian Space Economy Plan: The Mirror GovSatCom Programme

The use of the innovation partnership procedure requires a major effort both for the economic operators, who would have to engage in an operation with an uncertain outcome, but also for the public administration, which, if all the prerequisites required by the codified regulation were to be met, would have to commit itself to following a complex procedure requiring a high level of negotiation skills.

Indeed, according to the provisions of Art 65 of decreto legislativo 18 April 2016 no 50, contracting authorities are allowed to use the Innovation Partnership when they have to 'develop innovative products, services or works' and 'subsequently purchase the resulting supplies, services or works', if this need

cannot 'on the basis of a reasoned determination, be satisfied by recourse to solutions already available on the market'.

However, since this is an alternative procedure for selecting the contractor or contractors to the other main procedures provided for by the code, a reasoned determination by the contracting authority is required, which will involve in-depth knowledge of what the market already offers in order to rule out the existence of solutions that are likely to satisfy the public interest.[16]

From a purely procedural point of view, the economic operators invited by the contracting authority have to send their request to participate within thirty days from the publication of the contract notice; more precisely, it is a procedure in stages, according to the development of the research and innovation process, at the end of which the contracting authority assesses the objectives achieved, those still to be achieved and the consequent remuneration, and may also decide to terminate the innovation partnership or, if there are several operators, to reduce the number of them by terminating the individual contracts, provided that it has provided for this possibility in the tender documents.

Nevertheless, the possibility that the contract notice allows the contracting authority to unilaterally terminate the partnership has a dissuasive effect on economic operators who, after having invested in the project and research activity, could see their efforts thwarted by the administration's decision.[17]

However, the innovation partnership finds its principal expression in the phase of confrontation between the administration and the bidders called upon to propose innovative solutions, with a view to improving their content by means of negotiation phases for the initial bids and all subsequent bids, except for the final bids, which will be assessed for award on the basis of the 'best value for money'. This is a criterion that allows bids to be judged without rigid evaluation schemes, suitable for a procedure that enhances the negotiating autonomy of the parties.

What is required from the administration is to assume a highly specialised role, capable of understanding and identifying the most innovative solutions for its own needs.

Indeed, the administration should be able to position itself as an innovator. Moreover, it must be capable of understanding the alternatives present on the

[16] The contracting authority's justification appears to be a kind of *probatio diabolica*, because it should be not easy to demonstrate the non-existence of useful solutions to satisfy the public interest as regards a product, service or work. It is not possible to ignore the difference between the discipline provided for by the Italian Code and that contained in Art 31 of Directive 2014/24/EU, which does not provide for a specific reason in support of the choice, but only that 'In the tender documents the contracting authority shall identify the need for innovative products, services or works which cannot' be satisfied by purchasing products, services or works available on the market'. L. Marraccini and G. Terracciano, 'Partenariato per l'innovazione', in M.A. Sandulli and R. De Nictolis eds, *Trattato sui contratti pubblici* (Milano: Giuffrè, 2020), III, 133.

[17] F. Gambardella, n 4 above,149; C. Lamberti and S. Villamena, 'Nuove direttive appalti' *Urbanistica e appalti*, 875 (2015).

market for the goods, services or supplies it intends to acquire, and then to relate to the private party in a partnership relationship, for which there is still the risk of so-called information asymmetry.[18]

These and other issues are among the main themes analysed by the administrations that have participated and are still committed to the implementation of the national plan for the Space Economy,[19] which has identified the Partnership for Innovation as the most suitable tool for achieving innovative results through a complex system of funding.

More precisely, the national Space Economy plan is part of the strategy that, starting in 2014, has been promoted by the Italian Presidency of the Council of Ministers for the definition of the national policy in the space sector, since, in the view of the *Cabina di Regia* specifically created for its drafting, this plan will be able to enable Italy to transform the national space sector into one of the driving forces of the country's new growth.

It has been said that the Space Economy encompasses the production process that begins in the Upstream sector through research, development and construction of space infrastructures, which enables the implementation in the Downstream sector of related innovative products and services, such as services for environmental monitoring and weather forecasting.

It is clear that the growth of the Downstream sector is based on the quantity and variety of value-added services developed and managed by companies, especially SMEs with medium-high qualification staff.

From an operational perspective, the objectives set by the National Space Economy Plan require two parallel and complementary operations, one aimed at identifying the needs of economic operators, including those outside the sector, with regard to the development of new value-added services based on satellite data; the other aimed at organising the traditional channels of intervention of the national space policy with the resources and strengths of the regions interested in the spill-over effects of the Space Economy on their territories, operating mainly through the funding of space initiatives considered jointly suitable for this purpose.

These operations are based on joint co-financing using national funds, from the Development and Cohesion Fund, and regional funds, from the Regional Operational Programmes of the 2014-2020 programme.

The Plan, in fact, is based on a multi-regional cooperation programme aimed at promoting the supply of innovative technologies, services and products by enterprises and research competences from the respective territories, while the national intervention acts mainly on the innovative demand side by playing the

[18] About information asymmetry, see also L. Marraccini, 'The comparison of two project financing operations: the Line 5 of the Milan Metro and the London Tube' *Amministrativamente.com*, 11-12 (2017).

[19] Available at www.mise.gov.it.

role of Buyers Group, through the use of the Innovation Partnership, pursuant to Art 65 of *decreto legislativo* 18 April 2016 no 50.

The implementation of the National Space Economy Plan is carried out through specific programmes, among which the Mirror GovSatCom Programme stands out, which, as mentioned above, through the Innovation Partnership intends to build and operate the Ital-GovSatCom satellite system for the provision of innovative institutional telecommunications services, or alternatively, to build innovative elements of this system.

As envisaged by the Space Economy Plan, for the implementation of the Mirror GovSatCom Programme it was also necessary to resort to the Multi-Regional Plan of aid for research and development, functional to the implementation objectives of the Innovation Partnership signed by the Minister for Economic Development and the Regions and Autonomous Provinces interested in supporting the implementation of the Ital-GovSatCom satellite system.[20]

The contracting authority is the Italian Space Agency (ASI), which published the call for tenders in June 2018,[21] describing the four-stage procedure.

The first phase was the pre-qualification phase, open to all interested parties who, from the date of publication of the call for tenders, had thirty days to submit their applications to ASI.

Once the examination of the applications was complete, the contracting authority selected the operators to be invited to the second phase of initial bidding and negotiation, ie those who did not have any grounds for exclusion under Art 80 of the Italian Public Contracts Code and who met the general requirements under Art 83 and the specific requirements set out in the pre-qualification specifications.

Finally, the companies that passed the previous stages submitted the final offer, not subject to negotiation, through which the successful bidder was identified, resulting in the signing of the contract, which took place in July 2019.[22]

ASI's decision to make use of the Innovation Partnership is certainly due to evaluations of a legal nature, which have taken into account the issues related to the regulation of State aid and intellectual property rights, in relation to which the European directives and the Italian Public Contracts Code leave a wide discretion to the contracting authority.

Arts 107 and 108 of the Treaty on the Functioning of the European Union (TFEU), in fact, prohibit state subsidies that distort competition in the internal market and affect trade between member states in a manner contrary to the common interest, with consequent implications also on the regime applicable to intellectual property rights in innovation partnerships which, in light of the

[20] For further details about the Mirror GovSatCom Programme, see also the Detailed Operational Plan of MirrorGovSatCom available at www.mise.gov.it

[21] Available at www.asi.it

[22] 'Parte la Space Economy italiana. Assegnato il contratto per Ital-GovSatCom' *AirPress*, available at https://tinyurl.com/4rmb3vk2 (last visited 30 June 2021).

codified discipline, is defined by the tender documents according to the discretion of the contracting authority.

On this point, recital 49 of Directive 2014/24/EU provides that contracting authorities should not use PPI in such a way as to prevent, restrict or distort competition.

After all, if the contracting authority does not decide to hold the property rights exclusively, but to share them with one or more private partners, there would be a risk that the benefits obtained through the innovation partnership would be used to develop other innovative goods, services or works to be marketed in a market other than the one specifically created with the PPI, thus enjoying advantages that distort competition.

With the Communication 'Framework for State aid for Research and Development and Innovation' of 27 June 2014, the European Commission expressed an intention to harmonise procurement and State aid disciplines, reaffirming the importance of promoting research and technological development and the strategic nature of the Europe 2020 objectives, in line with Art 179 TFEU; it is therefore stipulated that for contracts for research and development services, the Commission will consider that no State aid has been granted to undertakings when the price paid for the services fully reflects the market value of the benefits obtained by the public purchaser and the risks assumed by the supplier.

In this sense, the Innovation Partnership could be assimilated to research and development service contracts, at least for the first phase of the procedure, where the research and development activities are concentrated.[23]

However, the 2014 Communication states that one of two conditions must alternatively be met: the first,

> all results which do not give rise to IPR may be widely disseminated, for example through publication, teaching or contribution to standardisation bodies in a way that allows other undertakings to reproduce them, and any IPR are fully allocated to the public purchaser;

the second, that

> any service provider to which results giving rise to IPR are allocated is required to grant the public purchaser unlimited access to those results free of charge, and to grant access to third parties, for example by way of non-exclusive licenses, under market conditions.

With regard to the Mirror GovSatCom PPI, the intellectual and material property rights deriving from the performance of the research activities of

[23] S. Bigazzi, 'Le "innovazioni" del partenariato per l'innovazione', in A. Fioritto ed, *Nuove forme e nuove discipline del Partenariato Pubblico Privato* (Torino: Giappichelli, 2016), 215-220.

common interest to the public and private partners are jointly vested in ASI and the successful economic operator, who will therefore become co-owners according to their share of the financial participation, with the right of free use for the institutional purposes of the public party and the right of the PPI Partnership to the commercial exploitation of the new solutions and technologies developed with recognition of royalties to the public party itself.[24]

The above considerations make it evident that the innovation partnership is not, in general, always a linear and simple operation, but usually presents complex structures distinguished by successive phases.

The first phase of the Mirror GovSatCom PPI is dedicated to the research and development of innovative satellite telecommunications solutions and applications, based on an innovative GEO platform, in response to emerging advanced institutional requirements; the second phase, on the other hand, involves the creation and deployment of an innovative satellite system, called Ital-GovSatCom, for the provision of telecommunications services, the provision and marketing of which are the subject of the third and fourth phases.

The success of the operation depends above all on the first phase, which will last no longer than 24 months, during which a platform for testing and validating the services will have to be set up to allow feedback from the Buyers Group.

In this sense, the end of the first phase will amount to a pivotal moment, in which the Buyers Group will assess the interest in the services that can be provided by the system, possibly expressing the willingness to use and/or purchase with its own financial resources.

Each of the above-mentioned phases corresponds to objectives subject to evaluation and negotiation, bearing in mind also that the PPI may be interrupted, at the end of each phase, if the contracting authority considers that the objectives cannot be achieved or further negotiated.

The first objective relates to research and development activities to identify innovative solutions in line with the requirements for satellite telecommunication services set out in the 'Satellite Communication to support EU Security Policies and Infrastructures – High Level SATCOM User Requirements', which identifies and describes the requirements of Institutional Users for satellite communications necessary to support EU policies in the field of security and infrastructure. Furthermore, given that these services are not exclusively of institutional interest, the economic operator will also have to develop innovative applications and solutions of interest to the private customer market.

The second objective is the implementation and operation of the Ital-GovSatCom system, based on the research and development activities developed under the first objective. As regards the types of service envisaged, reference must first of all be made to a 'basic' service of instant communication (voice and

[24] See also the Detailed Operational Plan of MirrorGovSatCom available at www.mise.gov.it

text) at reduced bandwidth over the entire access area, ie the national territory, the Mediterranean Sea and other geographical areas of interest. There must then be 'permanent' broadband services over the primary area of interest, eg for the management of infrastructures and the provision of communication to other space systems, while, finally, there must be 'advanced' broadband communication services on demand (such as real-time imaging and video, videoconferencing) over the entire access area.

Service provision is therefore the third objective, to be achieved once the satellite has been successfully launched, with its final GEO orbit and commissioning.

Finally, the fourth objective is commercialisation, with the purchase of services by the Buyers Group and business users.

After describing the objectives of the MirrorGovSatCom Programme, the Detailed Operational Plan of MirrorGovSatCom[25] considers a further scenario, whereby, in the event that the ItalGovSatCom system cannot be built, the research and development activities of the first objective will be useful for the creation of innovative components and subsystems for the construction of civil and dual telecommunications satellites.

Such a complex system of objectives is necessary because the architecture of the ItalGovSatCom system itself is complex, since it is composed of a Flight Segment, consisting of the GEO satellite based on the innovative telecommunications platform, a Ground Segment, ie the control centres for the satellite and the telecommunications mission, as well as an Application Platform, ie a Service Hub and a platform for the development of applications; finally, there is the User Segment, consisting of the user terminals for the various types of Users.

In this regard, it should be stressed that the objectives can be negotiated by the parties, unlike the minimum requirements for the innovativeness of the supply or services developed in the PPI, which, together with the award criteria, cannot be negotiated. Thus, the partnership must be terminated if the contracting authority considers, at the end of each phase, that the minimum requirements have not been met.

IV. Conclusion

In order to modernise the administration, the legislator has recently been encouraging the spread of models that allow, and sometimes require, the participation of the private sector, in accordance with an approach that is gradually becoming established in a number of forms in the relationship between the public administration and the private sector, intended as a citizen,

[25] Available at www.mise.gov.it

user or economic operator.

In the field of public contracts, public-private cooperation has encountered strong resistance in the case of certain activities, such as the construction and management of public works or the provision of services, in which the public partner can benefit from financial resources and private know-how which, especially in the technological field, appears difficult for the public sector to acquire.

In this sense, the reference is obviously to forms of public-private partnership, in which the administration moves away from a role of direct intervention in the market, and instead takes on a role of organisation, regulation and control.

However, the partnership with the private party may also take place in forms that are different from the PPP and that fall rather within the procedures for the selection of the contractor, among which the innovation partnership has been introduced.

Indeed, the fact that in the innovation partnership there is a form of partnership with a private party should not lead to any confusion between the procedure established *ex novo* by Art 31 of Directive 2014/25/EU and the public-private partnership.[26]

The innovation partnership is not part of the PPP scheme, whether purely contractual or institutionalised, since it is not intended to award the contract for the implementation or management of a public work or service, but merely to encourage the contracting authorities to work with the private sector to develop innovative products, services or works and subsequently to purchase what has been produced, provided that the product or service can be supplied or the works can be carried out to agreed performance levels and costs.

It is therefore an alternative way of choosing the contractor, albeit characterised by a more intense form of cooperation than the open and restricted procedures.

On the other hand, it would seem that the element of connection between the PPP and the innovation partnership is the fact that in both procedures the identification of the economic-financial balance is of crucial importance, since in both cases there is a real risk that the asymmetry of information between the public partner and the private partner will lead to the conclusion of opaque contracts, with a strong imbalance in the allocation of risks and responsibilities.

The special feature of the innovation partnership is that, compared to other procedures for the selection of contractors and, in particular, to the competitive dialogue, this instrument allows the best possible exploitation of the potential of the private sector, since the need of the contracting authority cannot be satisfied by solutions already available on the market.

As mentioned above, administrations need to develop innovative products, services or works and, therefore, initiate a long-term procedure aimed at

[26] L. Marraccini and G. Terracciano, n 16 above, 141.

achieving solutions, divided into successive phases with verification of intermediate results, and characterised by strong public-private cooperation for the negotiation of proposals and tenders.

In this way, a contracting procedure becomes a tool for smart, sustainable, inclusive growth, in line with the European Commission's guidelines contained in Communication COM (2010) 2020 final, 'Europe 2020 – A strategy for smart, sustainable and inclusive growth'.

With particular reference to the use of the instrument of the innovation partnership within the Italian Space Economy, it can be seen that it was a successful choice.

The procedure governed by Art 65 of the Italian Public Contracts Code, in fact, seems the most suitable for this kind of operations, where the public subject has to satisfy a need but there are no solutions available on the market and, consequently, a solution has to be developed in collaboration with the private sector.

In particular, the use of the innovation partnership seems to be more appropriate than competitive dialogue, which is more adequate in cases where there are significant information asymmetries between public and private partners.

In this case, however, it would not appear that there has been a disproportionate information asymmetry between the administration and the private sector, since the Italian Space Agency plays the role of contracting authority. Moreover, it ought to be taken into account that the Italian Space Agency is the national public body called upon, among other things, to promote national excellence in the research and development sector in the space field and a high level of competitiveness of the Italian industrial sector, with particular reference to small and medium-sized enterprises (SMEs), in order to be able to make the most of their competitiveness and capacity for innovation.

It was also essential that the Buyers Group carry out a proper needs assessment phase for the purposes of the technical and economic sizing of the project, with particular reference to the expected requirements for satellite communication services.

Moreover, the involvement of institutional users takes place constantly in all phases of the Innovation Partnership, both during the specification phase of the minimum innovative requirements of the supply, through the consultation process, and during the research and development phases, with the detailed indication of the user requirements by the Buyers group. This also involves additional institutional users potentially interested in the services; at the end of the research and development phases of the innovation partnership, on the other hand, the service offer proposed by the economic operator will be evaluated and possibly committing to acquire the services made available downstream of the implementation phase, while during the Development and Implementation phase, the involvement of users will take place with the operational verification

of the solutions implemented and the use of the services.

On the basis of the considerations above, it seems, therefore, that the innovation partnership represents a testing ground for an important part of the Italian administration which, with the PPI GovSatCom, is beginning to come closer to the vision that European law has of it, ie a player that knows the market for the goods, services or works that it intends to acquire and, in this sense, acts as an innovator, able to imagine and design products not yet available on the market.

At the end of the innovation partnership procedure, in fact, a close collaboration is established with the private partner, with different characteristics compared to those of the contractual relationship that binds the administration to the contractor, also thanks to the progressive negotiation of the conditions of the final offer.

In perspective, contracting authorities should be able to make technically complex choices, such as the prior identification of conditions to terminate a contract in the medium term, or the relative termination of a partnership, ie limited to individual partners, decisions for which adequate technical expertise is required.

From the Emissions Trading System to the Role of Private Law in Environmental Protection. Notes for Research

Andrea Nervi*

Abstract

The essay moves from a description of the emission trading system, as regulated by international agreements and European directives, focusing on the measures contemplated therein.

Starting from these premises, two aspects come to attention: first, the interaction between private and public law instruments in the construction and functioning of the market mechanisms; second, the effective suitability of such mechanisms to pursue the environmental purposes which, ultimately, represent (or should represent) the end purpose of the regulatory provisions.

There is, however, a problem of value consistency between the environmental purposes underlying the regulations of this market. Actually, the importance of the environmental issue seems to require more incisive forms of regulation than those that can be ensured by market dynamics.

I. Air Pollution and the Strategy to Contain Greenhouse Gas Emissions; Summary of the Regulatory Framework

In this contemporary era, the question of the contamination of the ecosystem and its consequences has reared its head in a sudden and even dramatic way. After decades, if not centuries, during which certain activities – industrial activities, transport services, etc – have been performed without particular concern for the environment, society today finds itself having to face new, unexpected events and phenomena, such as rising temperatures, the alteration of the normal hydrological cycle and sudden extreme weather phenomena, which sometimes place under great stress the structure of the terrain, as well as local infrastructures that man has constructed.

It is an obvious, but not irrelevant, observation that the extent of the above-mentioned events is indeed global, in the sense that they transcend national borders. Therefore they require a reaction and, more generally, the adoption of strategies which – at least in the first instance – must be on the level of international public law, which therefore develop into international agreements

* Full Professor of Private Law, University of Sassari. Published with the contribution of FAR2020NERVIA – Fondo di ateneo per la ricerca di base (COEP).

and – when possible, as in the case of the European Union – into regulatory acts of supranational scope. With reference to the specialised literature on the subject, one must perforce mention the Kyoto Protocol, dating back to 1997 and subsequently amended by the Paris Agreement of 2015,[1] as well as certain European directives and regulations;[2] the former (directives) have later been implemented in Italy.[3]

Said initiatives aim to limit the release into the atmosphere of so-called greenhouse gases, ie polluting emissions which – due to the 'greenhouse' effect – lead to the overheating of the earth's surface and oceans and, consequently, to alterations in the ecosystem. The pursuit of the aforementioned aim requires intervention *a posteriori* in anthropogenic activities that have now become consolidated practices, which often generate wealth and well-being for the subjects who control them or who, in any case, are involved in the same in various capacities (entrepreneurs, shareholders, workers, financiers, etc).

In an attempt to summarise extremely briefly a scenario that is sometimes very complex from a technical viewpoint, for the aforesaid purposes a programme of regulatory measures has been drawn up featuring a mixture of elements of private and public law,[4] with the intention of thus maintaining the reference to the traditional categories recognised by classical legal experts. This approach stems from the belief that purely authoritative tools will be insufficient and/or useless, and that it is therefore necessary to also adopt reward mechanisms, or incentives, based on market logics and, therefore, of a contractual nature.[5]

[1] See M. Montini, 'Riflessioni critiche sull'Accordo di Parigi sui cambiamenti climatici' *Rivista di diritto internazionale*, 719 (2017). See also A. Chiappetta and A. Gaglioti, 'Sviluppo sostenibile ed Emission Trading Scheme. Ipotesi ricostruttive per il mercato europeo della CO2' *Amministrazione in cammino*, 10 (2011).

In literature, the problem known as carbon leakage has quite rightly been pointed out; it refers to the re-location of companies from a regulated area to an unregulated part of the world, which leads to the risk of distorting the system itself. On this subject, see B. Pozzo, *Il nuovo sistema di Emission Trading comunitario* (Milano: Giuffrè, 2010), 83; G. Lo Schiavo, 'Emission trading e tutela dell'ambiente: quali obblighi per le imprese in vista dell'entrata in vigore della terza fase?' in G. Alpa et al eds, *Rischio di impresa e tutela dell'ambiente* (Napoli: Edizioni Scientifiche Italiane, 2012), 267; K. Peterková Mitkidis, 'Using Private Contracts for Climate Change Mitigation' 2 (1) *Groningen Journal of International Law*, 55 (2014).

[2] European Parliament and Council Directive 2003/87/EC of 13 October 2003 establishing a scheme for greenhouse gas emission allowance trading within the Community and amending Council Directive 96/61/EC [2003] OJ L275/32; and European Parliament and Council Directive 2009/29/EC of 23 April 2009 amending Directive 2003/87/EC so as to improve and extend the greenhouse gas emission allowance trading scheme of the Community [2009] OJ L140/63.

[3] See decreto legislativo 13 March 2013 no 30; decreto legislativo 4 April 2006 no 16. On the implementation of the directives in question under Italian law, see F. Giglioni, 'The Allocation of CO2 Emission Permits in Italy', in P. Adriaanse et al eds, *Scarcity and the State II* (Cambridge: Intersentia, 2016), 123; see also F. Gaspari, 'Tutela dell'ambiente, regolazione e controlli pubblici: recenti sviluppi in materia di EU Emission Trading Scheme (ETS)' *Rivista italiana di diritto pubblico comunitario*, 1155 (2011).

[4] On this subject, V. Jacometti, *Lo scambio di quote di emissione* (Milano: Giuffrè, 2010), 26.

[5] See C. Camardi, 'Cose, beni e nuovi beni, tra diritto europeo e diritto interno' *Europa e diritto*

Again extremely briefly, the public authority, specifically identified in advance by the legislator, establishes the maximum limit of emissions that can be allowed over a given period of time (the calendar year) in the area under its competence (generally determined on a national basis). Said limit is then distributed among the operators which, in the area in question, carry out the activities that generate the polluting emissions. As a result, each operator is assigned a certain annual emission allowance, which is to say a certain quantity of polluting emissions that may be produced by the activity of each single operator. If the established limit is exceeded, the operator in question will be fined by the competent authority.

However, a single operator does not necessarily produce the maximum amount of its allowed emissions. In fact, the emissions attributable to an operator's activity may remain below the established threshold if, for example, the operator has made investments to reduce emissions. In such a case, the operator in question does not take avail of its full 'allowance' of emissions produced by the relative polluting activities, and the surplus can be made available to other operators, which – *viceversa* – produce higher levels of pollution and find themselves unable to comply with the limits assigned to them by the competent authority. Thus, a demand for additional allowances is created by such operators, which can be met by the offer of allowances by a less-polluting operator.

In this way, a market for emission allowances arises,[6] in which, on one hand, there is a demand from the 'polluters' (ie, those who fail to respect the threshold set by the competent authority) and, on the other hand, an offer by the more 'virtuous' operators (ie, those who keep their emissions below their thresholds, due to their pollution-reducing efforts).[7] The mechanism is therefore designed to encourage operators to reduce greenhouse gas emissions, since they can thus monetise the reduction obtained, making it available to other operators.[8]

The description of the theoretical model can then be completed with the case in which, in addition to or as an alternative to the prior assignment of pollution rights, the competent authority organises an auction system, in which it is precisely the emission allowances that are for sale. In this scenario, the individual operator is

privato, 974 (2018); as well as M. Clarich, 'La tutela dell'ambiente attraverso il mercato' *Diritto pubblico*, 220 (2007). On this subject, see also M. Cecchetti and F. Grassi, 'Le quote di emissione', in R. Ferrara and M.A. Sandulli eds, *Trattato di diritto dell'ambiente* (Milano: Giuffrè, 2014), II, 304, which recalls the preceding preference for a system of environmental taxation and the successive evolution towards the current framework, which should balance the counter-posed needs of protecting the environment and of achieving sustainable development. Also mentioned by G. Mastrodonato, 'Gli strumenti privatistici nella tutela amministrativa dell'ambiente' *Rivista giuridica dell'ambiente*, 710 (2010).

[6] On this subject, F. Annunziata, 'L'atmosfera come bene negoziabile. I contratti di cessione di quote di emissione tra tutela dell'ambiente e disciplina del mercato finanziario', in M. Lamandini and C. Motti eds, *Scambi su merci e derivati su commodities* (Milano: Giuffrè, 2006), 778.

[7] See B. Pozzo, n 1 above, 5.

[8] On this subject, with clarity, see C. Camardi, n 5 above, 973.

faced with a choice: either to invest in pollution reducing technologies, or to allocate financial resources in order to purchase additional allowances and thus avoid the penalties otherwise applied for exceeding the limit.[9]

At this point, the system described above hinges entirely on the emission allowance, the legal nature of which has been the subject of discussion from the viewpoint of both private and public law. Recalling the literature already published on the subject,[10] it would seem that a solution to the question has been found by an arrangement in accordance with Commission Regulation (EU) no 1031/2010 of 12 November 2010 on the timing, administration and other aspects of auctioning of greenhouse gas emission allowances pursuant to Directive 2003/87/EC of the European Parliament and of the Council establishing a scheme for greenhouse gas emission allowances trading within the Community [2010] OJ L 302/1, which leads to the probability of emission allowances, once created, being assimilated to financial instruments. This, in turn, leads to their inclusion within the scope of application of the Markets in Financial Instruments Directive[11] and Markets in Financial Instruments Directive II[12], with all the consequences of the case in terms of the obligations of information, transparency and professional correctness on the part of the intermediaries involved in the individual transactions.[13]

In this way, a process has been completed which, from the outset, had recognised the need to develop a secondary market for the allowances.[14] In other words, a clear regulatory plan can be recognised, aimed at ensuring that emission allowances are considered 'liquid' assets, which is to say assets that can easily be traded on the market. It should be noted, not only between operators

[9] On this subject, V. Jacometti, n 4 above, 9, as well as M. Cecchetti and F. Grassi, n 5 above, 305.

[10] The debate is summed up by F. Giglioni, n 3 above, 126-127, where ample references are given; see also F. Gaspari, n 3 above, 1164.

[11] European Parliament and Council Directive 2004/39/EC of 21 April 2004 on markets in financial instruments amending Council Directives 85/611/EEC and 93/6/EEC and Directive 2000/12/EC of the European Parliament and of the Council and repealing Council Directive 93/22/EEC [2004] OJ L145.

[12] European Parliament and Council Directive 2014/65/EU of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU [2014] OJ L173/349.

[13] A clear illustration is given by F. Mocci and J. Facchini, 'La nuova disciplina delle quote di emissioni tra MiFID II e MAR' *Rivista di diritto bancario*, 11-14 (2016).

The evolution of the emission allowances market in the financial sense has been outlined perspicaciously by F. Annunziata, n 6 above, 778.

On this subject, see also V. Jacometti, n 4 above, 121.

[14] See M. Cecchetti and F. Grassi, n 5 above, 308, who observe that the negotiable allowances can, alternatively, be held as reserves (banking), or for borrowing (as assets to be borrowed/loaned) between different periods of regulatory obligation compliance. On this subject, see also G. Lo Schiavo, n 1 above, 270; as well as F. Mocci and J. Facchini, n 14 above, 2; V. Jacometti, 'La direttiva Emissions Trading e la sua attuazione in Italia: alcune osservazioni critiche al termine della prima fase' *Rivista giuridica dell'ambiente*, 280 (2008).

directly involved in the aim of reducing pollution and the greenhouse effect, but also among third parties,[15] such as financial intermediaries and, above all, their customers, ie private and public investors which – in this way – bet on the trend in the reduction of emissions and, therefore, on the impact of said emissions on the ecosystem.[16]

Before closing this technical description, it may be worth recalling that a system similar to that illustrated above also exists in other sectors,[17] such as that of renewable energy in which incentives exist for the use of energy that is 'clean' from an environmental viewpoint.[18] In this case, the relative legislation pursues the aim of inducing operators (ie, producers and importers of electricity) to ensure that a certain amount of energy produced or imported comes from renewable sources.[19] Once the target has been set, the operators in question can either produce renewable energy themselves or buy the equivalent quota from companies that carry out this activity. In this case, therefore, the quantity of renewable energy used is certified by the so-called green certificates, which therefore become a 'commodity' to be traded between electricity producers and industrial operators burdened by the obligation of introducing renewable energy into the system. The operators are thus called upon to choose whether to produce this energy due to their own structural investments, or to buy it from subjects which already produce it, and which will therefore be encouraged to produce more.

Also in this sphere, therefore, suppliers and customers of renewable energy quotas become counterparties, and the consequent trading of the quotas is fostered, in order to achieve the goal pursued by law which is, namely, an increasing use of renewable energy sources featuring modest environmental impact.

[15] On this subject, F. Giglioni, n 3 above, 128.

[16] On the secondary market which is thus generated, and on the relative functioning, see F. Mocci and J. Facchini, n 14 above, 7. Considerations on this point are expressed by B. Pozzo, 'Le nuove regole dello sviluppo: dal diritto pubblico al diritto privato', in *Benessere e regole dei rapporti civili. Lo sviluppo oltre la crisi* (Napoli: Edizioni Scientifiche Italiane, 2015), 97. The author specifies that the emission allowance instrument functions efficiently when the number of companies operating on the market is high; the presence of a limited number of operators could lead to a paralysis of the circulation of pollution rights, since companies could decide not to exchange rights, thus preventing new companies from entering the market.

[17] At this point, it is necessary to mention the extension of the scheme to air transport (European Parliament and Council Directive 2008/101/EC of 19 November 2008 amending Directive 2003/87/EC so as to include aviation activities in the scheme for greenhouse gas emission allowance trading within the Community [2009] OJ L8/3, implemented in Italy by decreto legislativo 30 December 2010 no 257), which has given rise to a lively dispute with overseas countries; on this subject, for an initial view, see F. Gaspari, n 3 above, 1156.

[18] See *amplius*, M. Clarich, n 5 above, 232-234.

[19] Extensive analysis in V. Colcelli, 'La natura giuridica dei certificati verdi' *Rivista giuridica dell'ambiente*, 179-181 (2012).

II. Issues: 1) The Role of the Public Entity

The system briefly described above reveals interesting aspects from a legal viewpoint, which, for that matter, becomes apparent from an examination of the literature on the subject. In particular, with regard to private law, attention is focused on the legal nature of the emission allowances and on the fact that the allowances can represent tradeable assets.[20]

This path of research confirms the sensitivity that contemporary civil scholars show towards the concept of the market and its dynamics, thus developing a profitable dialogue with other disciplinary fields, also outside legal science in the strict sense.

Starting from these premises, a different research path is illustrated below, focusing essentially on two aspects: first, the interaction between private and public law instruments in the construction and functioning of the market mechanisms resulting from the legislative framework.[21] Second, the effective suitability of such mechanisms to pursue the environmental purposes which, ultimately, represent (or should represent) the end purpose of the regulatory provisions.

Therefore, starting from the first of the theoretical issues, the debate that has developed so far on the subject aims to emphasise the artificial nature of the market that legislation has intentionally created progressively around emission allowances.[22] It is a market which, in fact, does not exist in nature, in the sense that it does not arise spontaneously from the relationships between individuals and/or other subjects recognised by the legal system. On the contrary, this market necessarily requires the existence, upstream, of an action carried out on the initiative of a public authority, which first and foremost defines the total amount of pollution allowances which it then distributes among the potential polluters.

It can be seen, however, that the role of the public authority goes much further, since it regulates the functioning of the market in an incisive way. It is sufficient to consider, at this point, the choices regarding the criteria for assigning the allowances to the various operators involved,[23] and in particular, the adoption of a system based on the previous industrial history of each specific operator (known as *grandfathering*), or the preference for an auction system,[24] therefore a

[20] See, in particular, C. Camardi, n 5 above, 972; as well as E. Lucchini Guastalla, 'Il trasferimento delle quote di emissione di gas serra' *Nuova giurisprudenza civile commentata*, 290 (2005).

[21] Considerations in P. Lazzara, 'La regolazione amministrativa: contenuto e regime' *Diritto amministrativo*, 342 (2018).

[22] See C. Camardi, n 5 above, 975; as well as M. Clarich, n 5 above, 225. On this subject, see also V. Jacometti, *Lo scambio* n 4 above, 9; M. Cecchetti and F. Grassi, n 5 above, 305.

[23] Extensively illustrated by V. Jacometti, *Lo scambio* n 4 above, 103-105, in which the so-called cap-and trade systems are examined, as well as the baseline-and-credit systems. On this subject, see also M. Cecchetti and F. Grassi, n 5 above, 306-307, as well as F.L. Gambaro, 'Il recepimento della direttiva "Emissions Trading" ' *Contratto e Impresa Europa*, 537 (2007).

[24] Which ought to be the preferable system, at least when functioning regularly; on this subject, A. Gratani, 'Le "quote" per inquinare: a titolo gratuito o oneroso?' *Rivista giuridica*

system based mainly on a competitive logic.[25]

The market mechanisms, ie, the private law provisions on which the market is based, gain relevance only after the decisions have been adopted by the aforementioned authority. For example, in a system of competitive auctions the commercial attractiveness of the 'legal asset' represented by the emission allowance is strictly linked to the way in which the public authority has calibrated the competitive value of the allowance.[26] In particular, it is linked to its capacity to render effectively comparable the alternatives of investing either in the financial instruments represented by the allowances or in the technologies that reduce polluting emissions.[27]

Therefore, it can be said, perhaps provocatively, that the market exists if, and to the extent that, it is made possible by the decisions adopted by the administrative authority.[28] This statement leads to considerations regarding a particularly delicate matter which, in essence, concerns the extent to which the decisions taken by the public administration in this specific sphere should be subject to control and review.[29] From an examination of case records, it appears

dell'ambiente, 396 (2013).

[25] This point is of particular importance, both in the relationship between companies competing on the same market and in the relationships between States with different levels of development. On this subject see F. Giglioni, n 3 above, 133; as well as V. Jacometti, *Lo scambio* n 4 above, 114. With regard to this second aspect of the problem, interesting considerations can be found in K. Peterková Mitkidis, n 1 above, 57.

[26] This point is illustrated by V. Jacometti, *La direttiva* n 14 above, 278.

[27] Discussed by B. Pozzo, *Le nuove regole* n 16 above, 97. The author states that the instrument of pollution rights minimises the individual and collective costs of reducing emissions, since reductions occur when their cost is lower. In addition, it represents an effective incentive for technical progress in the field of pollution control technologies. The public authority, in fact, cannot keep updated on all the technical possibilities available to individual plants, while the flexibility of tradable pollution rights makes it possible to exploit the full potential of the technological initiative of private operators, which is a strong stimulus for innovation. See also V. Jacometti, *Lo scambio* n 4 above, 108.

[28] C. Camardi, n 5 above, 975 and 987, focuses on the central role played by the discretion exercised by the appointed authority. On this subject, see also M. Cecchetti and F. Grassi, n 5 above, 305, in which the functions performed by the public authority in this sphere are illustrated in depth.

[29] Interesting statements were given, for example, by the European Court of Justice in Case C-203/12 *Billerud Karlsborg AB* v *Naturvårdsverket*, Judgment of 17 October 2013, available at www.eur-lex.europa.eu: *'27. The overall scheme of the directive is thus based on the strict accounting of the issue, holding, transfer and cancellation of allowances, the framework for which is provided for by Article 19 thereof and requires the establishment of a system of standardised registries through a separate Commission regulation. That accurate accounting is inherent in the very purpose of the directive, consisting in the establishment of a Community scheme for greenhouse gas emission allowance trading, which aims to reduce greenhouse gas emissions in the atmosphere to a level that prevents dangerous anthropogenic interference with the climate system, with the ultimate objective of protection of the environment (….). As observed by the Commission, in introducing itself a predefined penalty, the Community legislature wished to shield the allowance trading scheme from distortions of competition resulting from market manipulations.*

28. In that regard the Billerud companies' argument to the effect that they cannot be blamed for excessively environmentally harmful conduct must be rejected. Article 16(3) and (4) of the

that there have been a rather limited number of lawsuits, although the statements of principle are not devoid of interest, as will be illustrated further below.

Anyway, such cases arise from disagreement between, on one hand, the public administration vested with the power to 'set up' the market and, on the other hand, the single industrial operator which complains of prejudice to the organisation of his own productive activity. However, from the records of the (limited number of) lawsuits, a profile which would probably merit attention has not yet emerged: the reference regards the interests of the inhabitants and the consequences that they suffer – actually or potentially – as a result of the decisions taken by the public authority in this field. On further examination, moreover, the fixing of the emission allowances, both the total amount and the allowances of each single industrial plant, affects the population living in the area in which the polluting emissions occur, as well as the anthropic activities that take place in that area.

Therefore, there is some basis to the conclusion according to which society in the area concerned and/or the individual claimants have a legally relevant interest worthy of protection also as regards to the decisions taken by the public authority. Against this conclusion, it could be objected that the provisions of the competent national authority are, in fact, the 'offspring' of the system of international rules (regulations and/or agreements)[30] briefly referred to in the initial part of this contribution. The objection could have weight in a lawsuit, but – at least from the theoretical doctrinal point of view – it introduces another very delicate issue, which ultimately involves the identification of the subjects empowered to protect the environment,[31] in this specific case the atmosphere.

The subject is obviously too vast to be adequately examined here. However, at least one aspect deserves to be mentioned, namely concerning the transparency of the decision-making process through which the activity of the public entity in this field passes.[32] The assumptions and criteria that guide the authority in determining the total amount of emission allowances and their distribution should be available to and 'traceable' by the community concerned, precisely because of the direct impact of said allowances on the individuals and on their activities.

directive has as its object and effect to penalise not 'polluters' generally, but rather those operators whose number of emissions for the preceding year exceeds, as at 30 April of the current year, the number of allowances listed in the section of the surrendered allowance table designated for their installations for that year in the centralised registry of the Member State to which they report under Article 52 of Regulation No 2216/2004. This – and not the emissions which are per se excessive - is how the concept of 'excess emissions' is to be construed' (our underlining).

[30] On this subject, F. Giglioni, n 3 above, 125; B. Pozzo, *Il nuovo sistema* n 1 above, 6. Also mentioned in M. Cecchetti and F. Grassi, n 5 above, 320, who emphasise the progressively increasing role of the European Commission in respect of the Member States.

[31] The question of legitimacy with regard to private regulatory law is clearly expressed by M.W. Hesselink, 'Private Law, Regulation, and Justice' 22 *European Law Journal*, 693 (2016).

[32] On the risk of the "capture of the regulator" in this field, see F. Giglioni, n 3 above, 135; as well as P. Lazzara, n 21 above, 352.

The considerations expressed below are based on the belief that the protection of environmental assets (understood in this case – at least in the first instance – as the healthiness of the air) is a question of legally relevant interest, for the protection of which widespread legitimacy can be recognised. In my humble opinion, this is a necessary conclusion if one wishes to adopt a constitutionally-oriented interpretation of the provisions in question.

This first part of the analysis, therefore, aims to overcome the, as it were, solipsistic approach which seemed to appear in certain legal discussions on the subject. It is true that the market of emission allowances is an artificial market, based, first, on the decisions taken by the competent authority and, second, on the decisions adopted by the (industrial and financial) operators that trade on this market. The legal expert that pays attention to the system as such, however, cannot and must not forget that the ultimate reason behind the 'invention' of this market is the pursuit and, hopefully, the achievement of an environmental purpose, namely the reduction of polluting emissions and the protection of the ecosystem.

III. (Continued): 2) The Role of the Market

The time has now come to analyse more closely the role that the private law measures play in the regulatory strategy aimed at containing greenhouse gas emissions. As already mentioned, there is a tendency in specialist literature to emphasise that the implementation of these measures arises from the need to overcome the so-called mechanisms of command and control, of a purely authoritative nature.[33]

This approach is based on a certain number of needs: first, the awareness of the difficulty in altering, from the outside and during 'work in progress', the methods by which certain industrial and production activities are performed, activities that have developed (and which often prosper) due to the absence of specific provisions. Second, and relatedly, the need to avoid direct conflict with important economic operators and the consequent risk of litigation. More generally, emphasising the global, and therefore transnational, nature of the problem of atmospheric pollution, the adoption of purely authoritative instruments obviously clashes with the difficulty (*rectius*, impossibility) of identifying a recognised authority which can be considered as empowered to issue binding prescriptions to all potential polluters, regardless of their nationality and, therefore, of their respective roots in the jurisdiction of a specific national state.

In light of these premises, the regulatory strategy for the containment of polluting emissions was therefore developed on the basis of the introduction of measures of a consensual nature, which would allow the adoption of incentive mechanisms. In other words, instead of imposing an (impractical) authoritative

[33] See n 5 above.

reduction of emissions, a regulatory system has been created which – at least ideally – aims to reward the operator which achieves this result. The reward consists of avoiding financial penalties resulting from exceeding the limit and of (potentially) obtaining the additional economic benefits generated by the 'marketing' of unused emission allowances.

In the above-outlined scenario, the legal expert's technical work consisted firstly of precisely defining the notion of emission allowance, qualifying it as an asset that can be precisely identified and traced.[34] Second, the circulatory mechanisms of the asset thus identified were regulated; the culmination of this development was the 'arrival' of the asset within the sphere of financial intermediation, at which point complex rules became involved to regulate the functioning of a given market and to identify the behaviour which the relative players had to adopt.

It is not within the scope of this work to dwell on the technical aspects of this evolution and, above all, of the connection to the rules of financial intermediation. Nevertheless, this extension is not without consequences from a systematic point of view; the legal expert who pays attention to this perspective, therefore, must perforce make certain considerations, at least as regards the viewpoint of 'values' and 'principles'.

Proceeding step by step, it must first be emphasised that, in this way, the scope of circulation of the allowances has been extended to a wider circle of subjects than the polluters (virtuous and non-virtuous), since it also includes professional investors and third parties. This corresponds to a significant change in the overall approach of the regulatory strategy: the restricted circulation directly counterposed 'polluters' and virtuous operators, and therefore entailed easily measurable consequences at the environmental level.

However, circulation according to the techniques of financial intermediation also attracts and involves other operators, external to the environmental problem, as well as subjects with their own autonomous investment objectives.[35] At least at first sight, this extension lends itself to a double interpretation: for the virtuous operator, it fosters the circulation of the allowance, which becomes more attractive and, by effect, encourages tradeable wealth. For the non-virtuous operator, the opposite reaction is generated, since access to the allowance, as a financial instrument, is now easier and therefore more attractive, at least potentially, compared to the actual reduction of polluting emissions.

In practice, therefore, the fact that the allowances become tradeable assets introduces, within the initially envisaged incentive mechanisms, a new manner of measuring their value, which depends on their performance, as securities, on the market and, more generally, on the factors that influence investment decisions. In this scenario, the individual entrepreneur could find it more convenient to bet –

[34] On this subject, V. Jacometti, *Lo scambio* n 4 above, 428-429; M. Clarich, n 5 above, 229; F. Mocci and J. Facchini, n 14 above, 4.

[35] This point is examined by F. Gaspari, n 3 above, 1161-1162.

also in the medium-long term – on the performance of the asset instead of deciding to reduce the polluting emissions generated by his production activity.

This does not mean that market trading should be disparaged, nor the important role that it plays in contemporary society in supporting productive activities and in the creation of wealth and well-being.[36] However, there is a problem of value consistency between the environmental purposes underlying the regulations in question and the contribution that can derive from the connection with the logic and dynamics of financial intermediation.[37]

The potential conflict between the two perspectives is admirably considered by our administrative case law, with words that are useful to report in full:

> The monetisation of pollution allowances that are exceeded, through the purchase of green certificates, if it represents a legitimate alternative to the primary obligation to use renewable sources, cannot however constitute virtuous behaviour. In fact, as this Council of State has already advised (division VI, 6 July 2006, No. 4290), the behaviour of operators that have merely purchased those certificates cannot be rewarded, since this choice, in addition to being less virtuous, does not produce any increase in production capacity from alternative sources, which increase represents the real objective pursued by the legislator (European and Italian).[38]

The declaration regards renewable energy and green certificates, but there is no doubt that, in terms of value, the statement of principle contained therein can also be extended to the sphere of polluting emissions and greenhouse gases. With the intention of tracing the threads of the discussion so far, the above consideration provides food for thought on the continuing relevance of the 'polluter pays'[39] principle, according to which for some decades now has inspired the introduction of regulatory measures on environmental matters – at least at European level.[40] Already at first glance, it can be seen that the principle is based

[36] However, see the considerations of G. Ferrarini, 'Il Testo Unico della Finanza 20 anni dopo' *Rivista delle società*, 5 (2019).

[37] On this potential conflict see F. Annunziata, n 6 above, 798; the author recognises the difficulty of attaining a balance between protection of the environment and protection of the financial market. In decidedly more explicit terms, see M. Cafagno, 'Cambiamenti climatici tra strumenti di mercato e potere pubblico', in G.F. Cartei ed, *Cambiamento climatico e sviluppo sostenibile* (Torino: Giappichelli, 2013), 115.

[38] Consiglio di Stato 17 June 2014 no 3051, in the reasons; in the same sense, also Tribunale Amministrativo Regionale Lazio-Roma 16 March 2010, nos 4086 and 4090. All these decisions are available on the electronic database dejure.

[39] M. Meli, *Il principio comunitario "chi inquina paga"* (Milano: Giuffrè, 1996) is still up to date, especially in the part in which she illustrates the historic origin of the principle and its theoretic bases (see pages 26 and 51, with specific reference to atmospheric pollution).

[40] The point is explicitly discussed by G. Conte, 'Rischio di impresa e tutela dell'ambiente. Nuovi paradigmi di governo delle decisioni e nuovi modelli di ripartizione delle responsabilità', in G. Alpa et al eds, n 1 above, XXIII. Also mentioned in F. Fracchia, 'Cambiamento climatico e sviluppo sostenibile: lo stato dell'arte', in G.F. Cartei ed, n 38 above, 22.

on the assumption that there is some kind of equivalence between environmental values and market dynamics, which inevitably leads to supposing that the former can be substituted by the latter.

Certainly, from the historic viewpoint, the idea on the basis of which the protection of the environment can be monetised has performed a merit-worthy function, as it has induced the productive and entrepreneurial classes to internalise to some extent the environmental variable within their respective cost structures. Nevertheless, the facts lead us to maintain that the environmental problems are probably more serious than we are inclined to suppose and that, as a result, more incisive regulatory measures are needed than those that can derive from the recourse, albeit 'induced', to market dynamics.[41]

IV. Prospects for Investigation

The examination carried out in the above paragraphs opens up numerous prospects for investigation and research, entrusted to the sensitivity of the individual interpreter. Within this range, the legal expert who pays attention to the system as such probably tends to question the role and function of private law in today's socio-economic context.

The historic path followed by this branch of legal science gives us the idea of a system that tends to be self-referential, and in any case certainly autonomous and independent from the other partitions of the legal system. In the legal tradition of the western world, the zenith of this path is probably the Napoleonic coding, deeply characterised by the affirmation of bourgeois individualism on the surrounding reality, including – as far as relevant in this case – that composed of the environment and natural resources. As it is well known in Italy, this approach continues to be a significant feature of university courses.

In the meantime, however, a two-fold evolutionary path can be observed. On a purely internal level, the system of private law has been progressively enriched thanks to the 'contamination' of (*rectius*, interaction with) the Constitutional Charter and, above all, with the table of values represented therein, which is still highly relevant.[42] Because of this interaction, the implementation of constitutional principles and values passes (also) through the traditional institutions of private law, as can easily be seen by an examination of the evolution of case law trends on property, contracts and civil liability, not to mention the areas regarding family law and succession. Even without wishing to enter into

[41] Consistently, M. Meli, n 40 above, 72, observes that the literature in question falls back on so-called second-best solutions: it does not aim at an optimal level of contamination, but only at an acceptable level. This assessment is linked to considerations of a political nature, which confirm the difficulty of putting into practice the theoretical model described by the economists.

[42] On this subject see P. Perlingieri, *Il diritto civile nella legalità costituzionale secondo il sistema italo-comunitario delle fonti* (Napoli: Edizioni Scientifiche Italiane, 3rd ed, 2006), 192.

the merits of the individual issues, one cannot help noting that the institutions of private law have become instruments for the pursuit of purposes that lie outside its historically delimited sphere, being rooted in a different text, namely the Constitution.

Taking the liberty of using a provocative expression, there is thus a sort of exploitation of private law,[43] and this phenomenon is even more evident if one considers the evolution of European law.[44] In this regard, the relationship is, indeed, no longer represented in terms of values, but becomes purely technical, in the sense that under the European unitary system the institutions of private law become functional towards the economic policy objectives expressed by its own legislative acts and, in particular, its regulations and directives.[45]

As it is well known, the institutions of private law are characterised by their (generally) consensual nature, as well as by their main aptitude to satisfy idiosyncratic interests. These elements help to explain its wide use also by public authorities in order to meet the general needs underlying their respective institutional mission.

However, it is legitimate to ask whether this approach maintains its lasting effectiveness in the case of the protection of the environment and the safeguarding of natural resources, which – as mentioned in the introductory part of this work – call for increasingly more urgent and incisive measures. To put the question in more explicit terms, can environmental protection actually be achieved through the market (ie, the instruments of private law)? Or is it necessary to rethink the relationship between the idiosyncratic approach and the authoritative approach, moving the point of balance towards strengthening the role of the latter at the expense of the former?[46]

[43] Which, incidentally, leads to the need to rethink the fundamental categories of private law. For comments in this sense, with reference to the institution of the contract, see E. Gabrielli, 'La nozione di contratto e la sua funzione. Appunti sulla prospettiva di una nuova definizione di contratto' *Giustizia civile*, 309 (2019).

[44] Lucid considerations in G. Vettori, 'Il diritto privato europeo fra legge, Corti e diritti' *Rivista trimestrale di diritto e procedura civile*, 1350 (2018). On the difficulties, especially the methodological and interpretative difficulties, of this path, see the recent work of E. Bargelli, 'La costituzionalizzazione del diritto privato attraverso il diritto europeo. Il Right to respect for the home ai sensi dell'art. 8 Cedu' *Europa e diritto privato*, 59 (2019). In foreign literature, without pretending to be exhaustive, see H. Collins, 'The Revolutionary Trajectory of EU Contract Law Towards Post-national Law', in S. Worthington et al eds, *Revolution and Evolution in Private Law* (Oxford: Bloomsbury, 2018), 315; M.W. Hesselink, n 32 above, 683; P. Verbruggen, 'Introduction: Regulating Private Regulators: Understanding the Role of Private Law' 27 (2) *European Review of Private Law*, 177 (2019); O.O. Cherednychenko, 'Rediscovering the public/private divide in EU private law' 26 (1-2) *European Law Journal*, 6 (2019); Ead, 'Public and Private Enforcement of European Private Law: Perspectives and Challenges' 23 (4) *European Review of Private Law*, 481 (2015).

[45] Within the vast literature, see, in particular, H.W Micklitz, 'The Visible Hand of European Regulatory Private Law – The Transformation of European Private Law from Autonomy to Functionalism in Competition and Regulation' 28 (1) *Yearbook of European Law*, 28 (2009).

[46] On this subject, a clear-cut position is taken by M. Libertini, 'Persona, ambiente, sviluppo: ripensare la teoria dei beni', in *Benessere e regole dei rapporti civili* n 16 above, 481.

Although modest, it can be seen that the Italian experience provides a clearly positive answer to this last question if one merely considers the role of liability in tort, first exalted by legge 8 July 1986 no 349, strongly revised twenty years later in the Consolidated Law on the Environment,[47] in favour of the greater space given to public administration initiatives.[48] To come to the specific subject of this work, it is easy to add that the parties to a contract by which the emission allowance is exchanged (a trading scheme based on that of derivatives) simply aim to pursue their individual wealth,[49] without necessarily caring for 'external' environmental matters.

In fact, the use of a mercantile (or market-based) logic to prevent the contamination of environmental resources entails an underlying conceptual flaw, as it presupposes that environmental resources can have a cash value and, as such, can be the subject of trading transactions.[50] This conviction itself represents progress compared to an approach according to which these resources would be free for appropriation without the involvement of any counterparties,[51] as was believed for a long time, for that matter, and which continues to be maintained in many social and territorial contexts.

However, the importance, if not the urgency, of the environmental issue seems to require more incisive forms of regulation than those that can be ensured by market dynamics.[52] This opinion is also expressed by the arguments that pay attention to inter-generational balance and, therefore, to the duty of each generation to leave the 'common home' in (relative) order for their descendants.

It is therefore possible to catch sight of a *pars construens* of regulatory strategies and instruments which will inevitably also result from an in-depth review of the traditional institutions of private law. An arduous task, but certainly worthwhile, given the importance of what is at stake.

47 Decreto legislativo 3 April 2006 no 152.

48 On this subject, see U. Salanitro, 'Responsabilità ambientale: questioni di confine, questioni di sistema' *Juscivile*, 508 (2019).

49 See the clear comments of M. Barcellona, 'I derivati e la circolazione della ricchezza: tra ragione sistemica e realismo interpretativo' *Europa e diritto privato*, 1104 (2018); see also, however, the considerations of M. Pennasilico, 'Sviluppo sostenibile, legalità costituzionale e analisi "ecologica" del contratto' *Persona e mercato*, 38-39 (2015).

50 Comments in this sense in B. Pozzo, *Il nuovo sistema* n 1 above, 96. Recalling Harding's thesis notes on the so-called tragedy of the commons, the author states that the underlying idea of transferable pollution rights (economic instruments of a proprietary type) consists of the theory according to which environmental degradation is the result of the incomplete attribution of proprietary rights relating to the use of the natural resources.

51 Comments by M. Clarich, n 5 above, 219.

52 On this subject see also M. Libertini, n 46 above, 489.

Essays

Questioning Representative Sovereignty: The Italian Head of State in 'Post-State' Constitutional Law

Giuliano Vosa*

Abstract

The Italian constitutional order is undergoing a slight but salient shift as regards the role of the Head of State, who is called on to take delicate political positions while acting as a liaison between the national and supranational stages.

This work aims to investigate this shift and its consequences to analyse how a State's constitutional structure evolves as confronted with the post-State reality. Starting with an account of 'representative sovereignty' to locate Heads of State in contemporary parliamentary governments, it takes as reference a speech delivered by the President in 2018 and examines in this light the constitutional practices of the last ten years as well as some of the most recent activities involving the President. The picture the work aims to paint exposes the ties between the national and supranational levels.

Whether this picture coheres with the overall national constitutional architecture is doubtful; however, the fundamentals of 'representative sovereignty' as accounted for in the introductory part no longer work well together, and this challenges some of the cornerstones of contemporary constitutionalism.

I. Introduction. 'Representative Sovereignty': Domestic Rigidity and Supranational Openness

In one of his least famous works, a maverick of the early 1900's Italian legal philosophy, Giuseppe Capograssi, pointed to the breakdown of a key constitutional concept that he called 'representative sovereignty' to account for the decline of the liberal State.[1] In 1922 he argued that the mounting crisis of the State's constitutional arrangements lay in a relatively 'new' phenomenon: as social pluralism rose to an unprecedented magnitude, national institutions were facing growing difficulty in accommodating diverging interests by means of legislation – which, as a consequence, decreased their authority and

* *Manuel García Pelayo* Fellow, *Centro de Estudios Políticos y Constitucionales*, Madrid. I wish to thank Professors Iyola Solanke (University of Leeds) and Thomas Poole (LSE) as well as Professors Nicola Lacey and Martin Loughlin (LSE) for their crucial comments on earlier versions of this article. All mistakes remain my own.

[1] G. Capograssi, 'La nuova democrazia diretta', in Id, *Opere*, (Milano: Giuffrè, 1959), I, 475, 485-486.

effectiveness.[2] This failure to link the institutions engaged in law-making with a fast-changing society undermined the foundational legitimacy of national sovereignty by weakening its representative support.[3]

The proper remedy to the 1920's crisis, in Capograssi's view, was the construction of more inclusive architectures to channel the rising pluralism into mediation paths directed by representative institutions. Otherwise, he foresaw, executives would seek more direct, possibly non-parliamentary ties to communicate with society and to respond to its needs, with two consequences, seemingly at odds with each other, yet in fact concurrent.[4] First: the government and the Prime Minister would come to occupy the centre of domestic constitutional orders. Second: the role of the Head of State 'may be reconsidered, perhaps re-construed' in a denser political fashion.[5] In fact, he argued, the new social forces would look at the Head of State as their best-suited institutional interlocutor;[6] consequently, under the mounting pressure of the most powerful among such forces, she would be prompted to marginalise the Parliament and secure a new social pact in what would look like a renewed, authoritarian version of constitutional monarchies.[7]

Therefore, he held, Heads of State would claim a legitimacy of their own – in the mode of Constant's '*pouvoir neutre*' –[8] pursuant to which they would make political choices aiming at either inclusion or exclusion of legitimate interests. Yet, the latter option would add to the State's crisis:[9] absent sufficiently solid representative ties, the claims of the excluded would turn into a powerful element of destabilisation for the State itself, both in the international arena and at home.[10]

It is understood that, in comparison with his contemporaries, Capograssi pioneered an innovative approach to political representation;[11] his focus fostered a shift from the state-centred 'institutional' paradigm to address 'the people' in its multifaceted composition, and dared to look beyond the dominant

[2] For a comparison, see L. Duguit, *Traité de droit constitutionnel* (1912; Paris: Ancienne Librairie Fontemoing &Co., 3rd ed., 1921), I, 606-609.

[3] C. Vasale, *Società e Stato nel pensiero di Giuseppe Capograssi* (Roma: Edizioni di storia e letteratura, 1972), 121.

[4] G. Capograssi, n 1 above, 558.

[5] ibid, 559-560.

[6] ibid, 566. See comments in C. Vasale, n 3 above, 120.

[7] F.J. Díaz Revorio, 'La monarquía parlamentaria, entre la historia y la Constitución' 20 *Pensamiento Constitucional*, 65-106 (2015).

[8] B. Constant, *Réflexions sur les constitutions, la distribution des pouvoirs et les garanties dans une monarchie constitutionnelle* (Paris: Nicolle, 1814), 3.

[9] G. Capograssi, n 1 above, 570. See V.E. Orlando, 'La decadenza del governo parlamentare' 2 *Rassegna di scienze sociali e politiche*, 1, 589-600, 598 (1884).

[10] See C. Vasale, n 3 above, 133.

[11] *See H. Hofmann*, *Repräsentation. Studien zur Wort-und Begriffsgeschichte von der Antike bis ins 19. Jahrundert* (1974), translated by C. Tommasi, *Rappresentanza-rappresentazione: parola e concetto dall'antichità all'Ottocento* (Milano: Giuffrè, 2003), 415.

organicist conceptions of the State.[12] Yet, he wrote in an age of diffuse tensions within and among States, which resulted in the Second World War. Contrary to his hopes, the constitutional structures of liberal States did not make room for inclusive claims to representative sovereignty; rather the opposite, they mostly collapsed and paved the way to authoritarianism.[13] Domestically, constitutional flexibility could not counter the rise of autocratic governments that restricted, instead of broadening, social participation in the deliberation of the State's will.[14] Internationally, autarchic conceptions of sovereignty maintained by aggressive elites hardly tolerated any limitations to the States' power to defend their interests against one another.[15]

Consequently, representative sovereignty led to restricted participative spaces for law-making, whereas Heads of State went on to play an exclusive, rather than inclusive, role in refurbishing the ties between the State's power and a fast-changing society.[16] Far from protecting pluralism, the *Hüter der Verfassung* came to resemble the intimidating figure of Schmitt's *decider of last resort*, preserving national unity notwithstanding the interests of the excluded – even if their exclusion might lead to the infringement of established rights.[17]

After the war and with the awareness of the massive violations of human rights that had occurred worldwide, national flexibility and international autarchy were purposely abandoned. To counter the rise of autocratic and aggressive nationalistic governments, representative sovereignty was reformulated in a pluralistic fashion that may be described as a link between domestic rigidity and supranational openness. This link builds up a mutually positive relationship, where the former is instrumental to the latter and the

[12] *'Institutional representation' as a concept attached to the State's organs regardless of their actual ties to society is well-rooted in continental scholarship: see* V.E. Orlando, 'Du fondement juridique de la representation politique', 2(2) *Revue du droit public et de la science politique en France et à l'étranger*, 1-39 (1895) and V. Gueli, 'Il concetto giuridico della rappresentanza politica e la rappresentatività degli organi di governo' III-IV *Rivista italiana per le scienze giuridiche*, 239-256 (1942). This concept served as a ground for national representation in Fascist Italy: see O. Ranelletti, 'La rappresentanza nel nuovo ordinamento politico e giuridico italiano' 1(21) *Rivista di diritto pubblico e della pubblica amministrazione in Italia*, 199-206 (1929) and L. Paladin, 'Il problema della rappresentanza nello stato fascista' 1-2 *Jus*, 69-87 (1968). For comparison, see M. Stolleis, *A History of Public Law in Germany (1914-1945)* (Oxford: Oxford University Press, 2014), 17-20, 64.

[13] K. Polanyi, *The Great Transformation: The Political and Economic Origins of Our Time* (1944; Boston: Beacon Press, 2nd ed, 2001), 237-242.

[14] H. Heller 'Politische Demokratie und Soziale Homogenität', in H. Heller ed, *Probleme der Demokratie*, vol. I (Berlin: Walter Rothschild, 1928), 35-47, English ed: A. Jacobson and B. Schlink eds,*Weimar: A Jurisprudence of Crisis* (Berkeley: University of California Press, 2000), 256-258.

[15] B. Mirkine-Gützevich, *Droit constitutionnel international* (Paris: Sirey, 1933), chapter II.

[16] J.A. Sánchez Moreno, 'El Parlamento en su encrucijada: Schmitt *versus* Kelsen, o la reivindicación del valor de la democracia', 162 *Revista de Estudios Políticos*, 113-148 (2013).

[17] C. Schmitt, 'Der Hüter der Verfassung', 55(2) *Archiv des öffentlichen Rechts*, 161-237 (1929).

latter enhances the former:[18] as a result, a virtuous circle protects the self-determination of the human *person* as a supreme expression of her dignity[19] by allowing her a central position in the overall constitutional architecture.[20] In this framework, political and substantive rights stay in a mutually positive relation: the material content of substantive rights is a consequence of the exercise of political rights and of the constitutional structure that unfolds accordingly.

The functioning of this link can be roughly presented as follows. Domestically, rigid constitutions protect individual rights through a separation of powers:[21] the State's institutions accommodate plural interests via rational discourse according to the constitution, as the legislator's will is balanced by constitutional courts' substantive review.[22] At the supranational level, openness replaces autarchy:[23] bi-and multilateral agreements expand the scope of international law,[24] and human rights are protected against States' sovereign will.[25] Furthermore, on the European stage, the principles of primacy and direct effect lead to an increasing legal integration[26] – whereas, politically, the Union strives to encompass formerly rival States into a single order shaped by law rather than by pure power.[27]

[18] M. Luciani, 'La "Costituzione dei diritti" e la "Costituzione dei poteri". Noterelle brevi su un modello interpretativo ricorrente', in *Scritti in onore di Vezio Crisafulli* (Padova: Cedam, 1985), II, 497.

[19] See J. Habermas, 'The Concept of Human Dignity and the Realistic Utopia of Human Rights' 41:4 *Metaphilosophy*, 464-480 (2010) and P. Häberle, *Die Wesensgehaltsgarantie des Art. 19 Abs. 2 Grundgesetz* (Stuttgart: C.F. Müller Verlag, 1983), 179.

[20] The centrality of the person in the Constitution's order gave rise to what has been called the 'Constitution's sovereignty' doctrine: see L. Lacché, 'The Sovereignty of the Constitution: A Historical Debate in a European Perspective' 34 *Journal of Constitutional History*, 83-102 (2017), and G. Zagrebelsky, *Il diritto mite. Legge, diritti, giustizia* (Torino: Einaudi, 1992), 9-10.

[21] See Art 16 of the *Déclaration des droits de l'homme et du citoyen* (1789). On constitutional rigidity, A. Pace, 'La causa della rigidità costituzionale', in Id, *Potere costituente, rigidità costituzionale, autovincoli legislativi* (Padova: Cedam, 2002), 3-97; J.L. Requejo, 'El poder constituyente constituido. La limitación del soberano' 1 *Fundamentos*, 361-380 (1998). More recently, J. García Roca, 'De la revisión de las constituciones: constituciones nuevas y viejas' 40*Teoría y Realidad Constitucional*, 181-222 (2017) and Y. Roznai, 'Rigid (Entrenched) / Flexible Constitutions', *Max Planck Encyclopedia of Comparative Constitutional Law* (Oxford: Oxford University Press, 2018), 1-17.

[22] See P. Ridola, 'Libertà e diritti nello sviluppo storico del costituzionalismo', in P. Ridola and R. Nania eds, *I diritti fondamentali* (Torino: Giappichelli, 2001), I, 3-68, 35.

[23] C. Tomuschat, 'International Law: Ensuring the Survival of Mankind on the Eve of a New Century: General Course on Public International Law', 281(10) *Recueil des Cours – L'Aje*, 306 (1999).

[24] R. Lesaffer, 'Peace Treaties and the Formation of International law', in B. Fassbender and A. Pieters eds, *The Oxford Handbook of the History of International Law* (Oxford: Oxford University Press, 2012), 71-94.

[25] D. Shelton ed, *The Oxford Handbook of International Human Rights Law* (Oxford: Oxford University Press, 2013), 163.

[26] J.H.H. Weiler, 'The Transformation of Europe' 100(8) *Yale Law Journal*, 2403-2483, 2450 (1991).

[27] See J.-P. Jacqué, *Droit institutionnel de l'Union Européenne* (Paris, Dalloz, 5th ed, 2009),

In this framework, sovereignty is no longer thought of as a monolith in defence of State unity:[28] *multi-level* structures replace the state-centred model as the best-suited paradigm for pluralism to thrive at national and supranational levels.[29] Many scholars even argue that, given its ties to a State-centred conceptual background, sovereignty no longer suits the *post*-State scenario.[30]

Be that as it may, the constitutional arrangements that stem from the construction of multiple institutional channels between public powers and plural societies have found a better-defined role for Heads of State. Parliamentary governments[31] – the most diffuse constitutional structure within Europe's public space[32] – conceive of them as politically unaccountable counterweights acting as a last resort to resolve political conflicts beyond the majority's will.[33] In other words, Heads of State are called to play an *inclusive* role for the sake of national unity once no other political resource is available.[34] One might say that Capograssi's lesson has been embraced: the apex of institutional architectures is purposely designed to host socio-political pluralism in a peace-enhancing manner.

However, if one looks at the seething pluralism that ignites European societies from the viewpoint of today's Italian legal order, yet another of Capograssi's predictions seems to come true. Increasing social pluralism entails an increasingly political role for Heads of State, though designed as politically unaccountable; but, as he feared, this seems to come at the expense of inclusion, rather than fostering pluralism. Consequently, due to Heads of

87. Compare J.H.H. Weiler, 'The Community System: the Dual Character of Supranationalism' 1(1) *Yearbook of European Law* 267–306 (1981).

[28] J.H.H. Weiler, 'In Defence of the *Status Quo*: Europe's Constitutional *Sonderweg*', in J.H.H. Weiler and M. Wind eds, *European Constitutionalism beyond the State* (Cambridge: Cambridge University Press, 2003), 7-24.

[29] I. Pernice, 'Multilevel Constitutionalism and the Treaty of Amsterdam: European Constitution-Making revisited?' 36(4) *Common Market Law Review*, 703-750 (1999).

[30] N. MacCormick, *Questioning Sovereignty. Law, State and Nation in the European Commonwealth* (Oxford: Oxford University Press, 1999), 123-136.

[31] In the European Union (the UK being no longer considered as a member) there are 21 parliamentary governments out of 27; Cyprus is the only presidential government, whereas France, Lithuania, Poland, Portugal and Romania are listed as semi-presidential.

[32] See P. Ridola, 'Prime osservazioni sullo "spazio pubblico" nelle democrazie pluralistiche', in Id, *Diritto comparato e diritto costituzionale europeo* (Torino: Giappichelli, 2009), 31-49 and A. von Bogdandy, 'European Law beyond 'Ever Closer Union': Repositioning the Concept, its Thrust and the ECJ's Comparative Methodology' 22(4) *European Law Journal*, 519-538 (2016).

[33] S. Milačić, 'Le contre-pouvoir, cet inconnu' in *Mélanges Lapoyade-Deschamps* (Bordeaux: Presses universitaires de Bordeaux, 2005), 681.

[34] There are notable substantive differences among parliamentary governments themselves; see G. de Vergottini, *Diritto costituzionale comparato*, (Padova: Cedam, 9th ed, 2014), I, 613-620; N. Parpworth, *Constitutional and Administrative Law* (Oxford: Oxford University Press, 9th ed, 2016), 53-59. On the evolution of the royal prerogatives under English constitutional law, see R. Blackburn, 'Monarchy and the personal prerogatives' 3 *Public Law*, 546-563 (2004) and R. Brazier, 'Monarchy and the personal prerogatives: A personal response to Prof. Blackburn' 1 *Public Law*, 45-47 (2005); on the Spanish King's powers, F.J. Díaz Revorio, n 7 above, 75 and I. Torres Muro, 'Refrendo y Monarquía' 29 *Revista Española de Derecho Constitucional*, 43-70 (2009).

States' apical position in the constitutional architecture, presidential claims hinge on the representation of national unity; thus, their opponents suffer an exclusion without remedy, as countermeasures are simply unconceivable – Heads of State are *themselves* the ultimate 'countermeasure' in their own political-constitutional orders.[35]

Three years ago, in a speech whose influence in defining Italy's constitutional structure is still underrated, the Italian Head of State raised a *sovereignty* claim in alleged defence of national unity.[36] Today, the dust of a politically heated issue has blown over and two new parliamentary majorities have relegated that Cabinet in the 'history' section;[37] yet there are still crucial reasons to analyze that claim, as it questions the role of the Italian Head of State in fully-fledged *post*-national contexts.[38] Pressures coming from a space that is external to and independent from the range of a State prove to exercise a slight, but remarkable influence on that State's constitutional setting.[39] Therefore, this is not only an Italian concern, but one that is tied to a broader, 'post-State' scenario.[40]

Pursuant to a summary of those events, this work investigates the substance of the claim the President raised in light of its manifold implications for contemporary constitutionalism.[41] The points that will be discussed are

[35] T. Martines, *Governo parlamentare e ordinamento democratico* (Milano: Giuffrè, 1967), 152.

[36] 'Sovereign' as it pretends to define the sovereign interest of the State: N. Walker, 'Sovereignty Frames and Sovereignty Claims' 14 *University of Edinburgh Research Paper*, 1-26 (2013).

[37] In August 2019, the then Ministry for Home Affairs Matteo Salvini (Lega) ceased to support the Conte Cabinet in the hope of turning his party's increasing growth in the polls into an actual parliamentary majority; however, other major parties (*MoVimento Cinque Stelle* and the Democrats) though rivals in the 2018 campaign, agreed to form a new Cabinet, with Giuseppe Conte as Prime Minister and Democrats in crucial Ministries and an enhanced 'pro-Europe' attitude. See 'Governo, Conte annuncia i ministri', *La Repubblica*, 4 September 2019; 'With New Cabinet, Italy's Political Turmoil Ends, For Now, *The New York Times*, 4 September 2019. Eventually, the new 'crisis' triggered by *Italia Viva* and by its leader Matteo Renzi has led to the formation of a new, fully 'Europeanist' Cabinet led by Mario Draghi and supported by virtually all the political forces (with the exception of *Fratelli d'Italia*). See 'A Giant of Europe Prepares to Head Italy's New Unity Government', *The New York Times*, 12 February 2021.

[38] More in G. Vosa, 'La pretesa "responsabilità istituzionale" del Presidente della Repubblica: un'accorata denuncia dei mutamenti profondi che solcano il diritto dell'Europa' 4 *Rivista AIC*, 186-210 (2019).

[39] G. Scaccia, 'Espansione di ruolo del Presidente della Repubblica e funzione di rappresentanza dell'unità nazionale' 3 *Lo Stato*, 101, 110 (2014) and M. Luciani, 'Il Presidente della Repubblica: oltre la funzione di garanzia della Costituzione', in M. Luciani and M. Volpi eds, *Il Presidente della Repubblica* (Bologna: Il Mulino, 1997), 11.

[40] 'Post-State' being understood as a historical moment in which States have lost the monopoly of lawmaking due to the exclusivity principle's demise. The concept does not imply the irrelevance of States in the international scenario; compare M. Loughlin, 'Constitutional Pluralism: An Oxymoron?' 3-1 *Global Constitutionalism*, 9-30 (2014) and S. Cassese, 'The Rise and Decline of the Notion of State' 7(2) *International Political Science Review*, 120-130 (1986) though it alludes to the fading of formal equality among sovereign States in the shift from 'international law' to 'international relations' (M. Koskenniemi, *The Gentle Civilizer of Nations: The Rise and Fall of International Law (1870-1960)* (Oxford: Oxford University Press, 2002) 127, 440, 465).

[41] Compare A. Somek, *The Cosmopolitan Constitution* (Oxford: Oxford University Press,

anticipated hereinafter.

First: the President claims an 'institutional responsibility'. Yet, in the constitutional text, the only accountability to which he is subject arises in cases of betrayal and supreme violation of the Constitution, for which he would face impeachment. Accordingly, he has a correspondent duty to act only if, and when, the Constitution is at risk of supreme violation: his responsibility descends from the top of the institutional architecture and amounts to a substantive veto on the attributions of other constitutional organs.

Second: his claim refers to an alleged exercise of 'concrete sovereignty', the relevant arguments being *emotional,* rather than rational, in nature. The whole reasoning rests on a partial, politically controversial narrative of events from which the President strives to derive constitutional arguments.

Third: although the wording of the Constitution seems to preclude the development of such a claim, constitutional antecedents may be sought in the practice as evolved since 2011. In those circumstances, President Giorgio Napolitano shepherded fragile coalition governments to ensure compliance with the duties imposed by the European institutions to face the economic crisis. Yet, the stark opposition he had to confront – including with regard to troublesome, controversial political events that tarnished his mandate – would recommend the highest prudence in the evaluation of any arguments for a constitutional mutation that may stem from such practices.

Fourth: further confirmation of this claim has come from the response to another crisis, ie the CoVid-19 pandemic that has been shocking the world since early 2020, and in the aftermaths of the 2021 political turmoil that has led to the appointment of Draghi's Cabinet. President Mattarella has found occasions to strengthen his direct representational claim *vis-à-vis* the citizenry and to stress his role as a liaison between national and supranational orders.

As a result, in light of both recent and less recent circumstances, this slight *modification* of the presidential role seems to reveal an ongoing constitutional shift, which challenges the relation between constitutional rigidity and supranational openness at the roots of national orders. Thus, post-national constitutional arrangements maycontemplate a departure from what has been hitherto regarded as the cornerstone of contemporary constitutionalism.

II. The President's Powers and the Italian 2018 Elections: Innovative Practices in the Appointment of a Cabinet

The Italian President plays a crucial role in the appointment of the Cabinet.[42] It seems opportune to briefly recall the norms defining her *status*

2014), 176, and N. Krisch, *Beyond Constitutionalism: The Pluralist Structure of Post-national Law* (Oxford: Oxford University Press, 2010), 38.

[42] In the Italian literature, the span of presidential powers has been compared to the functioning

and attributions.

Under Art 87 (1) of the Constitution, the President is 'Head of State' and 'represents national unity'.[43] As Head of State, she acts as the linchpin of the national institutional machinery: she has to ensure compliance with constitutional law and practice at the institutional level. As representing national unity, she activates a direct link with the whole Italian society that she is to interpret according to criteria of sound 'humanitarian reasons'.[44]

These two provisions refer to two different channels fuelling the legitimacy of the presidential actions. The first channel has to do with her function as supreme 'constitutional magistrate' that is undergirded by a systematic reading of rules conferring constitutional attributions to the State's organs. The second is embedded in her institutional prerogative and stems from her apical position in the constitutional architecture. Whereas the former relates to interpretation and must be backed by reasonable legal arguments, the latter hinges on the personal qualities of the President herself: it virtually rests solely on her wisdom and sense of justice.[45]

For presidential powers to be prudently used in view of preserving the unity of the Italian State, a sound balance must be struck between these two channels.[46] Considering both a systematic and a strictly textual argument, the constitutional magistracy is mentioned first in the Constitution's wording and has logical and juridical priority over the national unity representation, which should be confined to mostly *symbolic* functions.[47] Hence, should the President violate the constitutional attributions of other organs on the basis of her link with the Italian society, she would probably trespass the boundaries of her legitimacy and steer Italy toward a *quasi*-presidential, monarchy-like government, which would most probably be in breach of the Constitution.[48]

As for the appointment of a Cabinet, it is accepted that the President has incisive powers – since abundant constitutional practices enriched the laconic provision laid down in Art 92(2).[49] In fact, the President's role has come to be

of a squeezebox (Giuliano Amato; see G. Pasquino, 'La fisarmonica del Presidente' 3 *La rivista dei libri*, 8 (1992)).

43 M. Luciani, 'Un giroscopio costituzionale: il Presidente della Repubblica dal mito alla realtà (passando per il testo della Costituzione)' 2 *Rivista AIC*, 18 (2017).

44 G. Scaccia, n 39 above, 101-115; M. Luciani, 'La gabbia del Presidente' 2 *Rivista AIC*, 1-10 (10 May 2013).

45 A. Sperti, *Responsabilità presidenziale e ruolo costituzionale del Capo dello Stato* (Torino: Giappichelli, 2012), 30-33.

46 L. Paladin, 'La funzione presidenziale di controllo' 2 *Quaderni costituzionali*, 309-327 (1982).

47 A. Sperti, n 45 above, 5-17.

48 In case of political stalemate, the President may use its attributions to force a way-out from the *impasse*; a renowned, although controversial theory attaches to his figure a power of *constitutional* direction, symmetric to – but, significantly, separate from – the Government's *political* direction. See P. Barile, 'I poteri del Presidente della Repubblica' 1 *Rivista trimestrale di diritto pubblico*, 295 (1958).

49 Accordingly, '(t)he President of the Republic appoints the President of the Council of Ministers

crucial:[50] it has been recognised that, in some circumstances, she might appoint a Cabinet even with no support in either chamber, something that has occurred a few times.[51]

As a general rule, the President regularly consults with parliamentary forces and may commission explorative mandates to political figures (*mandati esplorativi*) to seek a parliamentary majority; she may confer a pre-appointment (*pre-incarico*) on the person who could most likely receive parliamentary support as Prime Minister (*Presidente del Consiglio dei Ministri*). The candidate may accept (often with reserve) the pre-appointment, then seeks parliamentary consensus and may present within days a list of candidate Ministers whose appointment is discussed with the President.[52]

It is understood that, generally, the President proceeds to the appointment according to the wishes of the candidate President of the Council of Ministers that she had appointed. In fact, the latter acts as a spokesperson of a potential parliamentary majority and expresses the wishes of the respective MPs – the vote of confidence concerned (*voto di fiducia*) taking place in each chamber within ten days of the Cabinet's appointment, as laid down in Art 94 (3) of the Constitution.

In the case under discussion, however, things went differently. Due to the new electoral system, largely proportional, and to the rise of new forces in the face of the decline of the traditional centre-right/centre-left parties, the 2018 general elections in Italy left most analysts bemused because of the changeable political scenario. Multiple negotiation rounds occurred during several weeks and the whole process of appointment of the Cabinet acquired unprecedented visibility.[53] The Euro-critical focus taken during the campaign by some of the parties receiving the most votes raised some international concerns.[54] Eventually, *MoVimento Cinque Stelle* (Five Star Movement) and *Lega* came to an agreement

and, on his proposal, the Ministers'. See S. Galeotti, *La posizione costituzionale del Presidente della Repubblica* (Milano: Pubblicazioni Università S. Cuore, 1949), 10; G. Guarino, 'Il Presidente della Repubblica italiana. Note preliminari' 3 *Rivista trimestrale di diritto pubblico*, 3 b (1951).

[50] M. Carducci, 'Art. 94', in R. Bifulco, A. Celotto and M. Olivetti eds, *Commentario alla Costituzione* (Torino: Utet, 2006), II, 1810; A. Baldassarre, 'Il Capo dello Stato', in G. Amato and A. Barbera eds, *Manuale di diritto pubblico* (Bologna: Il Mulino, 1991), 461; F. Sacco, *La responsabilità politico-costituzionale del Presidente della Repubblica* (Roma: Aracne, 2012), 77.

[51] R. Ibrido, 'La nascita del Governo Fanfani VI ed i problemi costituzionali del governo privo della fiducia iniziale' *federalismi.it*, 1-18 (26 May 2013).

[52] A. Baldassarre and C. Mezzanotte, 'Presidente della Repubblica e maggioranza di governo', in G. Silvestri ed, *La figura e il ruolo del Presidente della Repubblica nel sistema costituzionale italiano* (Milano: Giuffrè, 1985), 92.

[53] C. Pinelli, 'Appunti sulla formazione del Governo Conte e sulla fine della riservatezza' 2 *Osservatorio costituzionale*, 1-10 (2018); M. Fichera, 'Formazione, funzionamento e struttura del Governo Conte: luci e ombre sugli sviluppi della forma di governo italiana', 3 *costituzionalismo.it*, *1-27* (2018).

[54] L. Fontana, ' Le responsabilità di chi ha vinto le elezioni' *Il Corriere della Sera*, 5 March 2018.

on a programme ratified by a written covenant,[55] the whole process being regarded as deeply innovative for Italian constitutional practices.[56]

The crucial events began on 23 May, as President Sergio Mattarella pre-appointed Giuseppe Conte (backed by *MoVimento Cinque Stelle* and *Lega*) who accepted with reserve. For some days, rumours mounted in the press about alleged disagreements on the candidate Ministers.[57] Eventually, on the afternoon of 27 May, Conte confirmed that he 'renounced' the appointment due to 'lack of an agreement' on the Ministers with the President.[58]

On that evening, President Mattarella took the initiative to present himself before the media in the most solemn form to deliver worldwide a speech of remarkable momentum.[59]

The peculiarities of that speech – leaving aside the circumstances leading to its delivery – can be summarized in two points. First, the conscious, evident attempt to sketch out a factual background in support of a specific narrative of the events related to the national and international scenarios. Second: such background – while resting on mostly emotional, rather than reason-based, arguments – was meant to support a clear-cut constitutional interpretation of the presidential powers that is designed to live beyond the specific circumstances and considerably expands the role of the President to the detriment of other organs. In other words: by an apparently unnecessary overexposure – perhaps even politically detrimental to his figure in the short term – President Mattarella gives reasons for the unprecedented role he is to play and strives to translate such reasons into stable constitutional foundations for presidential action.[60]

The *exordium* of the speech briefly recalls the events. The President reveals himself to 'have eased' political forces in the negotiations after the polls and provides accurate details of his actions, which nonetheless seems beyond constitutional practices.[61] He virtually directed two explorative mandataries in

[55] R. Bin, 'Il "contratto di governo" e il rischio di una grave crisi costituzionale' www.lacostituzione.info, 16 May 2018; G. Zagrebelsky, 'Contratto di governo? È patto per il potere' *Il Fatto Quotidiano*, 21 May 2019.

[56] M. Esposito, 'Spunti per un'analisi delle variazioni costituzionali percepibili nel procedimento di formazione del Governo Conte' 2 *Osservatorio Costituzionale*, 1-21 (2018).

[57] M. Damilano, 'La notte più buia della Repubblica e quei serpenti sulla Costituzione' *Editoriale L'Espresso*, 28 May 2018.

[58] T. Ciriaco and A. Cuzzocrea, 'Governo, il giorno della rinuncia di Conte. Ecco come è fallita la trattativa su Savona' *La Repubblica*, 28 May 2018.

[59] Vista – *Agenzia Televisiva Nazionale*, available at https://tinyurl.com/cxx29u7 (last visited 30 June 2021).

[60] As noticed (M. Dani and A. J. Menéndez, 'The "Savona Affaire": Overconstitutionalisation in Action?', available at www.verfassungsblog.de, 31 May 2018) he could have appointed Savona and reminded the public of his guarantee role as regards the State's compliance with international obligations; or else, he could have simply refused to appoint Savona without going public. However, he expressly chose to do otherwise.

[61] M. Esposito, n 56 above, 5.

seeking parliamentary majority, as well as the candidate Prime Minister.[62] In a more conventional reading of the Constitution, he may well exercise his 'influence' on the choices of political parties[63] – through *moral suasion*, or otherwise defined[64]– but may not interfere with their constitutionally provided tasks, ie seeking parliamentary consensus to build up the relation of confidence.[65]

The President underlines that he warned political forces – 'receiving no objection' – of the 'particularly high attention' he was 'to pay' to the choices of 'some Ministries'. Pursuant to this – he declares – he 'accepted all names' proposed, except the candidate Minister of Economy' (Paolo Savona, an internationally recognised economist with a prestigious *curriculum*, born 1936) 'in spite of the consideration' of his personal and professional profile. Yet, whether he has this power is questionable. There are precedents of the President exercising substantive scrutiny of specific candidate Ministers, but they are not comparable to this case.[66] First, because the reasons for such substantive scrutiny referred to moral or functional motives attaching to the person concerned, which in the case at debate the President has explicitly excluded. Second, because the relevant details were never aired to the public, rejection being the result of a cautious, discreet exercise of the President's influence – again: unlike what occurred in this case.[67] No rejection has been recorded that was explicitly grounded on the candidate's political ideas in relation to the relevant post; in a conventional understanding of the presidential figure, this would most probably amount to a political act interfering with the powers of political leaders in the formation of the Cabinet. Nevertheless, the President openly maintains the opposite view: he claims he has a 'guarantor-like role that has never, and could never, tolerate restrictions'.

Three further questions arise. First: who is to benefit from the 'guarantee'

[62] Conte specified that he was 'to renounce the charge' rather than 'not to accept the appointment'. This formula echoes the *Statuto Albertino* provisions (Art 65) referring to the King in a constitutional monarchy: see M. Esposito, n 56 above; compare M. Belletti, *Forma di governo parlamentare e scioglimento delle Camere. Dallo Statuto albertino alla Costituzione repubblicana* (Padova: Cedam, 2008), 363.

[63] L. Elia, 'Appunti sulla formazione del Governo' 2 *Giurisprudenza costituzionale*, 1170 (1957).

[64] D. Galliani, *Il Capo dello Stato e le sue leggi* (Milano: Giuffrè, 2013), II, 513; V. Lippolis and G.M. Salerno, *La Repubblica del Presidente. Il settennato di Giorgio Napolitano* (Bologna: Il Mulino, 2013), 14.

[65] B. Caravita di Toritto, 'I poteri di nomina e scioglimento delle Camere', in A. Baldassarre and G. Scaccia eds, *Il Presidente della Repubblica nell'evoluzione della forma di governo* (Roma-Bari: Laterza, 2013), 104.

[66] See the different positions in www.lacostituzione.info and 2 *Osservatorio costituzionale* (2018); also, D. Tega and M. Massa, 'Why the Italian President's Decision was legitimate', available at www.verfassungsblog.de, 28 May 2019; M. Dani and A. J. Menéndez, n 60 above.

[67] One may say that no 'judicialization' of the relevant positions has ever taken place: A. Stone Sweet, 'Judicialization and the Construction of Governance' 32(2) *Comparative Political Studies* 147-184 (1999).

attached to his role? The President's 'guarantor-like' role relates to his function of *contre-pouvoir* aiming to include minorities beyond the majoritarian will; but it is doubtful whether a majority or minority can be said to exist before a cabinet is appointed, obtains parliamentary support and undertakes any action. Second: how the President chooses the ministries for which he exercises such a 'particular attention' resulting in a veto-like power, for this implies that some ministries are 'more important' than others – contrary to the collegiality principle ruling the Council of Ministers (Art 92). Third: in which sense such a power has 'never tolerated restrictions' since it has never been exercised in these terms?

However, in the follow-up of the speech, there seemed to be no room to respond to any of these questions.

III. An 'Institutional Responsibility' in the Appointment of Ministries: in Search of 'New' Constitutional Grounds

The President's position appears to descend from a systematic constitutional reading that is deliberately *new*. He seemingly claims the existence of a constitutional unwritten norm formed in an extremely short time and without relevant practice; that is, absent the typical constitutive elements of customary norms.[68] This claim, already audacious, becomes manifest when he argues that he bears an 'institutional responsibility' in the selection of candidate ministers 'as the Constitution provides' which impelled him to refuse the appointment of Paolo Savona.

This 'institutional responsibility' is the key of the whole presidential stance. The Constitution contains no such reference, nor does the *genus* 'institutional responsibility' feature anywhere in the text. As the Head of State of a parliamentary government, the President is politically unaccountable: the only check on his actions lies in the cases laid down in Art 90 (high treason and supreme violation of the Constitution) for which he would face impeachment.[69] So long as the Constitution is in force, no other responsibility could attach to his office.

Therefore, his claim arguably exposes an issue of the utmost gravity: the responsibility that the President feels on his shoulders rests nowhere less than at the highest level of the State. His reported duty to act prompts him even to counter majoritarian political forces, as, should he fail to do so, his behaviour would fall within the scope of Art 90. Briefly: President Mattarella is asserting

[68] C. Esposito, 'Consuetudine (*dir. cost.*)' *Enciclopedia del Diritto* (Milano: Giuffrè, 1961), IX, 460; for comparison, A.V. Dicey, *Introduction to the Study of the Law of the Constitution* (8th ed, London: McMillan, 1915; repr: Indianapolis: Libertyfund.org, 1982), 277.

[69] L. Carlassare, 'Art. 90', in G. Branca ed, *Commentario della Costituzione* (Roma-Bologna: Foro Italiano-Zanichelli, 1983), II, 149-189.

on a prime-time broadcast speech that he is acting to prevent a supreme violation of the Constitution, which would arise from the mere appointment of Savona as Minister of Economy.

He deploys a lengthy description of the relevant factual background in support of this claim. The President reveals that for Minister of Economy he wished to appoint somebody who 'could not be seen as a supporter of the Italian exit from the Euro'. Then: *who* should not see Savona as a *No-Euro* supporter, and why would this bring such a menace to Italy? What obscure, yet threatening emergency is the President referring to, and – ultimately – what is the binding force that can be legitimately inferred from such an emergency?

The speech only contains generic explanations linked with a need to secure the 'trust of investors, Italians and non-Italians' and the rising spread rate. Yet, the menace is reported to have concrete implications: the losses on the stock markets are putting at risk the savings of Italian citizens and companies, the safeguard of which the President undertakes as a 'duty' of his own. Consequently, the binding force related to this menace is the highest.

Then, further argumentative support of the President's claim is provided, but mainly – if not wholly – in the form of overtly *emotional* grounds. Three issues are touched upon: indignation (occasioned by trivial comments in the German press);[70] Europeanism (endorsed by loud but vague proclamations of Italy being 'a founding member, and a protagonist' of the European Union); and personal feelings of the President himself ('I am not speaking with light heart').

In sum, the constitutional background that the Italian Head of State is offering to the citizens – and to the whole world – as a support for a sovereign claim in defense of Italy's national unity can be summarized as follows. Due to reasons linked with: 1) unspecified emergencies relating to the potential lack of trust from national and international investors; 2) generic duties to protect Italians' savings; and 3) vague pro-Europe sentiments, the appointment of a Minister of Economy with Euro-critical opinions, chosen by parliamentary actors, is deemed *per se*, and prior to any action (let alone, normative measure) taken by a Cabinet which is still to be appointed, a supreme violation of the Constitution that the President has the 'institutional responsibility' to prevent.

As a corollary, it must be acknowledged that in the President's view the political will of Parliament meets with *substantive* constraints, maybe equivalent to the 'forms and limits provided by the Constitution' to the people's sovereignty (Art 1 (2) of the Constitution) perhaps even to the 'Republican form' (Art 139) that cannot be modified without changing the Constitution.[71] Such

[70] On the numerous provocative headlines appeared in the German newspaper *Der Spiegel*, see 'Copertina con spaghetti a forma di cappio e la frase: "Ciao amore"' *Il Fatto Quotidiano*, 1 June 2018.

[71] In this regard, the President seems prudent: he specifies that 'leaving the Eurozone is a choice of fundamental importance' to be 'discussed openly and seriously, especially if it has not been

constraints, while endowed with nearly irresistible binding force, are vaguely enunciated; and reasons for their juridical nature look evanescent. In the relevant literature, two basic arguments have been formulated to strengthen their anchor in the current constitutional order: the first refers to the duty of protecting the 'savings' undertaken by the Italian Republic and the second to alleged EU/international constraints.

First: under Art 47 '(t)he Italian Republic encourages and safeguards savings in all forms. It regulates, co-ordinates and oversees the operation of credit'. This provision can be linked with the speech as the President mentioned the 'protection of savings' of the Italian citizens as a duty he has to comply with. However, at a closer look, this provision displays a loose relation to the case at hand. In particular, the second sentence shows that it applies to cases in which the savings' legal regulation falls within the State's competence. Yet, in this case, the connection between the protection of savings and the State's scope of action, including a refusal to appoint a candidate Minister, is far from evident. Is there any cause-effect relationship between the two? And, if so, does it bear constitutional relevance? Moreover, is it of such a magnitude to forbid a merely political act like the appointment of a candidate Minister? These questions remain unanswered.

Second: arguments referring to a duty to comply with European Union or international obligations are the most diffuse.[72] There are three arguably suitable constitutional bases. The first is Art 11, concerning the 'limitations of sovereignty' that Italy accepts as an EU member. Second comes Art 117 (1) providing that the legislative competence of the State and the Regions is bound by international and EU obligations. Third, Art 81 (as modified in 2012 pursuant to the 'Fiscal Compact' Treaty) contains the 'balanced budget' rule limiting resort to public debt financing.

The argument contends that these references provide a sufficiently solid constitutional support for the presidential refusal to appoint Savona; because, if read systematically, they prove the existence of legal constraints on the activity of Italian institutions stemming from Italy's membership in the EU and other international bodies.

Although carefully crafted, this line of reasoning is unpersuasive, for – unmistakably – the case at hand does not fall within the scope of any such provisions.

Art 81 is considered to lack actual binding value by most scholars, who wonder whether it may effectively serve its alleged purpose – ie working as a

on the table during the electoral campaign'. The door seems open for a future change; but this does not preclude that it may be considered a change *of* the Constitution rather than *in* the Constitution. See J. L. Requejo Pagés, *Las normas preconstitucionales y el mito del poder constituyente* (Madrid: CEPC-Estudios Constitucionales, 1998), 68.

[72] See S. Curreri, 'Le ragioni di Mattarella nel rifiutare quella nomina, ma lo ha fatto nella sede sbagliata' *www.lacostituzione.info*, 29 May 2018; D. Tega and M. Massa, n 66 above.

parameter to challenge legislation resulting in unbalanced budgets.[73] It has mostly worked as an additional parameter to support the constitutionality of austerity measures that cut off welfare expenditures.[74] In the current case, there is neither a piece of legislation nor a Cabinet calling for the Parliament to approve it, and no infringement of budgetary rules has occurred.

Art 11 reads as a basis to ensure the primacy of EU self-executing law;[75] but no such law is present or even cited in this case.

Art 117 (1) is more explicit: it binds 'State's and Regions' legislative competence' to compliance with EU/international law, and works as a parameter for constitutional adjudication in cases of incompatibility between non self-executing EU/international law and national law. Yet, again, no such law is anywhere at debate.[76]

Even if a systematic constitutional interpretation relying on all these articles is attempted to support a transfer of sovereignty to or a limitation of sovereignty in deference to EU/international institutions, there is still a significant distance between their range of application and the case at issue. The appointment of a Minister is by no means comparable to 'national law' for the purpose of the application of EU (let alone, international) treaties: it is a political act of a sovereign State. Therefore, if it can be constrained by virtue of Italian membership in the EU, or in other international legal orders, this means that the EU, or another international legal order, has authority to restrict the freedom of the Italian institutions to decide on their own composition. Such an authority does not rest on any EU legal basis; is hardly compatible with the purpose of any international treaty; finally, it cannot be justified in light of substantive harmonization of EU or international law, for the simple reason that it *does not harmonize law*, nor does even refer to it. It is a genuine takeover of national sovereignty on the *political* side; one which would certainly clash with Art 4 (2) of the Treaty on the European Union – pursuant to which '(t)he Union shall respect the equality of Member States before the Treaties as well as their national identities, inherent in their fundamental structures, political and constitutional, inclusive of regional and local self-government' – and would most probably operate in breach of the conferral as laid down in Art 5 of the

[73] G. Scaccia, 'La giustiziabilità della regola di pareggio di bilancio' 3 *Rivista AIC*, 1-20 (2012).

[74] L. Carlassare, 'Diritti di prestazione e vincoli di bilancio' 3 *Costituzionalismo.it*, 136-154 (2015).

[75] As interpreted after the Italian Constitutional Court's Judgment 170/1984: see G. Zagrebelsky, *Il sistema costituzionale delle fonti del diritto* (Torino: EGES, 1984), 142; F. Sorrentino, *Le fonti del diritto* (Genova: ECIG, 1997), 28; A. Celotto, 'Coerenza dell'ordinamento e soluzione delle antinomie nell'applicazione giurisprudenziale', in F. Modugno ed, *Appunti per una teoria generale del diritto*, (Torino: Giappichelli, 3rd ed, 2000), 129-270; an analysis of the effects in D. Gallo, *L'efficacia diretta del diritto dell'Unione europea negli ordinamenti nazionali* (Milano: Giuffrè, 2018), 163.

[76] F. Sorrentino, 'Nuovi profili costituzionali dei rapporti tra diritto interno e diritto internazionale e comunitario' 4 *Diritto pubblico comparato ed europeo*, 1355 (2003).

same Treaty.[77]

Briefly: none of these constitutional arguments seems to suit the case, which proves to be an unprecedented one in the seventy-year-long constitutional practice of the Italian Republic.

IV. 'Concrete Sovereignty': An Empirical Description, a Constitutional Argument?

The floor is open as to whether there is a constitutional argument that may support the presidential claim better than those discussed above. The remainder of the speech is of little help in this respect. The President only says that rejecting Savona's appointment is necessary to defend national unity and to protect 'concrete' national sovereignty.

However, 'concrete sovereignty' is not an identifiable concept; it would be misleading to associate it with an international treaty formally providing for the loss of a State party's sovereignty in case of non-compliance.[78] Thus, as no further legal ground is referred to, one is rather urged to look back at the factual substance of the President's claim, which is of the utmost gravity.

The President is openly saying that, while formally being a sovereign State, *concretely* Italy is, since an undetermined moment in time, no longer a sovereign State. To acknowledge that this loss of sovereignty occurred at some point in the past has three fundamental implications. First, a statement of facts: national political leaders have ignored or tacitly accepted both that loss and its consequences, and deliberately concealed this fact from the general public.[79] Second, a political point: representative coverage must be denied to those positions that, though fully lawful under the Italian constitutional order, simply cannot be upheld (anymore) because they clash with that very outer political source to which national sovereignty bows. Third, a legal aspect: the arguments to justify such an exclusion of political positions still in principle lawful cannot be formulated fully, because that exclusion depends on events ranging beyond the control of national institutions. Such events are to be taken as 'facts' that stand beyond the State's will and are not amenable to a 'public use of reason'.[80]

This is why it is up to the President to ensure respect of the duties

[77] K. Lenaerts and J.-A. Gutierrez-Fons, 'The Constitutional Allocation of Powers and General Principles in EU Law' 47(5) *Common Market Law Review*, 1629 (2010).

[78] See A. Cassese, *Diritto internazionale* (Bologna: Il Mulino, 2003), 211 and G. Del Negro, 'The Validity of Treaties concluded under Coercion of the State: Sketching a TWAIL Critique' 10(1) *European Journal of Legal Studies*, 39 (2017).

[79] Compare D. Chalmers, 'The Reconstitution of European Public Spheres' 9:2 *European Law Journal*, 127-189 (2003) and A. Somek, 'Delegation and Authority: Authoritarian Liberalism Today' 21:3 *European Law Journal*, 340 (2015).

[80] J. Habermas, 'Reconciliation Through the Public use of Reason: Remarks on John Rawls's Political Liberalism' 92(3) *Journal of Philosophy*, 109-131 (1995).

stemming from this source: standing at the top of the national institutional structure, he can evaluate how to better comply with them on a case-by-case basis and, by virtue of his prerogative, he would have his 'sound advice' translated into *modified* constitutional law whenever he deems it opportune, so as to adjust the national constitutional setting accordingly.

The impact of this construction on contemporary constitutional arrangements can hardly be overestimated. The concept of representative sovereignty that lies at the foundation of contemporary constitutions was held to construe multiple ties between public powers and a plural society based on a link between domestic rigidity and supranational openness. Now, this link is confronted with a severe challenge. In fact, should one return to the foundations of representative sovereignty and match them with the narrative that the President constructed, she would face an inextricable dilemma.

The presidential argument would proceed as follows. The openness of State constitutions, while aiming to protect pluralism by interconnecting constitutions in a supranational legal order, forced States to enter the global arena without the protection of the exclusivity principle.[81] This was *irreversible*, and also *indispensable* to reject aggressive autocratic nationalism that could jeopardize human rights. Thus, it is both *irreversible* and *indispensable* to stay within the supranational order; all the more so in times of crisis. In fact, a collective breakdown may lead to the demise of that order as a whole; including the disintegration of Italy as a political and constitutional unit, 'lost in translation' from a nation-State to a simple tile in a crumbling supranational mosaic.[82]

Therefore, the supreme violation of the Constitution for which the President feels responsible does not refer to the Italian Constitution only, but to this supranational construction. Decisions on how to confront the challenges caused by the crisis have of course been taken; but by others, although with formal participation of Italian representatives at the time of the decision, and cannot be changed unless others, too, agree on such changes.[83] As a result, Italy could not exercise its own sovereignty without impairing the sovereignty of some other 'sovereigns'; thus, finding itself as 'a clay pot among iron pots'[84] it must

[81] Pursuant to which all legal sources are contained in the Constitution, and there is no other source of law than those recognised by the Constitution. See C. Pinelli, *Costituzione e principio di esclusività* (Milano: Giuffrè, 1990), 61.

[82] For a comparison, G. Piccirilli, 'Il ruolo europeo del Capo dello Stato', in R. Ibrido and N. Lupo eds, *Dinamiche della forma di governo tra Unione europea e Stati membri* (Bologna: Il Mulino, 2018), 392-393, wonders whether presidential power may be arising in parallel with the responsibilities that international actors seem to attach to her figure; M. Ferrara, *Capo dello Stato, vincoli europei e obblighi internazionali. Nuove mappe della garanzia presidenziale* (Napoli: Editoriale Scientifica, 2019), 47-48, argues that such a responsibility may stem from the '*interconstitutional* nature' of the European legal order.

[83] C. Joerges, 'Europe's Economic Constitution in Crisis and the Emergence of a New Constitutional Constellation' 16(5) *German Law Journal*, 985-1027 (2014).

[84] This is a quote from a famous Italian novel: A. Manzoni, *I Promessi Sposi* (1827; 2nd ed, 1842); in the English translation Id, *The Betrothed* (London: Richard Bentley, 1853), 12: the parish

relinquish its autonomy in favour of other, more powerful countries.[85]

This is the apocalypse-like scenario the President discloses as both a framework for institutional action and a constitutional justification for a slight but ineluctable *constitutional mutation* he is to drive by himself, according to the circumstances.[86] Some scholars argue that warning signs of this scenario popped up in the German Constitutional Court's *Maastricht* ruling[87] and turned obvious in further judgments on the Economic and Monetary Union issued in Karlsruhe[88] and elsewhere.[89] Be that as it may, as a common background to all such cases, a general assumption stays unquestioned, to the extent that it is implicitly presumed as a *fact*: [90] there is *no alternative* than to yield to such constraints.[91] President Mattarella acts within this framework: in his view, the appointment of a Minister of Economy who may generate distrust in the markets would certainly cause a threat to the integrity of the State that the State alone could neither prevent nor fix. Clearly, he assumes there is no alternative to the predicted scenario, and thus he must use his prerogative in light of the genuine threat to the integrity of the State that he is to prevent.[92]

Nevertheless, if one looks at the political spectrum resulting from the 2018

priest *Don Abbondio* '…had found himself…like an earthen vessel thrown amidst iron jars'.

[85] See M.A. Wilkinson, 'The Specter of Authoritarian Liberalism: Reflections on the Constitutional Crisis of the European Union' 14(5) *German Law Journal*, 527-560, 542 (2013) and P. Craig, 'The Financial Crisis, the European Union Institutional Order and Constitutional Responsibility' 22(2) *Indiana Journal of Legal Studies*, 243-267, 256-257(2015).

[86] B. Ackerman, *We the People – II: Transformations* (Boston: Harvard University Press, 1998), Part II; compare F. Fernández Segado, 'Las mutaciones jurisprudenciales en la Constitución' 89 *Revista de las Cortes Generales*, 9-88 (2013); M. Luciani, 'Dottrina del moto delle costituzioni e vicende della Costituzione repubblicana' 1 *Rivista AIC*, 1-18 (2013) and M. Dogliani, 'Diritto costituzionale e scrittura', in Id ed, *La ricerca dell'ordine perduto. Scritti scelti* (Bologna: Il Mulino, 2015), 105-132.

[87] *BverfGe*, 2 BvR 2134/92 - 2159/9, 12 October 1993, *Brunner et alt.* v *European Union Treaty*. See C. Joerges, n 83 above, 1001, and C. Joerges and M. Everson, 'Who is the Guardian for Constitutionalism in Europe after the Financial Crisis?' 63 *LSE 'Europe in Question' Discussion Paper Series*, 9 (2013).

[88] See F. Scharpf, 'The Asymmetry of European Integration: Or Why the EU Cannot be a Social Market Economy' 8 *Socio-Economic Review*, 211-250 (2019).

[89] See, in general, T. Beukers, B. de Witte and C. Kilpatrick eds, *Constitutional Change through Euro-Crisis Law* (Cambridge: Cambridge University Press, 2017), Section III, 241-326. A detailed account of the Estonian case in C. Ginter, 'Constitutionality of the European Stability Mechanism in Estonia: Applying Proportionality to Sovereignty' 9 *European Constitutional Law Review*, 335-354 (2013); see also E. Chiti and P.G. Texeira, 'The Constitutional Implications of the European Responses to the Financial and Public Debt Crisis' 50(3) *Common Market Law Review*, 683-708, 695 (2013) and S. de la Sierra Morón, 'Límites y utilidades del derecho comparado en el derecho público: en particular, el tratamiento jurídico de la crisis económico-financiera' 201 *Revista de Administración Pública*, 69-99, 95 (2016).

[90] See W. Streeck, 'The Rise of the European Consolidation State' 15(1) *MPIfG Discussion Paper*, 1-28, 14 (2014).

[91] See C. Joerges and M. Weimar, 'A Crisis of Executive Managerialism in the EU: No Alternative?' in G. de Búrca, C. Kilpatrick and J. Scott eds, *Critical Legal Perspectives on Global Governance: Liber Amicorum David M. Trubek* (Oxford: Hart, 2014), 295-321.

[92] See C. Joerges, n 83 above, 1012; D. Tega and M. Massa, n 66 above.

elections, it is necessary to reconsider this assumption, however familiar and well-founded it might be considered. In fact, as many members of the new elected Parliament (significantly, a potential majority of them) did not agree on the overall narrative above elucidated, President Mattarella came to present as an unquestioned *fact* what was really a *political* position.

To put it clearly: many – potentially, a parliamentary majority – believe that the narrative endorsed by the President is untrue, wholly or in part. In particular, it may be untrue that Italy must be part of the net of intertwining constitutions composing the supranational legal order[93] – adherence to which is both *indispensable* and *irreversible*.[94] It may also be untrue that staying within that order requires obedience to unspecified external constraints that are to be translated into *modified* constitutional law at the word of the President.

In other words, what was presented as a shared, *neutral* framework is now a *partial*, controversial one: a heated *political* issue, which the President nevertheless treated as a plain, unquestioned one. In this regard, it does not matter whether the positions of those who challenge that framework differ, even radically, and do not offer an alternative view; the simple fact that such a challenge is raised bears witness to the political sensitivity of the issues concerned. As a consequence, the President finds himself to: 1) claim authority as Head of State to impose his stance in an exclusive manner, even against a potential parliamentary majority, and to do so 2) in light of a given viewpoint (an understanding of the events and an interpretation of constitutional provisions) that is no longer shared by political forces.

At this point, the question becomes whether precedents can be found in the institutional practice that may back such a claim for a stronger presidential figure.

V. National Political Instability and Incumbent Supranational Duties: Constitutional Precedents of a *Modified*, 'Stronger' President

It is hard to deny that a slight modification of the Constitution as regards presidential powers has occurred in the recent decades.[95] It has been argued that the President would be allowed to resort 'to the (un)expressed potential' enshrined in the Constitution to 'dismantle' an exasperated political

[93] R. Colliat, 'A Critical Genealogy of European Macroeconomic Governance' 18(1) *European Law Journal*, 6 (2012).

[94] See M.A. Wilkinson, 'Constitutional Pluralism: Chronicles of a Death Foretold?' *ARENA WP-7*, 1-28, 20-24 (2017).

[95] Compare O. Chessa, *Il presidente della Repubblica parlamentare: un'interpretazione della forma di governo italiana* (Napoli: Jovene, 2010), 52; G. Scaccia, *Il re della Repubblica. Cronaca costituzionale della presidenza di Giorgio Napolitano* (Modena: Mucchi, 2015), 63; V. Lippolis and G.M. Salerno, *La presidenza più lunga. I poteri del Capo dello Stato e la Costituzione* (Bologna: Il Mulino, 2016), Section I.

fragmentation.[96] However, such modification has admittedly arisen in times of crisis, precisely as in Carl Schmitt's most eloquent predictions,[97] so that its potential to support a constitutional mutation must be carefully weighed.

It seems that a line of continuity in this respect has been drawn and is discreetly but relentlessly being pursued by both Presidents Giorgio Napolitano (who can be seen as the initiator) and Sergio Mattarella.

Notably, the debate ignited as former President Napolitano started to make extensive use of his powers when the 2008 economic crisis began to affect Italy. That period was characterized by high national political instability and emerging supranational duties. Under such 'extreme' circumstances, he emerged as a 'stronger' institutional figure.

Since 2011 approximately, President Napolitano acted as a *de facto* political leader at both national and supranational levels. Domestically, following Berlusconi's resignation[98] and in view of the (in)famous letter received from the European Central Bank[99] he supervised the governments backed by both Democrats and Berlusconi's supporters.[100] On the European and international scene, his energetic presence in foreign policy via the chairmanship of the hitherto marginal Council of Supreme Defense[101] and through the power to concede pardon,[102] as well as his support to government coalitions thoroughly committed to European loyalty, were deemed crucial to secureing Italy's compliance with the rules deliberated at a supranational level to tackle the crisis.[103]

[96] E. Furno, *Il Presidente della Repubblica al tempo delle crisi* (Napoli: Editoriale Scientifica, 2021) at 23.

[97] C. Schmitt, 'Diktatur und Belagerungszustand: Eine staatsrechtliche Studie' 38 *Zeitschrift für die gesamte Strafrechtswissenschaft*, 138-161 (1916). See W.E. Scheuerman, *Carl Schmitt: The End of the Law* (New York: Roman & Littlefield, 1999), 33.

[98] See 'Silvio Berlusconi si è dimesso. La piazza in festa grida "Buffone"' *La Repubblica*, 12 November 2011.

[99] Signed by the then President J.-C. Trichet and by the future President M. Draghi. The full text was soon leaked to the press: see 'Il testo della lettera della BCE al Governo italiano' *Il Sole 24 Ore*, 29 September 2011, available at https://st.ilsole24ore.com/art/notizie/2011-09-29/testo-lettera-governo-italiano-091227.shtml?uuid=Aad8ZT8D (last visited 25 September 2020).

[100] Those forces amounted to what remained of the two coalitions that had been competing for power throughout the 90s and the first 2000s decade. Both were facing a serious hemorrhage of votes towards other political forces. An overview in P. Anderson, 'The Italian Disaster' 36(10) *London Review of Books*, 3-16 (22 May 2014).

[101] G. Scaccia, 'Il «settennato» Napolitano fra intermediazione e direzione politica attiva' 33(1) *Quaderni costituzionali*, 93-108, 101(2013).

[102] A. Pugiotto, 'Fuori dalla regola e dalla regolarità: la grazia del Quirinale al colonnello USA' 2 *Rivista AIC*, 1-6, 4 (2013); M. Luciani, 'La gabbia del Presidente' n 44 above.

[103] More recently, the topic of a supranational 'institutional' responsibility has opened the floor for a rich scientific debate: see G. Piccirilli, 'Il ruolo europeo del Capo dello Stato', and M. Ferrara, *Capo dello Stato, vincoli europei e obblighi internazionali*, both at n 82 above; see also M. Ferrara, 'La Presidenza Mattarella tra politica estera e garanzia interordinamentale' 2 *Quaderni costituzionali* 2020 388-391 and A. Spadaro, 'Dalla crisi istituzionale al Governo Conte: la saggezza del Capo dello Stato come freno al "populismo sovranista"', in A. Morelli ed, *Dal "contratto di*

Yet, acting as a patron of a fragile and heavily opposed coalition pact, his position was the most delicate; and it became all the more so when he requested the Italian Constitutional Court to deliver a harsh judgment proceeding from a controversial set of circumstances in order to back his 'style' in the exercise of presidential powers.[104]

That affair began with President Napolitano raising a *conflict of powers*[105] in July 2012 against the Prosecutor's Office based in Palermo, in a case of wiretapping in which the intercepted person was caught while speaking on the telephone with the President himself. This person – Nicola Mancino, who had formerly served as a Minister and as Head of the Judiciary Supreme Council, but was no longer in office at the time of the wiretapping – was reported to be troubled by the ongoing criminal investigations on the corruption scandals of 1992.[106]

Italian criminal procedural law states that the Prosecutor's Office must present all collected wiretapping records during an investigation (and previously authorized by a court) in a hearing in which all the parties are given the chance to listen to the whole set of wiretapped conversations. After that hearing, all the parties' allegations considered, the court admits or rejects each of the records in view of the actual trial.

In the case concerned, prior to such a hearing, information was leaked to the press regarding President Napolitano accidentally featuring in some records; Prosecutors were prompted to confirm these rumors and specified that in their view no penal relevance attached to those conversations. Nevertheless, they asserted that they were bound by a legislative duty to present all the records at the hearing and let the judge decide, before all the parties concerned.

Governo" alla formazione del Governo Conte (Napoli: Editoriale Scientifica, 2018), 19; as for the legislation on immigration (*Decreti Sicurezza*) and the President Mattarella's comments prior to the promulgation, see G. Azzariti, 'I problemi di costituzionalità dei decreti sicurezza e gli interventi del Presidente della Repubblica', 3 *Diritto pubblico*, 639-650, 646 (2019)

[104] Corte costituzionale 15 January 2013 no 1, *Cassazione penale*, 1319 (2013).

[105] 'Conflitto di attribuzioni tra poteri dello Stato' (Art 37, legge 11 March 1953 no 87).

[106] In February 1992 huge corruption scandals concerning politicians and members of the Italian financial elite arose all over Italy (famously dubbed *Tangentopoli*, the Town of Bribes); simultaneously, the conviction of Mafia bosses in the Maxi-Trial (30 January) and the utter political instability caused by the scandals triggered a gory reaction by the Mafia, which culminated in the murder of the two key Anti-Mafia Prosecutors (Giovanni Falcone and Paolo Borsellino, killed by explosives respectively on 23 May and 19 July). The current investigations in the framework of the 'State-Mafia Negotiations Trial' reveal that a pact among Mafia bosses and politicians was sealed in 1993 to halt these murders, and that this pact has oriented Italian politics from 1994 (I Berlusconi Cabinet) onward. In 1992 Nicola Mancino was Minister of Home Affairs; he was investigated and went acquitted in the First Instance Trial – unlike many others who were convicted, both politicians and Mafia bosses. Documents and reports in https://tinyurl.com/3f7jdndy (last visited 30 June 2021); comments on the judgment of First Instance (20 April 2018) in G. Amarelli, 'La sentenza sulla Trattativa Stato-Mafia: per il Tribunale di Palermo tutti i protagonisti sono responsabili del delitto di minaccia a un corpo politico dello Stato di cui all'art. 338 c.p.' *Diritto penale contemporaneo*, 7-8, *passim* (2018).

The President instead wanted those records destroyed beforehand and challenged the Prosecutors' action before the Court.[107]

Some commentators pointed out that the Constitutional Court was forced into an awkward position. Given both the importance of the case and the heated political context, to rule against the President would have urged him to resign, which was felt by majoritarian political forces to be nothing less than calamitous. Therefore, the President's institutional authority added to the ardent political support he enjoyed, which gave the Court no other choice than to endorse his position. President Napolitano knew he was making the Court an offer that could not be refused, and he did so – as a constitutional scholar commented – with the aim of sealing an institutional alliance that would have overwhelmed the Prosecutors, yet at the expense of the Court's credibility.[108]

The Court indeed engaged in an unprecedented, fully-fledged account of the presidential function. According to this decision, presidential powers are not exhaustively listed in the Constitution, but rest on a 'net of relationships' (*rete di raccordi*) that the President must be left free to entertain as she holds appropriate, in view of 'the unity of the legal order'. Thus, *as instrumental* to the free unfolding of such a net, all records in which the voice of the President is audible cannot be brought as 'evidence' before any court but must be destroyed prior to any adversary check; otherwise, her freedom to exercise her constitutional tasks would be undermined.[109]

In order to attain the desired outcome, the Court proceeded to an audacious 'constitutionally consistent' reading of criminal procedural rules meant to be derogatory in nature (thus, hardly interpretable extensively) and backed it by analogies with other norms limiting the judicial disclosure of records – namely, with the restrictions to disclosure applying to conversations between a doctor and his patient, a lawyer and his client, a Catholic confessor

[107] Opposite reactions from constitutional scholars featured in the press: G. Azzariti, 'Un conflitto senza regole' *Il Manifesto*, 17 July 2012; G. Zagrebelsky, 'Napolitano, la Consulta e quel silenzio sulla Costituzione' *La Repubblica*, 17 July 2012; M. Ainis, 'Le istituzioni e le persone' *Il Corriere della sera*, 17 July 2012; U. De Siervo, 'Ristabilire il senso del limite' *La Stampa*, 17 July 2012; F.P. Casavola, 'La tutela del Colle l'unico obiettivo' *Il Mattino*, 17 July 2012; A. Manzella, 'Conflitto di poteri: l'equilibrio smarrito' *La Repubblica*, 18 July 2012.

[108] See G. Zagrebelsky, n 106 above. Allegations were serious: the intercepted person had been implicated in the investigations on the negotiations allegedly entertained by Mafia members and State's officials after the bombs that devastated Italy in 1992-1993, and from the wiretappings he seemingly asked to speak with the President precisely about that issue. In addition, President Napolitano resorted to the Court few days before the 20th anniversary of the bombs killing the then Palermo Prosecutor Paolo Borsellino (19 July 1992) which obviously added to the momentum of the events and occasioned heated comments during the annual memorial: a banner was exhibited at the ceremony with the wording '*1992-2012: Romanzo Quirinale*' (*Romanzo criminale*, 'A Criminal Novel' is a popular movie and TV series based on an Italian criminal gang; the Roman hill '*Quirinale*' is the President's residence). See G. Pipitone, 'Mancino-Napolitano, un anno di *Romanzo Quirinale*: "Distruggere le intercettazioni"' *Il Fatto Quotidiano*, 18 April 2013.

[109] Corte costituzionale, n 103 above.

and the confessed person – which clearly aligned with the *charismatic*, far more than rational, legitimacy, that was being recognised as belonging to the President.

Many commentators warned that such arguments promised to go much further than that specific, yet highly controversial, case, for they could have been deployed in support of a new interpretation of the President's constitutional powers, which entailed two consequences. First, leaving the frontier of the President's mandate at her own disposal, as it was only defined in the teleological perspective of national unity. Second, disrupting the balance between the two channels supplying legitimacy to the President's action, above referred to as 'constitutional magistracy' and 'representative of national unity'.[110] Both consequences emphasized the emotional link between the President and the Italian people; and both resulted in increased presidential powers resting on a weakened reason-based argumentative support, to the detriment of other constitutional organs.[111]

Hence, material for constitutional arguments emerged from those troublesome days to endow the presidential figure with some 'stronger', politically relevant, traits. Another warning sign of an ongoing, yet controversial, mutation of the presidential role in the Italian constitutional architecture can be detected in the behaviour of President Mattarella during the current Covid-19 crisis. Although in a more discreet style than his predecessor's, President Mattarella has enhanced his emotional connection with the Italian people while making his voice clear and loud during the negotiations occurring at the EU level to decide on how to tackle such a crisis.

On 12 March 2020, in her monthly press conference, the President of the European Central Bank (ECB) Christine Lagarde answered a question on the ECB's role in the Covid-19 emergency by pointing out that 'the Bank is not to close spreads'.[112] Her declaration appeared to repudiate the ECB's robust stance as the bulwark against markets' speculations on the Euro's fall that former President Mario Draghi announced in his celebrated 'whatever it takes' speech.[113]

[110] See the relevant literature cited in G. Vosa, 'Percorsi di legittimazione del potere. La figura del Presidente della Repubblica nei primi mesi del bi-settennato di Napolitano, rileggendo C. cost., 1/2013' 1 *Rivista AIC*, 1-24 (2014).

[111] G. Scaccia, 'Il ruolo del Presidente della Repubblica dopo la sentenza della Corte costituzionale 1 del 2013', in L. Violini ed, *Il ruolo del Capo dello Stato nella giurisprudenza costituzionale*, 'Associazione Gruppo di Pisa' Annual Report (Napoli: Editoriale Scientifica, 2015), 39-71; M. Luciani, 'La gabbia del Presidente' n 44 above; A. Pace, 'Intercettazioni telefoniche fortuite e menomazione delle attribuzioni presidenziali' 6 *Giurisprudenza costituzionale*, 1267 (2013).

[112] Available at https://tinyurl.com/8vzwpxw5 (last visited 30 June 2021) the *verbatim* of the press conference. See comments: 'Italy furious at ECB's Lagarde 'not here to close spreads' comment', available at *Reuters.com*, 13 March 2020; 'ECB's plan to support eurozone banks is underwhelming' *The Guardian*, 13 March 2020.

[113] 'But there is another message I want to tell you. Within our mandate – within our mandate – the ECB is ready to do whatever it takes to preserve the Euro. And believe me, it will be enough'.

This caused turbulence to the stock market, to the detriment of the Italian financial position. Then, an official note from the *Quirinale* came rapidly on *Twitter* and soon became a trending topic: 'Italy is in a hard condition, while the Italian experience in fighting the virus will probably be of help for all countries. It is right to expect, in the common interest, that initiatives of solidarity be made, rather than moves that can hamper our action'.[114]

The quick reaction on social *media*, the peremptory expressions, the lack of formality and the sharp tones bordering on rudeness are highly exceptional to President Mattarella's usual style of communication. That message implicitly fostered an impression of political leadership vested in the Presidency as a representative of national unity. Indeed, that message aimed to defend the country as a whole in a skirmish of bargaining which will predictably be both difficult and delicate. Yet, what specific position Italy is to maintain is far from uncontroversial: political fractures are emerging as to what measures are to be taken and which of the recovery instruments discussed at the EU-Eurogroup[115] level is to be accepted or rejected.[116] Meanwhile, President Mattarella has gone silent; however, such an exposure of the President's role as a sort of gatekeeper between the domestic and the supranational plane[117] has 'fortuitously' aligned with a bizarre event resulting in an emotion-based boost of his personal figure. On 28 March, the President featured in a video to address the country during the emergency. The footage was broadcast in prime time, but unexpectedly looked unedited: it included failed attempts to record the speech where the President is mumbling, coughing, halting repeatedly, trying and stopping again, showing distress for the difficult circumstances, and joking with the cameraman: 'I do not go to the hairdresser either' (meaning: due to the lockdown) before delivering the message.[118]

Speech at the *Global Investment Conference*, London, 26 July 2012, available at https://tinyurl.com/hb3m4njx (last visited 30 June 2021).

[114] See the tweet at https://tinyurl.com/374eaepk (last visited 30 June 2021).

[115] On the relations between the 'Eurogroup' as a key institutional premise among the multiple intergovernmental articulations of the EU institutional architecture, P. Craig, 'The Eurogroup, power and accountability' 23(3-4) *European Law Journal*, 234-249 (2017).

[116] See the positions of the Italian MEPs at the European Parliament's Plenary Sitting (Brussels, 17 April 2020) in the vote on the Commission's proposal of a CoVid-19 Economic Package (2 April 2020, consisting of eleven different legislative proposals. The mutualization through common bonds (dubbed 'Coronabonds') of the debt for healthcare expenditures met with a 'no' from national opposition parties *Lega* and *Forza Italia* (Berlusconi's) while being the key negotiation target for the Italian Cabinet (as repeatedly pointed out by Prime Minister Giuseppe Conte). Within the majority's coalition, one party (*Movimento Cinque Stelle*) voted against the proposal referring to the use of ESM – European Stability Mechanism, perceived in their perspective as a 'Greece-like treatment' – and abstained on the overall package while the Democrats voted 'in favour of the ESM. All the documents are available at https://tinyurl.com/3ade4889 (last visited 30 June 2021).

[117] M. Ferrara, 'La Presidenza Mattarella tra politica estera e garanzia interordinamentale' at n 82 above, 390.

[118] See the video at https://tinyurl.com/h6kfyfn7 (last visited 30 June 2021).

This was claimed to be a mistake on the side of the operators, but it happened for the first time in the Italian Republic's history, and it was a mistake that admittedly contributed to strengthening an emotion-based channel for the President's direct communication with the Italian people.[119]

As a political crisis exploded in January 2021, President Mattarella went public twice. In the first occasion (29 January) he announced that 'a chance for a new Cabinet to be supported by the same political forces existed and was to be duly verified'; he conferred on the President of the Chamber of Deputies, Roberto Fico, an explorative mandate.[120] In the second occasion, speaking in a prime-time broadcast after a tense day, he confirmed that 'such a verification had delivered a negative result' and pointed to 'two possible solutions'. First: 'to immediately form a new Cabinet that is adequate to confront the serious ongoing emergencies. Second: 'elections'. The latter option 'must be carefully considered, as the polls are an exercise of democracy'; yet he declared he had a 'duty to highlight that they would coincide with a crucial time for Italy'. Then, he listed all the reasons why (he held) general elections were incompatible with 'a decisive development' of the vaccine campaign, as well as with a sound management of the EU pandemic funds amidst the mounting social emergencies, and so on. 'We cannot afford to waste this opportunity, for our future', he concluded. Finally, he addressed 'all the political forces' to support 'a high profile Cabinet' to confront these urgencies.

Soon afterwards, the Quirinale's Clerk officially mentioned the name of Mario Draghi as a candidate President of the Council of Ministers in a press conference.[121]

Among the commentators, some have highlighted that Mattarella's choice has been indeed of high political significance.[122] Others, too, have pointed out that this 'supermajoritarian Cabinet' represents a kind of constitutional anomaly.[123] It may also be highlighted that – apart from the obviously remarkable pressure he put on all the political forces to accept Draghi as the leader of a new Cabinet, which they did – he has silently excluded the option of a centre-right Cabinet that would have brought into power Eurosceptic forces;[124] anyhow, he has

[119] See M.C. Antonucci, 'Il barbiere di Mattarella, ovvero l'errore che lega istituzioni e cittadini (W il Presidente!)' available at *Formiche.net*, 28 March 2020, and the newspapers of those days: 'Mattarella e il barbiere: la normalità che ci aiuta. Quando la prima carica del Paese ha i tuoi stessi problemi, la distanza si abbatte' *Corriere del Mezzogiorno*, 29 March 2020.

[120] See 'Mattarella dà un incarico esplorativo a Fico: "Verificare l'esistenza dell'attuale maggioranza di governo"' Il Fatto Quotidiano, 29 January 2021.

[121] See the video at https://tinyurl.com/6dj9svtc (last visited 30 June 2021).

[122] A. D'Andrea, 'Decisioni neutrali che neutrali non sono. L'investitura del nuovo Governo' *LaCostituzione.info*, 14 March 2021

[123] S. Curreri, ' "Super-maggioranze" e "super-opposizioni" ' *LaCostituzione.info*, 14 February 2021.

[124] He could have conferred an equally short explorative mandate to the President of the Senate (Maria Elisabetta Alberti Casellati, Forza Italia) but such a possibility has apparently been excluded since the first round of informal meetings with the political leaders.

(again) furthered a top-down model of legitimacy that has been accounted for as 'risky'.[125]

In sum, President Mattarella walks a path that was inaugurated a few years ago; yet, like his predecessor, he is well aware of the liaison role that he has inherited, and of the increasing difficulties that it entails.[126] His task is coming to be the most delicate, perhaps even more than his predecessor's, given the multiple ongoing crises and the incumbent economic downturn: should he find himself to act as a political leader,[127] he would deprive of representation certain political interests, lawful nevertheless, simply because they apparently run contrary to his (and/or his party's) political positions.[128] Whether this is consistent with the Constitution – so long as the current Constitution stays in force – is a question that could find several grounds for a negative answer.

VI. Conclusions. Domestic Rigidity and Supranational Openness: What Foundations for Representative Sovereignty?

In the aftermath of the Second World War, according to common wisdom, the sense of pain as a result of the catastrophes that perturbed the planet urged States to agree on new arrangements designed to avoid the dangers of a totalitarian regime supported by aggressive nationalist States.[129] The key was to repudiate nationalist autarchy and to endow sovereignty with solid representative support, based on the centrality of the human person. In this vein, respect for human dignity inspired separation of powers and protection of rights[130]– both instrumental to one another – in drafting rigid constitutions. Likewise, openness towards a supranational public space fostered globalisation and triggered European integration as instrumental to economic liberties; so that a multi-level field of overlapping legal orders was created as 'a space of

[125] See G. Zagrebelsky, 'Il governo Draghi e tutti i rischi di questa 'democrazia dall'alto' at https://tinyurl.com/et2ep7e3 (last visited 30 June 2021).

[126] See the speech delivered for the 40th anniversary of the death of Giovanni Gronchi (3rd President of the Italian Republic) 'Il Capo dello Stato è la voce della Costituzione contro ogni smarrimento verso i valori della Carta' *Il Fatto Quotidiano* (18 October 2018).

[127] Some of his actions might be interpreted in this manner since the beginning of his mandate: see the invitation for a private colloquium delivered to Roberto Battiston (physicist, removed a few days prior from the Chairmanship of the Italian Space Agency by the new Cabinet) to 'speak about the autonomy of science' as reported in *La Repubblica* (9 November 2018).

[128] Sergio Mattarella has served as a member of the Democrat Party. Yet in some media, perhaps in unrequested outbursts of complicity, there are frequent contributions construing his public profile as a political leader: see G. Genna, 'Sergio Mattarella, ritratto di un presidente *pop*' *L'Espresso* (1 November 2018).

[129] See C. Möllers, ' "We are (afraid of) the People": Constituent Power in German Constitutionalism', in M. Loughlin and N. Walker eds, *The Paradox of Constitutionalism. Constituent Power and Constitutional Form* (Oxford: Oxford University Press, 2007 – repr 2012), 87-105.

[130] J. Habermas, *The Concept* n 19 above.

liberty'.[131]

Contemporary constitutionalism has rested on the idea that the constitutional structures derived from the positive, circular relationship of these concepts would work well to accommodate social and political pluralism at both national and supranational levels. Whoever doubted the functioning of such structures found solace in the idea of a common cultural *humus* – or at least a common constitutional culture[132]– which provided solidarity among peoples, as well as cooperation among institutions, with a more tangible background in cases of urgency.

Faced with the tough reality of the economic-financial crisis and of the increasing differences flourishing in a multi-faceted society, these hopes have proved largely optimistic.[133] Diverging interests generate tensions from within and outside States that constitutional devices struggle to modulate. Exacerbated political conflicts have led the Italian Head of State to curtail some interests from the spectrum of the State's unity, leaving them with no representative coverage. However, noticeably, those excluded interests are not ruled out because of their illegality. They are perfectly legitimate, but simply 'cannot be afforded' by the Italian State, which exposes obvious, painful asymmetries in European membership, and presents a challenge to the veracity of the overall narrative that aimed to cement the foundations of post-World War II constitutionalism.[134]

In fact, the events recounted in this article display an emerging fracture between domestic rigidity and supranational openness in the protection of socio-political pluralism within the European common space.

There are two possible alternatives: either A) the Constitution forbids the modified interpretation the President pursues as regards his own constitutional role; or B) it does not.[135] A proper answer to the question would require an in-depth analysis of the limits to the constitutional mutations triggered by supranational openness that would exceed the bounds of this article.[136]

[131] F. Álvarez-Ossorio Micheo, 'Europa como espacio integrado de libertad' 3(5) *Araucaria – Revista Iberoamericana de Filosofía, Política, Humanidades y Relaciones Internacionales*, 93-122 (2001).

[132] P. Häberle, *Verfassungslehre als Kulturwissenschaft* (Berlin: Duncker & Humblot, 1998), 320.

[133] See K. Jaklić, *Constitutional Pluralism in the EU* (Oxford: Oxford University Press, 2013), 102, 217; among others, G. della Cananea, 'Is Multilevel Constitutionalism really "Multilevel"?' 70 *Zeitschrift für ausländisches öffentliches Recht und Völkerrecht*, 283-317 (2010).

[134] J. Shaw, "Shunning' and 'seeking' membership: Rethinking citizenship regimes in the European constitutional space' 8(3) *Global Constitutionalism*, 425-469 (2018); A.J. Menéndez and E.D.H. Olsen, *Challenging European Citizenship. Ideas and Realities in Context* (London: Palgrave-McMillan, 2020), 59.

[135] *Ex plurimis*, see B. Ackerman, *Revolutionary Constitutions: Charismatic Leadership and the Rule of Law* (Boston: Harvard University Press, 2019), 27 (on Italy: 131) and R. Albert, *Constitutional Amendments: Making, Breaking, and Changing Constitutions* (Oxford: Oxford University Press, 2019), 18, 61.

[136] A. Morrone, 'I mutamenti costituzionali derivanti dall'integrazione europea' 20

However, it can be argued that domestic rigidity and supranational openness find themselves in conflict or lead, if taken together, to the suppression of political pluralism on either the national or the supranational plane – precisely that political pluralism that they aimed to protect.

If A) is correct, either constitutional rigidity is safe at the expense of supranational openness – because President Mattarella, in order not to violate the Constitution, must refrain from his position and repudiate the supranational ties that Italy has subscribed to – or, for the sake of supranational openness, the national constitution must become flexible, ie allowing for a presidential breach of the constitutional architecture. In the first case, political pluralism is preserved at the national level but repudiated on a supranational plane; in the second, it might well be preserved at the supranational level but at the expense of legitimate positions on the national level.

If B) is correct, the presidential action to prevent a supreme violation of the Constitution can be regarded as the optimal solution to reconcile domestic rigidity with supranational openness; but legitimate political interests are left with no representative coverage. Thus, the combination of the two leads to the suppression of political pluralism, rather than to the protection thereof.

The constitutional mutation that is being triggered seems to overturn the positive relation between political and substantive rights; political rights risk being denied in order for substantive ones to be granted.[137] In fact, President Mattarella links to 'concrete sovereignty' and to the 'protection of savings' his rejection of Savona's appointment as a Ministry of Economy. Hence, in a certain point of the European legal space, along the curvature of the Italian constitution, some rights do not find their source in the law-based functioning of representative institutions, but elsewhere. This new source has an eerie characteristic: it requires obedience to an inscrutable authority whose legitimacy is justified on emotional arguments rather than on rational discourse.[138] In this perspective, the departure from the moorings of contemporary constitutionalism, including self-determination as a supreme expression of human dignity, could hardly be more evident: [139] to put it bluntly, the source of authority is getting so far and unfathomable than it could well go transcendent without much difference.

As a consequence, substantive rights would not be *granted* – in the sense of being *recognised* as belonging to the citizenry as a community of human persons – but, rather, genuinely conceded, *octroyés*, *otorgados*, *elargiti* by a

Federalismi.it, 24 October 2018, 1-27.

[137] M. Wilkinson, *Authoritarian Liberalism and the Transformation of Modern Europe* (Oxford: Oxford University Press, 2021), 95, 147.

[138] J. Habermas, *Die Einbeziehung des Anderen. Studien zur Politischen Theorie* (Frankfürt: Suhrkamp Verlag, 1996), 65.

[139] See M. Wilkinson, 'Authoritarian Liberalism: The Conjuncture behind the Crisis' *LSE-WP* 5/2018, 1-21.

virtually unquestionable, superior law-maker, in the like of pre-modern absolute monarchies[140] – or, else, in the XIX century's *Allgemeine Staatslehre* fashion.[141] This, to be sure, would tie the so-called *multilevel* protection of rights to certain political and constitutional theories whose commonality with the totalitarianisms of the *Short Twentieth Century*[142] has been well documented in a book published as recently as 2003.[143] In that book, the editors Christian Joerges and Navraj Singh Ghaleigh wonder whether Europe's constitutional legacy includes certain dark aspects that were going unnoticed. Around two decades later, this sounds like a warning that ought not be neglected.

[140] See M. Dawson and F. de Witte, 'Self-determination in the constitutional future of the EU' 21(3) *European Law Journal*, 371-383 (2015).

[141] For comparison, see C.F. von Gerber, *Grundzüge eines Systems des deutschen Staatsrechts* (3rd ed, Leipzig: Tauchtnitz, 1880), 44-75, 190-243 and G. Jellinek, *System der subjektiven öffentlichen Rechte* (Tübingen: Mohr Siebeck, 1892), 39-50.

[142] The quote is from E. Hobsbawm, *The Age of Extremes: The Short Twentieth Century, 1914–1991* (New York: Random House, 1994).

[143] C. Joerges and N. Singh Ghaleigh, *Darker Legacies of Law in Europe: The Shadow of National Socialism and Fascism over Europe and Its Legal Traditions* (Oxford: Hart, 2003).

Online Unfair Commercial Practices: A European Overview

Mariacristina Zarro*

Abstract

The supranational economic paradigm considers the weak user a tool for the realization of the market: through his choices (contracts) he rewards companies that contribute to offering products at the best quality-price ratio, thus playing a central and propulsive role in the European common market. To do this, however, he needs correct information and conduct that today we try to guarantee through the integrated regulation, always of European derivation, relating to information obligations in contracts with consumers, that on misleading and comparative advertising, and especially that which governs the phenomenon of unfair commercial practices. The fight against the latter becomes indispensable for the purpose of creating the internal market. Yet, to date there is still no regulation concerning the fact that conflicts with these practices are carried out online. The numerous cases brought to the attention of the antitrust authorities in recent years require to analyze these practices, even when these are subtly perpetrated online. It is necessary to investigate whether the discipline, including the more recent European one, is able to respond to the new way of being of the markets and whether the current binary system of public and private enforcement is suitable to deal with the fight against these practices that are harmful to both the consumer and the internal market.

I. Introduction

A closer scrutiny to the reforms to the Treaty establishing the European Economic Community reveals that the original mercantile prospective has been abandoned in favour of a more attentive and respectful approach towards the protection of human rights.[1] It has been

* Assistant Professor of Private Law, Marche Polytechnic University.

[1] See P. Perlingieri, *La tutela del consumatore nella Costituzione e nel Trattato di Amsterdam*, in P. Perlingieri and E. Caterini eds, *Il diritto dei consumi* (Napoli: Edizioni Scientifiche Italiane, 2004) I, 12, emphasising the role of Treaty of Amsterdam in spreading the culture of individual rights and confirming the interest of the European in themes which are not exclusively economic. Such a trend is also visible from the amendments to the Treaty on European Union and, in particular, by the replacement of the existing seventh recital of the Maastricht Treaty with the new text: 'Determined to promote economic and social progress for their peoples, taking into account the principle of sustainable development and within the context of the accomplishment of the internal market and of reinforced cohesion and environmental protection, and to implement policies ensuring that advances in economic integration are accompanied by parallel progress in other fields'. Also very important is the amendment of Art F, para 1, with the following provisions: 'The Union is founded on the principles of liberty, democracy, respect for human rights and

> a revenge of our Constitutional lawmaker (and maybe not only of ours) against the founding fathers of the European Union. Once it has been made clear that the market is an instrument for protecting human dignity and the rights associated to it (rather than a value in itself), the original mercantile prospective has been abandoned.[2]

Nevertheless, competition and consumer protection, which are essential instruments for the internal European market, remain core objectives of the European lawmaker.[3] Their paramount importance is stated in in a great number of Directives and Regulations, as well as in Art 38 of the EU Charter of Fundamental Rights, providing that 'Union policies shall ensure a high level of consumer protection'.[4]

fundamental freedoms, and the rule of law, principles which are common to the Member State'. On such themes, see A. Tizzano, *Il Trattato di Amsterdam* (Padova: CEDAM, 1998), 37; S. Negri, 'La tutela dei diritti fondamentali nell'ordinamento comunitario alla luce del Trattato di Amsterdam' *Diritto dell'Unione europea,* 782, 773-793 (1997); A. Adinolfi, 'Le innovazioni previste dal Trattato di Amsterdam in tema di politica sociale' *Diritto dell'Unione europea*, 563-569 (1998).

[2] P. Perlingieri, *Relazione conclusiva*, in P. Perlingieri and L. Ruggeri eds, *Diritto privato comunitario* (Napoli: Edizioni Scientifiche Italiane, 2008), 401.

[3] See CaseC-67/96, *Albany International BV* v *Stichting Bedrijfspensioenfonds Textielindustrie*, Judgment of 21 September 1999, available at www.eurlex.europa.eu, whereby it has been highlighted that the 'task of the Community is to promote throughout the Community a harmonious and balanced development of economic activities and a high level of employment and of social protection'. See also S. Giubboni, 'Da Roma a Nizza. Libertà economiche e diritti sociali fondamentali nell'Unione europea' *Quaderni di diritto del lavoro e delle relazioni industriali*, 9-35 (2004); Id, *Diritti sociali e mercato* (Bologna: Il Mulino, 2003); J. Shaw, *Social law and policy in an evolving European Union* (Oxford: Hart Publishing, 2000).

[4] The first ones in chronological order are European Parliament and Council Directive Directive 84/450/EEC of 10 September 1984 relating to the approximation of the laws, regulations and administrative provisions of the Member States concerning misleading advertising [1984] OJ L250/17 and European Parliament and Council Directive 85/577/EEC of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council [2011] OJ L304/64 to protect the consumer in respect of contracts negotiated away from business premises. Then there are European Parliament and Council Directive 90/314/EEC of 13 June 1990 on package travel, package holidays and package tours [1990] OJ L158/59; European Parliament and Council Directive 92/59/EEC of 29 June 1992 on general product safety [1992] OJ L228/24; European Parliament and Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts [1993] OJ L95/29; European Parliament and Council Directive 93/22/EEC of 10 May 1993 on investment services in the securities field [1993] OJ L141/27; European Parliament and Council Directive 94/47/EC of 26 October 1994 on the protection of purchasers in respect of certain aspects of contracts relating to the purchase of the right to use immovable properties on a timeshare basis OJ L 280/83; European Parliament and Council Directive 95/46/EC of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31; European Parliament and Council Directive 97/7/EC of 20 May 1997 on the protection of consumers in respect of distance contracts [1997] OJ L144/19; European Parliament and Council Directive 1999/93/EC of 13 December 1999 on a Community framework for electronic signatures [2000] OJ L13/12; European Parliament and Council Directive 2000/31/EC of 8 June 2000 on certain legal aspects of information society services, in particular

However, despite such a great emphasis, dissatisfaction as to the current status of consumer protection and the effectiveness of the remedies available to that purpose emerges from many sources. This is the case, *inter alia*, of the *Communication from the Commission*, entitled *Ten Years of Antitrust Enforcement under Regulation 1/2003: Achievements and Future Perspectives*, concerning the effectiveness of decentralised enforcement of competition rules by National Competition Authorities.

Pursuant to the study conducted by the European Commission, ten years after the introduction of Council Regulation (EC) No 1/2003, there is still room for a significant improvement in the effectiveness of the enforcement of competition rules. The 2018 Work Program holds that 'the success of the internal market ultimately depends on trust. This trust can easily be lost if consumers feel that remedies are not available in cases of harm'. Accordingly, the Commission considers enhancement of judicial enforcement and out-of-court redress of consumer rights, and a more effective action by national consumer authorities, as crucial elements of its agenda.[5]

electronic commerce, in the Internal Market [2000] OJ L178/1; European Parliament and Council Directive 2001/95/EC of 3 December 2001 on general product safety [2002] OJ L11/4; European Parliament and Council Directive 2005/29/EC of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market [2005] OJ L149/22; European Parliament and Council Directive 2008/48/EC of 23 April 2008 on credit agreements for consumers, striking to find the right balance between a high level of consumer protection and the competitiveness of businesses[2008] OJ L133/66; European Parliament and Council Directive 2015/2302/EU of 25 November 2015 on package travel and linked travel arrangements, amending Regulation (EC) No 2006/2004 and Directive 2011/83/EU of the European Parliament and of the Council and repealing Council Directive 90/314/EEC [2015] OJ L326/1; European Parliament and Council Directive 2008/122/EC of 14 January 2009 on the protection of consumers in respect of certain aspects of timeshare, long-term holiday product, resale and exchange contracts[2009] OJ L33/10; European Parliament and Council Regulation 261/2004/EC of 11 February 2004 establishing common rules on compensation and assistance to passengers in the event of denied boarding and of cancellation or long delay of flights, and repealing Regulation (EEC) No 295/91[2004] OJ L146/1 of 14 January 2009 on a Code of Conduct for computerised reservation systems and repealing Council Regulation (EEC) No 2299/89 [2009] OJ L35/47introducing a Code of Conduct for computerised reservation systems with common criteria and rules for the fixing of the price within the internal market; European Parliament and Council Directive 2012/27/EU of 25 October 2012 on energy efficiency, amending Directives 2009/125/EC and 2010/30/EU and repealing Directives 2004/8/EC and 2006/32/EC [2012] OJ L315/1; European Parliament and Council Directive 2013/11/EU of 21 May 2013 on alternative dispute resolution for consumer disputes and amending Regulation (EC) No 2006/2004 and Directive 2009/22/EC [2013] OJ L165/63; European Parliament and Council Regulation 2017/1369/EU of 4 July 2017 setting a framework for energy labelling and repealing Directive 2010/30/EU [2017] OJ L198/1.

[5] The same has already been observed in the past, when the growing trust in ADR Procedures spurred the enactment of European Parliament and Council Directive 2013/11/EU of 21 May 2013 on alternative dispute resolution for consumer disputes and amending Regulation n. 2006/2004/EC and Directive 2009/22/EC, as well as the related Regulation No 524/2013/EU of the European Parliament and of the Council, adopted on the same date, concerning online dispute resolution for consumer disputes and amending Regulation No 2006/2004/EC and Directive 2009/22/EC (Directive on consumer ADR)[2013] OJ L165/63.

Directive 2019/2161 EU reached the same conclusion,[6] declaring that

> consumer protection law should be applied effectively throughout the Union. Yet, (…) the Commission in 2016 and 2017 (…) concluded that the effectiveness of Union consumer protection law is compromised by a lack of awareness among both traders and consumers and that existing means of redress could be taken advantage of more often.[7]

The failure to achieve the objectives set by the European Commission and the European Parliament may be ascribed to many causes, one of the most relevant being the fast technological development occurred in the last decade, which has led to significant changes in the structure of the markets. In fact, an increased use of social media, along with improvements of information processing systems, reshaped in the last years the very nature of economic transactions. In this context, the fast rate of technological development has made it increasingly clear how statutory rules are often unable to keep up with the speed of an internet-based society, requiring thus, with the utmost urgency, a regulatory reform aimed at enhancing the effectiveness of consumer protection and promoting a fair competition.

In accordance with the idea that the enforcement of consumer regulation is a key element for a well-functioning market, the correctness of the decision-making process has been considered crucial for the consumer to be able to make a rational choice, so to improve the efficiency in the supply chain of products and services. Consequently, as far as an opaque and uncompetitive market does not encourage consumers to make reasonable decisions, accuracy of information and fairness in commercial practices have become crucial for the safeguarding of the proper functioning of the market.[8]

These goals are pursued through a detailed discipline of information requirements, tight rules concerning comparative and misleading information, and the prohibition of unfair commercial practices on the market.[9] Even though

[6] European Parliament and Council Directive 2019/2161/EU of 27 November 2019 amending Council Directive 93/13/EEC and Directives 98/6/EC, 2005/29/EC and 2011/83/EU of the European Parliament and of the Council as regards the better enforcement and modernisation of Union consumer protection rules [2019] OJ L 328/7. See below, para 3.

[7] Also the Report Procedural Protection of Consumers, published on 26th January 2018 and available at europe.eu, as well as the Study for the Fitness Check of EU consumer and marketing Law – Final report Part 1: Main report, European Commission (REFIT), published on 29th May 2017, available at europa.eu, show some inefficiencies in consumer protection, especially in terms of access to justice.

[8] See C. Biasior, *Pratiche commerciali scorrette*, in F. Casucci and G.A. Benacchio eds, *Temi e istituti di diritto privato dell'Unione europea* (Torino: Giappichelli, 2017), 187.

[9] National European legislation set a detailed discipline about misleading and comparative advertisement, then after European Parlaiment and Council Directive 2005/29/EC, provided for rules against unfair commercial practices, which have been transposed into Italian law with decreto legislativo 2 August 2007 no 146 as 'unfaithful commercial practices'. A. Vanzetti and V. Di Cataldo, *Manuale di diritto industriale* (Milano: Giuffrè, 8th ed, 2018), 132.

the fight against such abusive behaviours is a central element for the improvement of the internal market, strikingly, there is no legislation covering unfair commercial practices on the internet.

The first part of this article analyses the most relevant decisions rendered by National Competition Authorities concerning unfair commercial practices[10] on the internet.[11] The second part is dedicated to the most recent rules enacted by the European Union, with the aim of understanding the possible effects deriving therefrom and improving their transposition into the domestic legal system. More specifically, after some introductory observations, paragraphs 3 and 4 explore cases concerning unfair commercial practices on social media (*Facebook*, *WhatsApp* e *TripAdvisor*); paragraph 5 investigates the recent Directive 2019/2161 UE. The last section conveys conclusive remarks on the expected effectiveness of consumer protection from unfair commercial practices on the internet.

Before digging into the above-mentioned issues, it is worth reminding that the vast majority of business to consumer (B2C) transactions take place on the internet[12] and that the current legal framework on the enforcement of consumer protection is given by Art 18 and et seq of the Consumer's Code. Furthermore, domestic rules on these matters derive from the European legislation, with Directive 2005/29 EC as one of the most important sources. Such a Directive has been transposed into Italian law by the decreto legislativo 2 August 2007 no 145.[13]

[10] See, *ex multis*, M. Clarich, 'Le competenze delle autorità indipendenti in materia di pratiche commerciali scorrette' *Giurisprudenza commerciale*, I, 688-705 (2010); M.R. Raspanti, 'Il nuovo assetto di competenze in materia di pratiche commerciali scorrette: "reddite quae sunt auctoritatis auctoritati" ' *Concorrenza e mercato*, 155-182 (2015); G. Barozzi Reggiani, 'Pratiche commerciali scorrette, regolazione e affidamento delle imprese' *Diritto amministrativo*, 683-718 (2016); A. Fachechi, 'Gli orientamenti dell'Autorità garante della concorrenza e del mercato in materia di pratiche commerciali scorrette (anni 2014-2015)' *Concorrenza e mercato*, 497-523 (2016).

[11] European Parliament and Council Directive 2005/29/EC of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive') [2005] OJ L149/22. The Commission Staff working document guidance on the implementation/application of European Parliament and Council Directive 2005/29/EC on unfair commercial practices, accompanying the document communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – a comprehensive approach to stimulating cross-border e-commerce for Europe's citizens and businesses – clarifies that 'On-line platforms work according to many different business models: their behaviours range from merely allowing users to look for information supplied by third parties to facilitating, often against remuneration, contractual transactions between third party traders and consumers or advertising and selling, in their own name, different kinds of products and services including digital content'.

[12] See data on e-commerce by the *Politecnico di Milano*, available at *mark-up.it*.

[13] See, ex multis, E. Minervini, Codice del consumo e pratiche commerciali sleali, in E. Minervini and Rossi Carleo eds, Le pratiche commerciali sleali. Direttiva comunitaria ed ordinamento italiano (Milano: Giuffrè, 2007), 75; R. Di Raimo, 'Note minime sulle implicazioni sostanziali dell'art. 14 della direttiva 2005/29/CE a margine di una proposta per il suo recepimento'

II. Unfair Trade Practices on the Internet. The Most Relevant Cases Dealt with by the Italian Competition Authority and the Phenomenon of Influencer Marketing

The evasive and ever-changing character of unfair trade practices on the internet reflects on the variety of the case-law. Along with cases concerning the usual exchange of goods and services on the internet,[14] there are much more complex situations,[15] such as those regarding influencer marketing.[16] The latter

Contratto e impresa Europa, 91-101 (2007); A. Gentili, 'Pratiche sleali e tutele legali: dal modello economico alla disciplina giuridica' Rivista di diritto privato, 58, 37-67 (2010); A. Fachechi, Pratiche commerciali scorrette e rimedi negoziali (Napoli: Edizioni Scientifiche Italiane, 2012).

[14] In 2016, the Italian Competition Authority has initiated and then concluded several proceedings regarding RG Group, Mobile Store S.r.l, WM S.r.l.s, Aquila S.r.l, Sami S.r.l.s, due to unfair practices perpetrated online. The national Authority established that these companies used to sell technological products, such as televisions, tablets and smartphone, declaring they were immediately available. Consumer bought products but, soon afterwards, difficulties emerged for the delivery of the good or restitution of the price. All these proceedings ended with an administrative sanction against the investigated companies. See ICA, decision 13 April 2016 no 25975; decision 13 April 2016 no 25976; decision 13 April 2016 no 25977; decision 4 August 2016 no 26159; decision no 26163 dated 4 August 2016. All these decisions are available at *agcom.it*. In the same 2016, the competition watchdog has sanctioned Amazon EU Sàrl and Amazon Services Europe Sàrl, which is in charge of the management of the Amazon marketplace, connecting consumers with third party sellers. the Italian Competition Authority found that the two companies had not provided, or had provided inadequately, relevant information during the purchase phase; in particular: mandatory pre-contract information and information concerning conformity legal guarantees provided for by the Consumer Code. More in particular, Amazon Eu Sàrl did not adequately provide users, before contractual obligations and in an easy and accessible way, with a specific pre-contract document offering information concerning rescission and related terms and exclusions, the existence and conditions of a post-sale customer service, besides a remind on legal guarantees. Moreover, Amazon did not adequately inform consumers about the real identity of their counterparts, so that consumers thought they had concluded a contract with Amazon rather than a third party. Accordingly, decision no 25911, dated 9 March 2016, the Competition Authority: *a*) Amazon EU Sàrl and Amazon Services Europe Sàrl had violated Arts 49 and 51 of the Consumer Code and, for that reason, has ordered to cease such an infringement; *b*) levied sanctions for € 80,000 against Amazon EU Sàrl and € 220.000 against Amazon Services Europe Sàrl; *c)* required both the companies to communicate to the Authority the measures adopted to comply with its decision, within thirty days from the summoning.

[15] Somehow more complicated appears the Trenitalia case. The AGCM defined as an unfair commercial practice Trenitalia's omission of regional cheaper trains from the results of online searches, on automatic vending machines and on its app. Investigations ascertained that the algorithm governing the search engine had misleading effects on consumers, who were not informed of alternative and generally less expensive solutions. Decision no 26700, dated 19 July 2017, branded such a commercial practice in contrast with Arts 20 and 21, para 1, lett *b*), of the Consumer Code, levied a € 5 million fine against Trenitalia and ordered the company to immediately stop its conduct.

[16] AGCM, Press release of 24 July 2017; Id, Press release of 6 August 2018, available at *agcm.it*. 'Bloggers or so-called influencers (ie widely followed social media personalities) support or endorse specific brands through photos, videos or comments posted on blogs, vlogs and social media such as Facebook, Instagram, Twitter, YouTube, Snapchat, Myspace, thus generating an advertising effect. This form communication – initially used only by celebrities – is now becoming more and more common on social networks among a considerable number of users who do not have a particularly high number of followers'. For a definition of influencer marketing, see M. Fiocca, 'Il

has recently attention in scholarly writings and in case-law, both at the national and international level.

Influencer marketing involves public figures with a high number of followers, generally active in the show business or in the fashion industry, endorsing products on social media such ad Facebook, Instagram, Twitter, YouTube, Snapchat or Myspace and inducing consumers to choose some brands, rather than others. Difficulties with domestic and European legislation begin when advertising and endorsement disclosure is missing.[17] Such a phenomenon, originally limited to the most-followed figures, is currently expanding so that also persons with a limited number of followers are now involved in the business. The spread of marketing through digital solutions, social media and influencers has posed many sensitive questions as to consumer protection, particularly regarding unfair commercial practices and the protection of privacy on the internet. Due to the contrast between the fast technological development and the length of the law-making process, the effectiveness of consumer protection is under dispute and the further enhancement of social media in the future will increase the importance of the question.[18]

At first, in two occasions, the Italian Competition Authority (the AGCM)[19] has used moral suasion,[20] inviting influencers, undertakings, and their commercial partners to disclose advertisement purposes.[21] Next, the AGCM started investigations[22] against Alitalia S.p.A., Aeffe S.p.A., Alberta Ferretti and other *influencers* for possible hidden advertisement, which were later terminated after the subjects of these proceedings committed themselves to change their behaviour. Investigations concerned the alleged dissemination by various influencers

binomio digitale "influencer story-telling": la nuova pubblicità e la tutela dei consumatori' *Cyberspazio e diritto*, 436, 431-456 (2018), saying that thanks to influencer marketing, brands use the public figures social visibility to prevail over their competitors. In other terms, influencer marketing ends up in brand marketing.

[17] See Art 5 of decreto legislativo no 145 of 2007, transposing into the Italian legal system the rule contained in Art 14 of European Parliament and Council Directive 2005/29/EC, amending Directive 84/450/EEC on misleading advertisement, prescribing that advertisement always need to be easily recognisable.

[18] See M. Fiocca, n 16 above, 432.

[19] Even before the arousal of AGCM's attention, with the enactment of the 2016 code of conduct named 'Digital Chat', the Italian Authority on advertising standards ('IAP' – Istituto dell'Autodisciplina Pubblicitaria) was already well aware of the importance of disciplining the phenomenon of influencer marketing. In fact, the IAP stated that 'whenever comments or opinions regarding some products or brands, expressed by celebrities, influencers or bloggers have a commercial nature, the Code of conduct applies. For example, there is a commercial nature if a celebrity/influencer/blogger has signed an advertisement agreement with the owner of the band (or its commercial partners)'.

[20] On positive effects of moral suasion, see records of the hearing on 4th May 2017, reporting the beneficial outcome in the Amazon case, where the Authority invited the company to provide a more accurate information about the purchase process to its users.

[21] AGCM, 'ICA closes second moral suasion on influencers and brands, yet opens investigation into possible hidden advertising', available at *agcm.it*.

[22] AGCM, decision 22 May 2019 no 27787, available at agcm.it.

(including Alessia Marcuzzi, Chiara Biasi, Martina Colombari, Federica Fontana, Carlo Mengucci, Elena Santarelli, Giulia De Lellis, Cristina Chiabotto) on their Instagram profiles of posts artificially featuring the Alitalia logo printed on Alberta Ferretti clothing, worn in the commercial.

Such a conduct has been considered in violation of Arts 22 and 23, para 1, lett *m)* Consumer Code in that it represented a fraudulent omission with misleading effects on consumers, because of the failure to disclose commercial intent and the hidden advertising.[23] Furthermore, endorsement of products from specific brands was even more deceptive as it emerged from the influencer's apparently normal and relaxed everyday routine.[24] In so doing, for the purpose of the applicability of the rules contained in the Consumer Code, the competition watchdog has implicitly considered influencers as professional traders.[25]

Investigations have been terminated and no sanction has been imposed, for the Italian Competition Authority held commitments submitted by the undertakings subject to the proceedings to be able to meet the concerns expressed to them. More in particular, Alitalia has committed to '1) rigorously abide by the rule prohibiting unfair trade practices, with a specific reference to hidden advertisement, and that the top managers involved with influencer marketing shall issue a strong recommendation to avoid the occurrence of similar circumstances; 2) adopt Guidelines aimed at clarifying the rules of conduct that influencers will need to observe in their relationship with the Company. These Guidelines will be an essential part of the cooperation agreement among the parties so that, in case of violation, Alitalia will be entitled to impose to its counterparty a penalty determined on the basis of the value and characteristics of the contract; 3) introduce a standard clause in the agreements concerning the licensing and co-marketing of the Alitalia trademark, providing that commercial partners shall adopt all the measure and precautions required for preventing hidden advertisement, and remind to influencers the importance of complying in good faith with the requirements of the law'.[26]

Notwithstanding that no sanction has been issued,[27] it is clear how labile

[23] According to scholarly writings, the so-called 'misleading omissions' are voluntary misconducts that could justify annullability of the contract pursuant to Art 1439 of the Civil Code. See M. Nuzzo, 'Pratiche commerciali sleali ed effetti sul contratto: nullità di protezione o annullabilità per vizi del consenso?', in E. Minervini and L. Rossi Carleo eds, *Le pratiche commerciali sleali* (Milano: Giuffrè, 2007), 240, 235-244.

[24] In 2016, the US Federal Trade Commission has published its Endorsement Guides, regulating online advertising and marketing. Pursuant to such a Guide, influencers have to disclose advertisement and commercial purposes.

[25] For a definition of trader see the comment of the Case C-105/17 *Komisia za zashtita na potrebitelite* v *Evelina Kamenova*, Judgement of 4 October 2018 by A. Aiello, 'Nozioni di professionista e di pratiche commerciali nella giurisprudenza della Corte di giustizia dell'Unione europea' *Rivista diritto media*, 282-286 (2019); and by C. Scapinello, 'La nozione di "professionista" nel commercio elettronico' *Giurisprudenza italiana*, 1813-1823 (2019).

[26] See AGCM, n 22 above.

[27] The requirements of manifest misbehaviour and severeness were not met and the ICA

could be the balance between (influencer's) freedom of expression and consumer (as well as market) protection. On one hand, the correctness of consumers' decision-making process needs to be safeguarded from the threats posed by of hidden advertisements. On the other hand, individuals' ability to freely express themselves for non-commercial purposes also needs to be assured.

III. The Facebook and WhatsApp Cases

Since the above-mentioned case concerns the advertisement of goods and services meant for consumers, the consistency of the conduct of the Italian Competition Authority is undeniable. However, in respect of the measures adopted against WhatsApp and Facebook (nos 26596, 26597 and no 27432), the situation is different.

In such cases, the legitimacy of AGCM's decision depends on the possibility of defining the commercialisation of personal data on the social network in terms of a consumer relationship.[28]

With two decisions dated 11 May 2017, the Antitrust Authority ascertained that WhatsApp Inc. *de facto* forced the users of its service to fully accept the new Terms of Use, and specifically the provision to share their personal data with Facebook, since without granting such consent, they could not have been able to use the service anymore. In these decisions, the Provider's conduct has been regarded as a violation of Arts 20, 24 and 25 of Consumer Code and, consequently, a four million euros fine has been levied (then reduced to three, due to the precautionary suspension of data sharing with Facebook). Most importantly, the Italian Competition watchdog has ascertained that personal data, information and contents generated on *social media* have an economic value, and may well be used as consideration, instead of a monetary price, thus constituting part of a consumer relationship.[29]

considered the commitments submitted by influencers and undertakings to be a sufficient countermeasure.

[28] During Auditions in April 2018, Senator Orrin Hatch famously asked to Mark Zuckerberg: 'How do you sustain a business model in which users do not pay for your service?', who in turn replied: 'we run ads'. The Cambridge Analytica scandal unveiled how personal data were disclosed to third parties for profiling 87 million users. See C. Goanta and S. Mulders, 'Move Fast and Break Things: Unfair Commercial Practices and Consent on Social Media' *Journal of European Consumer and Market* Law, 136-146 (2019); C. Langhanke and M. Schmidt-Kessel, 'Consumer Data as Consideration' *Journal of European Consumer and Market Law*, 218-223, 220 (2015).

[29] See F. Bravo, *Il commercio elettronico dei dati personali*, in T. Pasquino et al eds, *Questioni attuali in tema di commercio elettronico* (Napoli: Edizioni Scientifiche Italiane, 2020), 83; G. Giannone Codiglione, 'I dati personali come corrispettivo della fruizione di un servizio di comunicazione elettronica' *Il diritto dell'informazione e dell'informatica*, 419, 390-425 (2017). See also M. Schmidt-Kessel, *Consent for the Processing of Personal Data and its Relationship to Contract*, in A. De Franceschi and R. Schulze eds, *Digital Revolution – New Challenges for Law* (München: Verlag C.H. Beck, 2019), 76, addressing the so-called bunding prohibition and

More recently, in December 2018, the Italian Competition Authority[30] levied a five million euros fine against Facebook for unfair commercial practices. The decision by the AGCM emphasised the deceiving nature of the advertisement 'it's free and will always be',[31] in so far as it failed to disclose to the users that their personal data would be used for commercial purposes. Furthermore, the social network also failed to provide genuine and accurate information on whether personal data would be used for profiling or personalised advertisement purposes.

A second violation consisted in the undue conditioning of users, whose personal data have been made available to third parties without any explicit authorisation. More specifically, through the pre-selection of the 'Active Platform function' Facebook pre-set users' ability to access websites and external apps using their accounts, thus enabling the transmission of their data to websites or apps, without any express consent.

Facebook has challenged the measures enacted by the Italian Competition Authority before the Regional Administrative Tribunal competent for the Region of the Lazio, which held that 'personal data have an economic value and Facebook failed to comply with information and disclosure requirements'. The Administrative Tribunal further specified that 'personal data may well represent a disposable asset, capable of being exchanged in consideration of goods and services within contractual relationships'.[32] The Tribunal, furthermore, held that Facebook 'needs to comply with existing legislation safeguarding accuracy,

highlighting that it concerns the case 'where the object of the consent has nothing to do with the contract but is only of accessory nature'.

[30] Public enforcement in the field of data processing is demanded to the cooperation of the Italian Competition Authority and the Communication Regulatory Authority, which is an independent authority established by legge 31 July 1997 no 249 regulating and controlling audio and video communication, as well as postal services and the press. It is composed of a President, a Commission for networks and infrastructures, a Commission for products and services, and a Council. Commissions are composed of the President and two Commissioners, while the Council is constituted with the presence of the President and all Commissioners. As for the adjudication power by the Italian Communication Regulatory Authority, see P. Rossi, 'Il nuovo Regolamento Agcom per la risoluzione delle dispute fra operatori, nella simbiosi tra regulation e adjudication' *Amministrazione in cammino*, 29 November 2017, 1-17; G. Nava, *Regolamentazione e contenzioso tra operatori nelle comunicazioni elettroniche* (Torino: Giappichelli, 2012), 108; F. Donati, *L'ordinamento amministrativo delle comunicazioni* (Torino: Giappichelli, 2007), 212; G. Della Cananea, 'Regolazione del mercato e tutela della concorrenza nella risoluzione delle controversie in tema di comunicazioni elettroniche' *Diritto pubblico*, 612, 601-618, (2005); see also, *si vis*, M. Zarro, *Le decisioni delle autorità amministrative indipendenti nelle controversie tra utenti e imprese* , in D. Mantucci ed, *Trattato di diritto dell'arbitrato* (Napoli: Edizioni Scientifiche Italiane, 2020), 650.

[31] After the proceedings, the anodyne motto 'Create an account. It's quick and easy' appears.

[32] Tribunale amministrativo regionale Lazio 10 January 2020 no 260, *Diritto e Giustizia* (2020), 13th January 2020. Decision no 261 does not seem to be published. The two decisions have the same content and motivation and also the challenged document is the same: ie decision no 27432 rendered in PS/11112 by the Italian Competition Authority during the hearing on 29 November 2018, notified on 7 December 2018 and published in *Rassegna di diritto Farmaceutico e della salute*, 205 (2019). In Judgment no 260 the challenging party is Facebook Inc., in the following Judgment no 261, the subsidiary Facebook Ireland Limited takes that role.

clearness and the non-deceiving nature of information provided to consumers', qualifying as misleading its advertisement[33] that its services would have been free of charges. Far from being free, in exchange of the services received, users pay a consideration consisting in the transfer of their personal data (such as personal identity, contacts, pictures, geographical address and preferences from previously visited websites).[34] These decisions from the Italian Competition Authority and the Administrative Tribunal are particularly important and deserve to be praised for their rigorousness. Even though, considering that 98% of Facebook's revenue derive from advertisement, similar steps could – and perhaps should – have been taken earlier.

The cases discussed above lead to the conclusion that there is reasonable ground for holding that the lack of a monetary transaction does not necessarily, and in all cases, entail that the product is 'free'. On the contrary, Recital 13 of Proposal for a Directive 634/2015 EU on certain aspects concerning contracts for the supply of digital content establishes that

> in the digital economy, information about individuals is often and increasingly seen by market participants as having a value comparable to money. Digital content is often supplied not in exchange for a price but against counter-performance other than money, ie by giving access to personal data or other data.[35]

The final version of Directive 2019/770/UE does not make any reference to the concept of 'consideration', as it would probably have been inappropriate after the European Data Protection Supervisor stressed that

> fundamental rights such as the right to the protection of personal data cannot be reduced to simple consumer interests, and personal data cannot be considered as a mere commodity.[36]

Instead, Art 3 refers to the case in which

> the trader supplies or undertakes to supply digital content or a digital service to the consumer and the consumer provides or undertakes to provide personal data

and states that such an exchange triggers the applicability of the Directive just

[33] Tribunale amministrativo regionale Lazio 10 January 2020, n 32 above.

[34] See C. Langhanke and M. Schmidt-Kessel, n 28 above, 223: 'A performance promised in exchange for consent to process with personal data is therefore not gratuitous. There is a valuable consideration, which leads to a synallagmatic contract'.

[35] The same conclusion is also reached in Commission's Guidance on the implementation/application of European Parliament and Council Directive 2005/29/EC.

[36] European Data Protection Supervisor, Opinion 4/2017 on the Proposal for a Directive on certain aspects concerning contracts for the supply of digital content, 14 May 2017.

like the payment of a price. Consequently, there is an express textual recognition of the economic value of personal data, which imposes to consider the relationship between the users and internet service providers under the light of consumer law. This is the reason why Authorities involved in public enforcement need to investigate the distortive effects on the market of misleading advertisement, alluring consumers with the promise of free services and jeopardizing the functioning of the Internal market. Thus, a synergic approach towards personal data protection, consumer protection and a correct functioning of the market should be implemented, meaning that also the market of personal data should be defined a 'regulated market' due to the penetrating powers of regulatory authorities.[37]

IV. The TripAdvisor Case

Unfair commercial practices are also linked to online reputation.[38]

[37] This is also in line with the German *Bundeskartellamt* 6th February 2019, B6-22/16, available at *bundeskartellamt.de*, with comments by G. Colangelo and M. Maggiolino, 'Antitrust Über Alles. Whither Competition Law after Facebook?' *World Competition*, 355-376 (2019); M. Midiri, 'Privacy e antitrust: una risposta ordinamentale ai Tech Giant', available at *federalismi.it*, 209-234 (2020) sanctioning the abuse of dominant position by Facebook for making the use of its social network by private users, who also use related services such as WhatsApp and Instagram, conditional on the collection of data and the combining of such information with the use's accounts, without an explicit and genuine consent. According to the *Bundeskartellamt* such a conduct would violate § 19 of *Gesetz gegen Wettbewerbsbeschränkungen*, prohibiting the abuse of a dominant position. Facebook's dominant position has been established on the grounds of the high number of daily active users and the low level of supply-side substitutability. Furthermore, thanks to its market position and the inappropriate processing of data, Facebook had access to a large number of further sources, securing its competitive edge over the competitors and increasing market entry barriers. Facebook's market power reflected on end customers. According to *Bundeskartellamt*, users were not in a position to voluntarily give their consent to the treatment of personal data due to Facebook's dominant position. In fact, users would have to refrain from using a variety of services, if they did not want to add any more data to their extensive Facebook database. And such an option greatly reduced their choice. The company has thus violated not only competition law, but also European Parliament and Council Regulation 2016/679/EU of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC [2016] OJ L119/1, since consent was not freely given, for users had no genuine or free choice, being unable to refuse or withdraw consent without detriment. Also the Australian Competition and Consumer Commission, in a preliminary report published in 2018, the Australian Competition and Consumer Commission, *Digital Platforms Inquiry, Preliminary Report*, available at *accc.gov.au* (2018), has expressed its concerns over the collection via social media of a huge amount of personal data with a relevant economic value, through means which go beyond the genuine and unfettered consent of the users. The Australian Competition and Consumer Commission especially highlighted the threats and the complexities hidden in terms of service and wishes for an urgent legislative reform, aimed at providing consumers with a better understanding of their rights and knowledge of the functioning of similar online platforms.

[38] See G. D'Alfonso, 'Recensioni 'diffamatorie' in rete e lesione della reputazione digitale d'impresa. Illecito aquiliano e valutazione comparativa degli interessi dell'impresa e degli internauti, alla luce degli indirizzi giurisprudenziali sui limiti all'esercizio del diritto di critica' *Diritto, mercato, tecnologia*, 2, 1-48 (2019), pointing out that, with the evolution of the internet, business reputation

The starting point of the enquiry regarding online reputation are the features of the contemporary internet-based society, whereby citizens, undertakings, and companies are more and more connected with each other, exchanging information and data nature through social media networks. The enormous amount of information available on the internet, such as reviews, comments, and pictures may have both a positive or a negative effect on consumers' preferences on specific products or services. In such a context, companies do feel the need to exploit the opportunities offered by online markets.[39] Several empirical studies proved that reputation significantly influences the overall value of a company, with economists trying to establish whether there is a biunivocal relationship between online reputation and a company's performance (not only in financial terms). Reputation, as a matter of fact, is a diversifying element from other competitors, and it is non-replicable being, therefore, of the utmost importance for competitive advantage. Consequently, a link between reputation and the economic performance of a company really seems to exist.

A new 'reputation system' has emerged in the current economic and legal scenario.[40] It is a system based on online platforms operated on a peer-to-peer basis, relying on qualitative reviews and numerical ratings from users.[41] Triangular relationships are established between sellers, consumers, and the operator of the platform, aimed at building trust among strangers using a collaborative marketplace. As a result of these triangular transactions, a great deal of information and data is uploaded on the internet and everything becomes capable of being measured:

> hotels and restaurants (*Yelp*, *Tripadvisor*, *Booking*), sellers (*Ebay*, *Amazon*), professionals – starting from lawyers (*Avvo*) and scholars (*RateMyProfessors*) – and private individuals (*Airbnb*, *Uber*) are rated.[42]

has assumed a different meaning. Digital reputation describes the idea of internet users, based on information gathered online. See, *si vis*, M. Zarro, 'La tutela della reputazione digitale quale «intangible asset» dell'impresa' *Rassegna di diritto civile*, 1514, 1504-1531 (2017); A. Ricci, *La reputazione: dal concetto alle declinazioni* (Torino: Giappichelli, 2018), 174; Id, 'Il valore economico della reputazione nel mondo digitale. Prime considerazioni' *Contratto e impresa*, 1298, 1297-1316 (2010); N. Di Stefano and F. Giannone, *Manuale sulla web reputation. Dall'identità digitale all'economia della reputazione*, available at https://tinyurl.com/rrer22w (last visited 30 June 2021); L. Carota, 'Diffusione di informazioni in rete e affidamento sulla reputazione digitale dell'impresa' *Giurisprudenza commerciale*, I, 629, 624-638 (2017); A. Fusaro, *Informazioni economiche e "reputazione d'impresa" nell'orizzonte dell'illecito civile* (Torino: Giappichelli, 2010), 1.

[39] G. Atti et al, La quarta rivoluzione industriale: verso la supply chain digitale: Il futuro degli acquisti pubblici e privati nell'era digitale (Milano: Franco Angeli, 2018).

[40] L. Carota, n 38 above, 624; G. Smorto, 'Reputazione, fiducia e mercati' *Europa e diritto privato*, 199-218 (2016). On digital reputation, see A. Ricci, n 38 above, 168.

[41] C. Busch, Crowdsourcing consumer confidence. How to regulate online rating and review systems in the collaborative economy, in A. De Franceschi ed, *European Contract Law and the Digital Single Market* (Cambridge: Intersentia, 2016), 223.

[42] G. Smorto, *Reputazione, fiducia e mercati*, n 40 above, 169. See also C. Busch, n 41 above, 224.

Even though reputational feedback systems are being carefully studied to increase their reliability, there are still several problems with the reliability of ratings, reviews, and comments that need to be addressed in an effective way.[43]

The most important limit of these reputational feedback systems is the lack of any preliminary control on the accuracy and reliability of the information entered by users. Because of the lack of preventive mechanisms, users of the platform and traders may indulge in abusive conducts with distortive effects on the markets. Instigation may come from many different circumstances: eg, due to the direct relationship between online reputation and earnings, a company may be interested in artificially inflating its ratings, or discrediting a rival undertaking on the platform; the same consumers/users may be positively or negatively biased towards a specific product. However, notwithstanding the different reasons, as a matter of fact, abusive conducts, biased comments and fake feedbacks negatively affect trust among users, dissolving the reputational bonds necessary for the functioning of the marketplace. For this reason, they represent an unfair commercial practice, with misleading effects on consumers relying on the accuracy of information provided by the online platform.

One of the most relevant decision on these issues is the TripAdvisor case, concerning the relationship between business reputation and unfair commercial practices. As widely known, this online platform offers to the public user-generated contents such as tourist information, feedbacks of hotels, restaurants *et alia*. To post a review, users older than 13 years of age simply need to register on the website and accept terms and conditions. Users can even register with more than one account,[44] and are not required to have actually purchased the product they review. This lack of regulation leads to a higher risk that users could provide or be provided themselves with misleading information, so to nullify the positive effects of the collaborative platform.[45]

With decision no 25237, dated 12 December 2014, the Italian Competition Authority levied a € 500,000 fine against *TripAdvisor* LLC and *TripAdvisor Italy* s.r.l. for unfair commercial practices. According to the investigations conducted by the Competition Authority, TripAdvisor was found responsible for the spread of misleading information, because

> while stating that it does not check the facts set out in the reviews, and while aware that (…) on the said website fake reviews, both positive and negative in their judgments, are published by users who have not actually

[43] G. Smorto, n 40 above, 173.

[44] Authentication of personal identity is not a requisite for registration, to the extent that the identity of the review remains secret. See, L. Vizzoni, 'Recensioni non genuine su TripAdvisor' *Responsabilità civile e previdenza*, 710, 706-722 (2018).

[45] ibid 711, recapitulating unfair practices such as: boosting, ie the artificial boost on online platforms of a company's profile; digital vandalism, when a company libels a competitor's product; and optimization, consisting in buying a stock of positive reviews by specialized agencies.

> availed themselves of the services provided by the facilities included in the database, uses particularly assertive information, capable as such of increasing the consumers' trust in the authentic and genuine character of the reviews published by users.[46]

One of the commercials under the scrutiny of the Italian watchdog, still flagged on the website, reads

> it does not matter if you prefer chain hotels or niche resorts: on TripAdvisor you can find many reviews, true and authentic, which you can trust. Millions of travellers have published online their most sincere views on hotels, beds & breakfast, pensions and much more still.

The Competition Authority had no doubts that false information has been provided as investigation proceedings established that many reviews concerned inactive undertakings, or were traceable to a fantasy character, or were related to a period during which the advertised activity happened to be close.

Both TripAdvisor LLC and TripAdvisor Italy s.r.l. challenged the above-mentioned fine before the Administrative Tribunal competent for the Region of Lazio. With judgement no 9355/2015,[47] the Administrative judges reversed the arguments of the Italian Competition Authority, holding that the platform property advertised the functioning of the website, also due to the express warning that the trustworthiness of the reviews could not be certified and that reviews and comments by users were mere opinions.

The Competition Authority appealed the judgement rendered by the Tribunal, and the State Council confirmed the fine levied against TripAdvisor, holding that commercials published on TripAdvisor's website were able to mislead consumers regarding the truthfulness of the ratings and reviews posted by users. However, the Court seated in Palazzo Spada reduced the sanction from € 500,000 to € 100,000.[48]

There is no doubt that TripAdvisor should be considered as a hosting provider for the purposes of decreto legislativo 9 April 2003 no 70,[49] as it administers the platform and earns its profits from pay-per-click advertisement; the price set to its commercial partners is directly proportional to the number of clicks generated from viewers logged into the system. As a hosting provider, according to Art 18, para 1, b, of Consumer Code, TripAdvisor may well be

[46] See, the comments by B. Blasco, 'Falsità delle recensioni in internet, astroturfing e scorrettezza delle pratiche commerciali' *I contratti*, 231-242 (2017).

[47] Tribunale amministrativo regionale Lazio 13 July 2015 no 9355 with note by E. Della Bruna, 'Ingannevolezza della comunicazione commerciale, (in)adeguatezza organizzativa e responsabilità degli internet provider (il caso TripAdvisor)' *Rivista di diritto dell'impresa*, 417-430 (2016). See also L. Vizzoni, n 44 above, 706.

[48] AGCM, 15 July 2019, Decision no 4976, available at *agcm.it*.

[49] E. Della Bruna, n 47 above, 418.

considered as a professional for the purposes of the applicability of consumer protection law.[50]

Technological developments and the rise of social media have changed not only the traditional features of marketplaces, but also the common way of doing business. Social studies have demonstrated the economic significance of reputation. From a legal point of view, it now has to be determined how and in what terms reputation may be considered as an intangible asset, fully protected by the legal system and capable of being transferred.

Being online business reputation an incorporeal factor of production, it may well be considered as an intangible asset, which needs to be protected as

[50] See, Opinion of Advocate General Szpunar, 31st May 2018, C-105/17, available at *europa.eu*. The Advocate proposed to interpret Art 2(b) of European Parliament and Council Directive 2005/29/EC as meaning that a natural person, such as the defendant in the main proceedings, registered on an online platform for the sale of goods cannot be classified as a 'trader' when publishing, on that website, eight advertisements at the same time for the sale of different products. However, added the Advocate, it is for the referring court to assess whether such a person may be defined as a 'trader' and, therefore, whether the activity she carries out constitutes a 'commercial practice' within the meaning of Art 2(d) of the Directive 2005/29 EC. The Advocate has reached such a conclusion based on the argument that the simultaneous publication on an online platform of a total of eight advertisements for the sale of different new and used products does not seem to be sufficient to allow use of the classification of 'trader' within the meaning of EU directives. However, it has been noted that 'the classification of 'trader' requires 'a case-by-case approach' and that it is therefore appropriate, for the referring court to carry out a specific analysis to establish whether a person is covered by the definition of 'trader'. That analysis will seek, in particular, to establish whether the online platform sale was made in an organised manner and for profit; whether that sale occurs over a certain duration and with a certain frequency; whether the seller has a legal status which enables her to engage in commercial transactions, and to what extent the online sale is connected to the seller's commercial activity; whether the seller is subject to VAT; whether the seller, acting in the name of a specific trader or on his behalf or through any other person acting in her name or on her behalf, received remuneration or an incentive; whether the seller purchases new or used goods with a view to selling them on, thus making that a regular, frequent and/or simultaneous activity in relation to her trade; whether the amount of profit generated on the sales confirms that the transaction made falls within the scope of a commercial activity, and/or whether the products for sale are all of the same type or value, in particular, whether the offer is focused on a limited number of products. It should be noted that those criteria are neither exhaustive nor exclusive, and therefore, in principle, meeting one or more of the criteria does not, in itself, establish the classification to be used in relation to an online seller with regard to the concept of 'trader'. It will therefore be necessary to make an overall assessment taking account of all the relevant criteria in order to decide on the classification to be used. Those criteria will thus enable the national courts to determine whether a person is carrying out a commercial activity which places him in a stronger position than the consumer and, consequently, whether there is an imbalance between the trader and the consumer. However, it is for the referring court, in view of the foregoing considerations, to assess, on the basis of the facts available to it and based, inter alia, on the criteria set out in the preceding points, whether that person may be classified as a 'trader' within the meaning of those directives». The European Court of Justice, following the arguments laid down by the Advocate General, held that «a natural person who publishes simultaneously on a website a number of advertisements offering new and second-hand goods for sale can be classified as a 'trader', and such an activity can constitute a 'commercial practice', only if that person is acting for purposes relating to his trade, business, craft or profession, this being a matter for the national court to determine, in the light of all relevant circumstances of the individual case'.

such, according to its characteristics and in compliance with the purposes it is deemed to serve. However, from a different point of view, online business reputation has its own peculiarities, so that it cannot be simply equated to that of natural persons. As a personality right, reputation of natural persons cannot be merchandised and may not be considered under any circumstance as a factor of production. This argument implies that the grounds of protection of online business reputation, as well as the remedies for its enforcement, are necessarily different from those of natural persons.

This is even more important, considering the evolution of business models in last decades and that a complex network of relationships based on mutual trust and cooperation has developed against a highly hierarchical and vertically integrated company structure – a twist branded by the German scholars as V*erbund – Verband – Verkehr*.[51]

The spread of a business network based on a cooperative approach, whereby every interaction allows the members to add a valuable contribution to the overall efficiency of the system, significantly increases the importance of trust, the lack of which determines the collapse of the whole structure. Furthermore, trade libel or unfair commercial practices not only affect a company's position in the complex web of interactions of which markets are made but, due to the way communication flows on the internet and its rapidity, also makes it almost impossible a restore the *status quo ante*. This being so, it would prove ineffective to search a remedy against such violations by subsuming the specific case under a general rule, instead of adopting a case-by-case approach able to properly evaluate the specific circumstances of the case. Holding the contrary would mean to inadmissibly conceive the issue of online business reputation, personal data and know how protection as an *ius singularis*, confined in their own compartments.

A discipline based on general rules and detailed technical regulation would not only be inadmissible, as it would bureaucratise the protection of fundamental values for society, but it would also be inadequate. In fact, due to the fast development of new technologies, a technical and excessively detailed regulation would soon become obsolescent.[52] An example in this sense is given by the decreto legislativo no 70/2003,[53] transposing into the national legal system Directive 2000/31/EC on certain legal aspects of information society services. This legal instrument provides for a limitation of the liability of intermediary

[51] G. Teubner, "Verbund', 'Verband' oder 'Verkehr'?: zur Außenhaftung von Franchising-Systemen' *Zeitschrift für das gesamte Handelsrecht und Wirtschaftsrecht* , 154 (1990); transl in Id, 'Beyond Contract and Organization? The External Liability of Franchising Systems in German Law'*,* in C. Joerges ed, *Franchising and the Law: Theoretical and Comparative Approaches in Europe and the United States* (Baden-Baden: Nomos, 1991), 105.

[52] P. Perlingieri, 'Privacy digitale e protezione dei dati personali tra persona e mercato' *Il Foro napoletano*, 483, 481-488 (2018).

[53] Regarding hosting providers' liability see M. Gambini, *Principio di responsabilità e tutela aquiliana dei dati personali* (Napoli: Edizioni Scientifiche Italiane, 2018), 24.

service providers, unless they have actual knowledge of the illegality of the activity or information stored at the request of a recipient of the service. Such a Decree is now obsolete,[54] as it is unable to prevent potential risks coming from users, traders and marketers. As seen in the TripAdvisor case, social media and other kinds of online cooperative platforms indiscriminately make it possible for anyone to publish reviews and comments, even if containing false, misleading or deceptive statements. For this reason, judges called to solve similar cases struggled to ground service provider's liability on the above-mentioned legislation. Instead, to oblige information society service providers to remove manifestly wrong or defamatory information, judges availed themselves of the ordinary causes of action, especially Art 700 of the Code of Civil Procedure, which, however, requires the applicant to establish that the risk of an imminent and irreparable harm.

V. Consumer Protection from the Unfair Trade Practices on the Internet. A Recent European Legislative Measure

The cases discussed in the previous sections show how evasive unfair commercial practices and misleading advertisement could be and how hard it may be, even for the most prepared and up-to-date consumer, to recognise similar phenomena when they occur. It is thus clear that private enforcement of consumer law risks to be ineffective, in as much as no one could ever enforce rights and redress wrongs of which he is not aware. Such a gap between consumers and traders cannot be filled by the internet, and requires an intervention by the State so to create valid tools for consumer law enforcement in the era of digital marketplaces and avoid distortions.[55]

[54] See European Parliament and Council Directive 2000/31/EC of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market [2000] OJ L178/1, which has been transposed into the Italian legal system with decreto legislativo 9 April 2003 no 70. See also, C-18/18, *Eva Glawischnig Piesczek* v *Facebook Ireland Limited*, Judgement of 3 October 2019, available at www.eur-lex.europa.it At the national level, case-law focuses on the definition of provider and the scope of its responsibility, see Corte d'Appello di Roma 29 April 2017 no 2833 with comments by G. Cassano, 'Nozione di provider e delimitazione della responsabilità: la giurisprudenza prende una direzione' *Il diritto industriale* 181, 185-187 (2018); Tribunale di Torino 7 April 2017 no 1928 commented by V. Vozza, 'La responsabilità civile degli Internet Service Provider tra interpretazione giurisprudenziale e dettato normative' *Danno e responsabilità*, 78-86 (2018). The author highlights the urgent need of a legislative reform, due to the spread of online platforms offering the opportunity to upload potential illicit materials. In such a context, European Parliament and Council Directive 2000/31/EC shows all its limits, which are only in part due to the quick developments of an internet-based society. See also Art 13 of European Parliament and Council Directive 2016/0280/UE on copyright in the Digital Single Market, concerning the use of protected content by information society service providers storing and giving access to large amounts of works and other subject-matter uploaded by their users.

[55] See A. Quarta, 'Il diritto dei consumatori ai tempi della peer economy. Prestatori di servizi e *prosumers*: primi spunti' *Europa e Diritto Privato*, 667-681 (2017). The Author agrees with the

This new approach, aimed at promoting the proper functioning of the market through the instruments of contract law, is exemplified by the allocation to national competition authorities of supervisory powers over unfair clauses in contracts concluded with consumers and unfair commercial practices.[56] Due to the interdependency of competition and consumer law, strategies for the protection of consumers and marketplaces cannot be compartmentalised. A synergistic approach should instead be preferred, so to modulate interventions in an efficient way using the limited resources available.[57]

For this purpose, public enforcement of consumer law has been further improved by the EU legislation. National Authorities have the power to ascertain, prohibit – also in the form of an interim measure – and levy sanctions, once limited to unfair commercial practices (Art 27 Consumer Code), has also become available in cases involving contracts negotiated away from business premises (Art 66, para 1, Consumer Code), as well as timeshare and long-term holiday contracts (Art 79 Consumer Code).[58]

The latest steps taken in this direction are Directive 2019/1/EU [59] and

solution suggested in the text highlighting, however, that 'collaborative markets could suffer from information asymmetry, being thus a public intervention strongly recommended'.

[56] Decreto legislativo 6 September 2005 no 206, introducing Art 37 bis of the Consumer Code, has given to AGCM the power to evaluate unfairness of clauses put into terms and conditions of standard contracts. See, L. Rossi Carleo, Sub *art. 37* bis, in E. Capobianco et al eds, *Codice del consumo annotato con la dottrina e la giurisprudenza* (Napoli, Edizioni Scientifiche Italiane, 2nd ed, 2019), 240; A. Barenghi, , Sub *art. 37* bis, in V. Cuffaro ed, *Codice del consumo* (Milano: Giuffrè, 5th ed, 2019), 364.; E. Minervini, *Dei contratti del consumatore in generale* (Torino: Giappichelli, 3rd ed, 2014), 147; D. Achille and S. Cherti, *Le clausole vessatorie nei contratti tra professionista e consumatore*, in G. Recinto et al eds, *Diritti e tutele dei consumatori* (Napoli: Edizioni Scientifiche Italiane, 2014), 96; M. Angelone, 'La tutela amministrativa contro le clausole vessatorie alla luce dell'attività provvedimentale condotta dall'AGCM nel triennio 2013-2015' *Concorrenza e mercato*, 525-551 (2016); Id, 'La nuova frontiera del «public antitrust enforcement»: il controllo amministrativo dell'Agcm avverso le clausole vessatorie' *Rassegna di dritto civile,* 9-40 (2014)

[57] See AGCM, Relazione annuale sull'attività svolta nel 2017, available at agcm.it, 221.

[58] M. Angelone, 'La «degiurisdizionalizzazione» della tutela del consumatore' *Rassegna di diritto civile,* 723-728 (2016); Id, 'Diritto privato «regolatorio», conformazione dell'autonomia negoziale e controllo sulle discipline eteronome dettate dalle authorities' *Nuove autonomie*, 453, 441-461 (2017); A. Tucci, 'Strumenti amministrativi e mezzi di tutela civilistica: verso un superamento della contrapposizione?' *Rivista di diritto bancario*, 84, 75-100 (2020), highlighting the European trend towards public enforcement through Independent Authorities.

[59] European Parliament and Council Directive 2019/1/EU of 11 December 2018, to empower the competition authorities of the Member States to be more effective enforcers and to ensure the proper functioning of the internal market [2019] OJ L11/3. Such a new Directive counts sixty-seven recitals and is aimed to increase independence, autonomy of National Authorities, as well as the effectiveness of investigations and sanctions. Independence is defined by Art 4 as the lack of any conditioning by Governments, politicians, which could influence the decisions of National Authorities. Fort this purpose, Art 4, para 4, states that 'members of the decision-making body of national administrative competition authorities are selected, recruited or appointed according to clear and transparent procedures laid down in advance in national law'; they should not 'take any instructions from government or any other public or private entity'» (Art 4, para 2, lett. b, Directive 1/2019). Furthermore, independence is guaranteed through the strengthening of competition authorities' financial, technical, infrastructural, and human resources, so to be able to properly serve

2019/2161/EU,[60] pursuing the modernisation of the existing consumer protection rules, by filling the gaps in national legislations regarding the deterrence and sanctioning of intra-Union infringements.[61] They also point out some more effective remedies against unfair commercial practices and make up for the shortages of injunctive reliefs in consumer protection, trying to go beyond the limits of Directive 2009/22/EC.[62]

It is thus necessary to evaluate whether the promise for stronger public and private enforcement tools and better redress opportunities,[63] made by the Commission with its 2018 Communication titled 'A new deal for consumers', are kept.

Reading the Recitals of the Directive, especially no 17 and followings, a specific focus on unfair commercial practices in online marketplaces emerges therefrom. Due to the need of more effective consumer protection measures, the Directive adapts the law to fit the developments of a digital economy and the increasing economic importance of personal data.[64]

More specifically, Recital 18 states that a

> higher ranking or any prominent placement of commercial offers within online search results by the providers of online search functionality

their statutory purposes. Moreover, Art 10 gives to National Competition Authorities the power to impose any behavioural or structural remedies which are proportionate to the infringement committed and necessary to bring the infringement effectively to an end. For a comment on the text of the Directive, see, F. Ghezzi and B. Marchetti, 'La proposta di direttiva in materia di rete europea della concorrenza e la necessità di un giusto equilibrio tra efficienza e garanzie' *Rivista italiana di diritto pubblico comunitario*, 1015-1075 (2017), European Parliament and Council Directive 2019/1/EU needs to be examined together with European Parliament and Council Regulation (EU) 2017/2394 of 12 December 2017 on cooperation between national authorities responsible for the enforcement of consumer protection laws and repealing Regulation (EC) No 2006/2004 [2017] OJ L345/1, trying to promote a closer cooperation among National Authorities from different Member States.

[60] European Parliament and Council Directive 2019/2161/EU n 6 above. For a comment of the newly enacted Directive, see A. Cilento, '"New deal" per i consumatori: risultati all'altezza delle ambizioni?' *Contratto e impresa*, 1195-1216 (2019); R. Caponi, 'Ultime dall'Europa sull'azione di classe' *Il Foro italiano*, 332 (2019); M. Loos, 'The Modernization of European Consumer Law (Continued): More Meat on the Bone After All' *European Review of Private Law*, 407-423 (2020); B. Duivenvoorde, 'The Upcoming Changes in the Unfair Commercial Practices Directive: A Better Deal for Consumers?' *Journal of European Consumer and Market Law,* 219-228 (2019); J. Van Duin,' The Real (New) Deal: Levelling the Odds for Consumer Litigants: On the Need for a Modernization' *European Review of Private Law*, 1227-1249 (2019).

[61] The issue has already been dealt with by the European lawmaker with European Parliament and Council Regulation (EU) 2017/2394, regarding cooperation between national authorities responsible for the enforcement of consumer protection laws.

[62] Recital 3, European Parliament and Council Directive 2019/2161/EU n 6 above.

[63] Cfr Communication from the Commission, 11 April 2018, addressed to the Parliament and to the Council [COM (2018) 183 final].

[64] See, Z. Efroni, 'Gaps and Opportunities: The Rudimentary Protection to 'Data-Paying Consumers' under New EU Consumer Protection Law" *Common Market Law Review*, 799-830 (2020).

has an important impact on consumers.[65]

Whereby,

> ranking refers to the relative prominence of the offers of traders or the relevance given to search results as presented, organised or communicated by providers of online search functionality, including resulting from the use of algorithmic sequencing, rating or review mechanisms, visual highlights, or other saliency tools, or combinations thereof.[66]

According to the new Directive, therefore, it should be clearly stated that

> practices where a trader provides information to a consumer in the form of search results in response to the consumer's online search query without clearly disclosing any paid advertising or payment specifically for achieving higher ranking of products within the search results should be prohibited.

In other words, the duty to disclose relevant information to consumers and transparency should be enhanced, so that online platforms will be required to declare whether search results contain any paid advertisement. Accordingly, the query result will still appear in the online search, but consumers will be advised that the reason why that query search appears is not because it is more or less fit to their preferences, but because it is a paid advertisement.

As per Recital 22,

> traders enabling consumers to search for goods and services, such as travel, accommodation and leisure activities, offered by different traders or by consumers should inform consumers about the default main parameters determining the ranking of offers presented to the consumer as a result of the search query and their relative importance as opposed to other parameters. That information should be succinct and made easily, prominently and directly available.

The law also weighed the need to improve the efficiency of consumer protection with business trade secrets, providing that 'the information requirement regarding the main parameters determining the ranking is without prejudice to Directive (EU) 2016/943'.[67] As a result, traders should not be required to disclose

[65] Recital 18, European Parliament and Council Directive 2019/2161/EU n 6 above.

[66] Recital 19, European Parliament and Council Directive 2019/2161/EU n 6 above.

[67] European Parliament and Council Directive 2016/943/EU of 8 June 2016, on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure [2016] OJ L157/1. See, G. Chiappetta, 'La proposta di direttiva e le linee di evoluzione della disciplina del Know-how *rectius* delle informazioni aziendali riservate' *Il diritto industriale*, 169-188 (2016); V. Falce, 'Tecniche di protezione delle informazioni riservate.

the detailed functioning of their ranking mechanisms, including algorithms.

A great deal of attention is dedicated to fake reviews and endorsements and, therefore, it is provided that

> when traders provide access to consumer reviews of products, they should inform consumers whether processes or procedures are in place to ensure that the published reviews originate from consumers who have actually used or purchased the products.

As a consequence, stating that reviews were submitted by consumers who actually used or purchased a product, when no reasonable and proportionate steps were taken to ensure that they originate from such consumers, shall constitute an unfair commercial practice because of its misleading effects on consumers.[68]

Recommending to Member States to strengthen private enforcement of consumer law, the Directive provides that consumers should be able to avail themselves of several remedies to seek redress for unfair commercial practices. Member States are encouraged to make available remedies for consumers so to ask for compensation of damages and, where appropriate, a price reduction or termination of the contract, in a proportionate and effective manner. Furthermore, Member States are not prevented from maintaining or introducing rights to other remedies, such as repair or replacement in order to ensure full removal of the effects of unfair commercial practices.[69]

The question to be answered is whether it will be really possible to deliver the level of effectiveness envisaged by the Commission in the launching a new deal for consumers.[70] The objective is to overcome the limits of the previous Directive

Dagli accordi TRIPs alla direttiva sul segreto industriale' *Il Diritto Industriale*, I, 129- 157 (2016); D. Mastrelia, 'La tutela del know-how, delle informazioni e dei segreti commerciali fra novità normative, teoria e prassi' *Il Diritto Industriale*, 519, 513-523 (2019). *See* also, *si vis*, M. Zarro, 'Notazioni in tema di possesso degli «intangibles»: il caso del «know how»' *Il Foro napoletano*, 183-204 (2018), regarding the differences among Member States as to the protection of know how. The common law of Confidence protected all kind of confidential information, either they were commercial, technical or personal. Remedies against violation of trade secrets were even more staggered. European Parliament and Council Directive 2016/943/EU has been transposed into the Italian legal system with decreto legislative 11 May 2018 no 63, which significantly amended the Code of Industrial Property. The current Art 98 C.I.P., as recently amended, defines the meaning of "trade secret" requiring that all the following elements should be met: corporate and technical information, which is not disclosed to the public, have an economic value, are object of protection.

[68] Recital 47, European Parliament and Council Directive 2019/2161/EU n 6 above

[69] Recital 16, European Parliament and Council Directive 2019/2161/EU n 6 above.

[70] See, G. Vettori, 'Effettività delle tutele (diritto civile)' *Enciclopedia del diritto*, (Milano: Giuffrè, 2017), Annali X, 381; Id, 'L'attuazione del principio di effettività. Chi e come' *Rivista trimestrale di diritto e procedura civile*, 939-959 (2018); Id, 'Il diritto ad un rimedio effettivo nel diritto privato europeo' *Rivista di diritto civile*, 666-694 (2017); Id, 'Controllo giudiziale del contratto ed effettività delle tutele. Una premessa' *Nuova giurisprudenza civile commentata*, 151-160 (2015); Id, 'Contratto giusto e rimedi effettivi' *Rivista trimestrale di diritto e procedura civile*, 787-815 (2015); M. Libertini, 'Le nuove declinazioni del principio di effettività' *Europa e Diritto Privato*, 1071-1096 (2018); G. Carapezza Figlia and S. Sajeva, 'Responsabilità civile e tutela

2005/29/EC , which failed to address the issue of remedies available against unfair commercial practices, and leaved the matter entirely to public enforcement.[71] At the same, it is also questionable whether the European lawmaker will succeed in harmonising procedural aspects of consumer protection, since the magnitude of the effects of online unfair commercial practices would suggest the adoption of a common frame of rules and principles.[72]

As for the first question raised, it is clear that over-reliance upon compensation for damages shows that the European lawmaker is far from having a clear scenario of what needs to be done to improve consumer protection enforcement. To be precise, the Commission's new deal for consumers needs to be assessed in the Dieselgate context, whereby a compensatory remedy seemed to be the more appropriate remedy to redress the harm suffered by consumers.[73] However, not every case has the same characteristics and compensatory relief is not always the best remedy available. Another example in this direction is given by unfair clauses in loan agreements, where the most appropriate remedies is relative nullity, possibly with retroactive effects, as recently recognised by the Court of Justice of the European Union.[74]

Regarding the second question, the efforts put into the Directive do not appear sufficient for a real harmonisation of procedural rules on consumer protection. Member States being free to choose whether to provide consumers with the option for repair or replacement of the product, in addition to the remedies already available, does not really militate in favour of a harmonisation of the remedial framework, nor of an improvement of private enforcement.

Regarding public enforcement, the Directive mandates Member States to provide for effective, proportionate and dissuasive penalties to contrast infringements of the principles enshrined into it. While remedies for the private enforcement of consumer protection envisage unfair commercial practices,

ragionevole ed effettiva degli interessi', in G. Perlingieri and A. Fachechi eds, *Ragionevolezza e proporzionalità nel diritto contemporaneo* (Napoli: Edizioni Scientifiche Italiane, 2017), I, 161.

[71] A. Fachechi, *Pratiche commerciali scorrette e rimedi negoziali* (Napoli: Edizioni Scientifiche Italiane, 2012), 31.

[72] See, J. Van Duin and C. Leone, 'The Real (New) Deal: Levelling the Odds for Consumer Litigants: On the Need for a Modernization', Part II *European Review of Private Law,* 1230, 1227-1249 (2019), H.W. Micklitz and N. Reich, 'The Court and Sleeping Beauty: The Revival of the Unfair Contract Terms Directive' *Common Market Law Review*, 771-808 (2014).

[73] See I. Garaci, 'Il dieselgate. Riflessioni sul private e public enforcement nella disciplina delle pratiche commerciali scorrette' *Il diritto industriale*, 61-76 (2018).

[74] Joined Cases C-154/15, C-307/15 and C-308/15 Francisco Gutiérrez Naranjo v Cajasur Banco SAU, Ana María Palacios Martínez v Banco Bilbao Vizcaya Argentaria SA (BBVA), Banco Popular Español SA v Emilio Irles López and Teresa Torres Andreu, Judgement of 21 December 2016, with comments by S. Pagliantini, 'La non vincolatività delle clausole abusive e l'interpretazione autentica della Corte di giustizia' *I contratti*, 11-25 (2017), 11. See also G. D'Amico, Mancanza di trasparenza di clausole relative all'oggetto principale del contratto e giudizio di vessatorietà (Variazione sul tema dell'armonizzazione minima), in G. D'Amico and S. Pagliantini eds, *L'armonizzazione degli ordinamenti dell'Unione europea tra principi e regole. Studi* (Torino: Giappichelli, 2018), 89.

remedies available for public enforcement apply also in respect of Directive 2011/83/EU on consumer rights, Directive 93/13/EEC on unfair terms in consumer contracts, and Directive 98/6/EC, on consumer protection in the indication of the prices of products offered to consumers. Furthermore, to ensure the deterrent effect of these remedies, Directive 2019/2161/EU provides that Member States should set in their national law the maximum fine for widespread infringements at a level that is at least four percent of the trader's annual turnover in the Member State or Member States concerned. However, proportionality principle shall always be guaranteed.

With specific reference to unfair commercial clauses, Member States may set sanctions when companies put in standard contracts clauses, which are in the so-called 'black-list' or have already been declared abusive. On the other hand, however, Member States may also attribute to National Authorities the power to ascertain whether a specific clause is unfair and, consequently, to sanction the violation. Moreover, the strengthening and modernization of class actions,[75] which may be filed before jurisdictional as well as administrative Courts, also need to be credited.[76]

The Directive sets forth class action proceedings able to provide injunctive and compensatory relief, so to protect consumers interests in fields where the unfair conduct of market players could cause them more harm. Reliefs encompass injunctive measures, requiring the addresses to perform or to refrain from performing a specific action, as well as specific performance or *restitutio in integrum*, if the circumstances so allow. The variety of remedies discourages moral hazard by companies as they are able to tackle many different situations, skimming off profits from unfair and abusive conducts.

VI. Conclusive Remarks

In light of the above, it emerges how the new Directive insists on the importance of a coordinated private and public enforcement of consumer protection and devotes a great deal of attention to the modernisation of the discipline concerning the digital market, which currently appears at least laconic.[77]

[75] The importance of class actions in unfair commercial practices has already emerged in the so-called *Dieselgate* case, which pointed out the importance for European Union Law of a collective remedy against the harms caused by widespread unfair commercial practices. See A. Cilento, n 60 above, 1199.

[76] See R. Caponi, 'Ultime dall'Europa sull'azione di classe (con sguardo finale sugli Stati uniti e il "Dieselgate")' *Il Foro italiano*, 332-340 (2019); A. Palmieri, 'Perdite seriali dei consumatori e tutela collettiva risarcitoria: dove si dirige l'Europa?' *Il Foro italiano*, 205-210 (2018); L. Serafinelli, 'Ancora sulla tutela del consumatore, anche in forma collettiva' *Nuova giurisprudenza Civile commentata*, II, 612-620 (2019) concerning the introduction of a European framework for class actions.

[77] See, U. Von Der Leyen, *A Union that Strives for More. My Agenda for Europe*, Political

It is undeniable that the Directive addresses a vast number of issues, showing to be aware of the many implications deriving from misleading advertisement and unfair commercial practices, which are not merely confined to e-commerce. More precisely, the Directive gives a clear picture of what should be considered as an unfair commercial practice, and further describes what is misleading information and provides for stricter controls on social networks and online service providers, such as the duty to disclose commercial and advertisement purposes when providing their services. In this direction, the Directive extends its scope so to encompass also digital services, apparently rendered on a free of charge basis.[78] The Commission's new deal announced that consumers would enjoy rights, in respect of pre-contractual information and withdrawal, both in the case of paid and free online services. This promise has been transposed into the Directive, providing that it will be applicable to online trade in all circumstances when a monetary price is paid, or even when personal data are exchanged in return of online services. The idea of 'price' is thus widened, so to include also personal data.[79]

Some doubts about profiling of consumers are legitimate. Recital 45 provides that 'traders may personalise the price of their offers for specific consumers or specific categories of consumer based on automated decision-making and profiling of consumer behaviour allowing traders to assess the consumer's purchasing power'. In such a case, 'consumers should therefore be clearly informed when the price presented to them is personalised on the basis of automated decision-making, so that they can take into account the potential risks in their purchasing decision'. Compelling questions may arise should price be personalised based on grounds other than a consumer's purchasing power, such as nationality, religion or sexual orientation. Only time will show what the future may bring.

In general, the Directive is expected to deliver a better and more modern discipline of consumer law, repressing unfair commercial practices and improving the effectiveness of the right deriving therefrom. This is true even if the previous paragraph has shown that something more could be done in terms of private enforcement.

That being said, there are other parts of the European Union law that remain obscure. It is, for instance, hard to explain why some countries prohibit to put terms and conditions in online transactions, making it impossible for big corporations like Facebook, Google, Amazon or TripAdvisor to define a common

Guidelines for the next European Commission 2019-2024, 2019, 13, available at https://tinyurl.com/5fyjmypa (last visited 30 June 2021). The newly appointed President of the European Commission stresses the importance of Digital Markets of the purposes of her political agenda.

[78] See, A. De Franceschi, *La vendita con elementi digitali* (Napoli: Edizioni Scientifiche Italiane, 2020), 15.

[79] Art 4(2)(b), European Parliament and Council Directive 2019/2161/UE n 6 above.

approach valid throughout the European Union. A better harmonisation of the contractual and remedial framework of consumer protection is therefore due.

With all probability, the same directive on unfair clause after being into force for more than twenty-five years needs to be revised and updated.[80] In the meanwhile, national lawmakers and legal professionals need to provide consumers with effective individual remedies and, possibly, enhance the synergy between public and private enforcement. In other words, decisions by National Competition Authorities should be recognised as binding in civil proceedings files by consumers, harmed by unfair commercial practices. After all, this is what Art 9 of Directive 104/2014/EU provides in respect of infringements of competition law provisions by stating that 'Member States shall ensure that an infringement of competition law found by a final decision of a national competition authority or by a review court is deemed to be irrefutably established for the purposes of an action for damages brought before their national courts'.[81] Furthermore, due to the fact that many online commercial practices produce their effects in more than one Member State, it is necessary to consider decisions rendered by Competition Authorities of other States, and the Commission itself, at least as a *prima facie* evidence of the fact that an infringement of competition law has occurred. This would immensely help consumers in fulfilling the burden of the proof and would probably also contribute to the harmonisation of the procedural framework of consumer law among Member States.

[80] See, M. Loos, n 60 above, 423, who expresses a similar point of view.

[81] Reference is made to Art 9 European Parliament and of the Council Directive 2014/104/EU of 26 November 2014 on certain rules governing actions for damages under national law for infringements of the competition law provisions of the Member States and of the European Union [2014] OJ L 349/1, transposed into Italy by Art 7 decreto legislativo 19 January 2017 no 3. See, *si vis*, M. Zarro, 'La tutela risarcitoria da danno antitrust: nuovi sviluppi per il sistema misto di enforcement' *Rivista di diritto dell'impresa*, 669, 657-677 (2017); G. Villa, 'L'attuazione della Direttiva sul risarcimento del danno per violazione delle norme sulla concorrenza' *Corriere giuridico*, 445, 441-449 (2017); G. Bruzzone and A. Saija, '"Private e public enforcement" dopo il recepimento della direttiva. Più di un aggiustamento al margine?' *Mercato Concorrenza Regole*, 29, 9-36 (2017); P. Fabbio, 'Note sull'efficacia nel giudizio civile delle decisioni delle Autorità della concorrenza nazionali dopo il "Decreto enforcement" (d.lgs. 19 gennaio 2017, n. 3)' *Analisi Giuridica dell'Economia* 367, 367-390 (2017); L. Miccoli, 'Tra "public and private enforcement": il valore probatorio dei provvedimenti dell'AGCM alla luce della nuova Direttiva 104/14 e del d.lg. 3/2017' *Jucivile*, 348-368 (2017); E.A. Raffaelli and A. Croci, 'La prova nel private antitrust enforcement', in M.C. Malaguti et al eds, *Politiche antitrust ieri, oggi e domani* (Torino: Giappichelli, 2017), 153; G. Alpa, *Illecito e danno antitrust. Casi e materiali* (Torino: Giappichelli, 2016), 5; R. Chieppa, 'Il recepimento in Italia della Dir. 2014/104/UE e la prospettiva dell'AGCM' *Il diritto industriale*, 319, 314-321 (2016); F. Pasquarelli, 'Da prova privilegiata a prova vincolante: il valore probatorio del provvedimento dell'AGCM a seguito della direttiva 2014/104/UE' *Il diritto industriale*, 252-264 (2016). See also for more detailed analysis of the relationship between Independent Authorities and Judges, M. Angelone, 'Giudici e Autorità indipendenti: concorrenza e sinergia tra rimedi' *Rassegna di diritto civile*, 403-424 (2020).

Early Repayment of Loans Under EU Law: The *Lexitor* Judgment

Enrico Baffi* and Francesco Parisi**

Abstract

Recent changes in EU law provide flexibility and protection to consumers, facilitating early repayment of loans, when the consumer is no longer interested in continuing a credit relationship. From an economic point of view, early repayment of loans should be facilitated, because it allows money that is no longer needed to be put to other desirable uses. Under current EU law, as recently interpreted in the *Lexitor* judgment by the Court of Justice of the European Union, upon early repayment of a loan, consumers can obtain a pro-rated reimbursement of all the up-front and recurring costs of the loan, including origination fees and ancillary service costs. In this brief article, we take a critical look at the current EU approach to this issue, suggesting that, while the legislature made the pragmatic choice in permitting partial reimbursement of up-front costs, this leads to several economic inefficiencies that can ultimately hurt the consumer. Repayment of up-front costs, and of any other sunk cost associated with the creation of the loan, can lead to a suboptimal mix of lending contracts. Some consumers could, in fact, take out a lower interest rate long-term credit, even though they may only need a short-term loan, and this would create a disadvantage for the overall class of consumers. In order to actually protect the economic interest of consumers and carry out the intent of the legislature, we conclude that the up-front costs for non-mandatory ancillary services (such as brokerage fees, etc.) should not be included in the costs that are eligible for reimbursement in the event of early repayment of the loan. By excluding these costs, consumers will not be incentivized to overconsume such services, minimizing the negative externalities imposed on other consumers.

I. Introduction

Under current EU law, consumers have the right to fully or partially repay loans early and are entitled to obtain a pro-rated reimbursement of all the fixed and recurring costs of the loan. As currently applied, this pro-rated reimbursement includes origination fees, brokerage and legal fees, and all other fixed costs that are associated with the creation of the loan (hereinafter, we shall refer to this category of costs as 'up-front costs'). But the question remains whether allowing for the reimbursement of up-front costs in the event of early loan repayment is

* Professor of Law and Economics, Guglielmo Marconi University, Department of Economics.
** Oppenheimer Wolff & Donnelly Professor of Law, University of Minnesota, School of Law, and Professor of Economics, University of Bologna, Department of Economics.

economically efficient or even intended by the European legislature.

The rules governing early repayment of loans have emerged gradually in the EU over the last several years. The European legislature recently issued Directive 2008/48/EC in an effort create 'a well-functioning internal market in consumer credit'.[1] As part of its goal to establish a well-functioning market, the European legislature granted protection to consumers who wished to 'discharge fully or partially their obligations arising from (a) credit agreement before the due date'.[2] In other words, the Directive granted the consumer the right to repay loans fully or partially at any time. It additionally provided that consumers are 'entitled to a reduction in the total cost of the credit, such reduction consisting of the interest and the costs for the remaining duration of the contract'.[3] The total cost of the credit, as defined in Art 3(g) of the Directive, includes 'interest, commissions, taxes and any other kind of fees which the consumer is required to pay in connection with the credit agreement'.[4] The total cost of credit also includes any fees for ancillary services if 'the conclusion of (that) service contract is compulsory in order to obtain the credit or to obtain it on the terms and conditions marketed'.[5]

In the well-known *Lexitor* case,[6] the Court of Justice of the European Union (CJEU) interpreted these provisions to mean that consumers have the right to receive a pro-rated reimbursement of both the fixed, up-front costs and the recurring costs[7] of a loan in the event of early repayment. The decision was met with criticism and reluctance to comply by credit institutions.[8]

The Advocate General in his opinion, and the CJEU in its judgment, asserted that the choice made by the European lawmaker to reimburse both up-front costs and recurring costs was a pragmatic choice rather than an ideal one.[9]

[1] European Parliament and of the Council Directive 2008/48/EC of 23 April 2008, on credit agreements for consumers and repealing Council Directive 87/102/EEC [2008] OJ L 133/66.

[2] European Parliament and of the Council Directive 2008/48/EC, n 1 above, Art 16(1).

[3] European Parliament and of the Council Directive 2008/48/EC, n 1 above, Art 16(1).

[4] European Parliament and of the Council Directive 2008/48/EC, n 1 above, Art 3(g).

[5] European Parliament and of the Council Directive 2008/48/EC, n 1 above, Art 3(g).

[6] Case C-383/18, *Lexitor* v *Santander Consumer Bank S.A.*, Judgment of 11 September 2019, available at www.eur-lex.europa.eu.

[7] Recurring costs are the costs incurred by the lender while the loan is in progress. In other words, they are the costs that 'depend objectively on the duration of the contract'. Case C-383/18 n 6 above, para 24.

[8] See, eg, Intesa Sanpaolo, 'Information on risks and relative hedging policies', available at https://tinyurl.com/t262ujx6 (last visited 30 June 2021) (stating that it believed the Italian banking law Article 125-*sexies* could not be interpreted in a manner that complies with the *Lexitor* ruling and that while it would comply, it 'reserves the right to reconsider this operational stance in light of future developments'); Prawo, 'Polish banks are disregarding the Court of Justice of the European Union judgment regarding loans; Civil Rights Ombusman set to intervene', available at https://tinyurl.com/54nvvyzs (last visited 30 June 2021) (discussing that some Polish banks 'are proving loath to refund all the costs incurred by them' despite the CJEU's decision).

[9] Case C-383/18, n 6 above, paras 33-34, 55, 63-65, 68. In para 53 of his Opinion, the Advocate General stated that the rule establishing that up-front costs are not reimbursable while

First, there was the practical problem of distinguishing between up-front and recurring costs.[10] Second, there was a risk that if it established that up-front costs were non-reimbursable but recurring costs were reimbursable, credit institutions could use their discretion in invoicing costs to increase up-front costs to artificially lower the amount of reimbursable costs.[11] In other words, the credit institution could categorize a recurring cost as an up-front cost and charge the consumer when setting up the loan to avoid any potential reimbursement in case of early repayment. Because of these issues, the Advocate General and the CJEU concluded that the European legislator chose a rule that is easier to apply, under which all the up-front and recurring costs of the credit, as defined by the European Directive, are reimbursable on a pro-rated basis.[12] But as discussed in this article, allowing for the reimbursement of the up-front costs of a loan can create several inefficiencies, hurting both the consumer and the market for credit. While these inefficiencies exist because of the pragmatic choice of the legislature, the inefficiencies of this rule can be limited by making non-mandatory ancillary services — those voluntarily purchased by the consumer — ineligible for reimbursement. This conclusion is supported both by the intent of the legislature as well as by considering the economic incentives.

The remainder of this article is structured as follows. In Section II, we discuss the economic incentives created by the EU rules regarding early repayment of loans, providing special focus on whether granting consumers the right to early repayment is economically efficient and whether mandating partial reimbursement of up-front costs, as interpreted by the CJEU, is economically efficient. In Section III, we consider the question of whether consumers should be entitled to receive a reimbursement of the cost of third-party services that were not required for obtaining the credit, ie, non-mandatory ancillary services, such as the cost of financial advice and intermediation or additional insurance. We analyze both the intent of the European legislature in drafting Directive 2008/48/EC as well as examine the economic incentives under such a rule. In Section IV, we provide concluding thoughts on whether Directive 2008/48/EC and the *Lexitor* judgment should be interpreted to allow partial reimbursement

costs dependent on the duration of the loan are reimbursable may 'appear (…) at first sight to be relatively simple and therefore interesting, (but) its practical application will probably give rise to considerable difficulties of a practical nature. Indeed, as highlighted by the referring court in its request, credit institutions rarely specify which of the costs they incur are covered by the costs charged to consumers and, even when this occurs, the consumer would be entitled to dispute the accuracy of such specification'. The Advocate General added that maintaining the distinction between up-front costs and costs dependent on the duration of the loan may be impractical because 'in the event of a dispute over the amount of the reduction to which the consumer is entitled in the event of early repayment, national courts (would) have to call on the services of accounting experts, even if, by their nature, the costs in question are relatively modest'. Op. A.G., Case C-383/18, *Lexitor*, paras 53, 55.

[10] Case C-383/18, n 6 above, paras 33, 53, 55.

[11] ibid paras 31, 54.

[12] ibid paras 37, 66, 68.

of voluntarily purchased services in the event of early repayment. In sum, we conclude that the European legislature did not intend for non-mandatory ancillary services to be partially reimbursable in the event of early repayment, and the economic incentives support this conclusion.

II. The Law and Economics of Early Loan Repayments

The European Directive 2008/48/EC gave consumers the right to repay a loan early and to receive a pro-rated reimbursement of the 'total cost of the credit'. The Lexitor judgment interpreted this to conclude that in the event of early repayment of a loan, the lending credit institution must provide a pro-rated reimbursement of both recurring and up-front costs to the consumer. But allowing for the reimbursement of up-front costs creates several inefficiencies in the market for credit. While this choice may not be the ideal choice, as described by the Advocate General and the CJEU, it was the pragmatic choice.

1. The Pragmatic Choice of the European Legislature in Enacting Directive 2008/48/EC

Under Directive 2008/48/EC, in an aim to provide flexibility and protection to consumers, the European legislature granted consumers a right to early repayment of loans when the consumer is no longer interested in continuing a credit relationship. From an economic point of view, this may be considered efficient by allowing the capital to be put to more desirable uses. But if such a clause were truly efficient, then one would expect it to be included in every loan contract. Yet the legislature granted this right to overcome persistent information asymmetries in the market for lending. Additionally, the CJEU, in the *Lexitor* judgment, concluded that Directive 2008/48/EC granted consumers the right to partial reimbursement of both recurring and up-front costs of credit in the event of early repayment. While allowing for partial reimbursement of up-front costs can alter consumer decisions leading to market inefficiencies, the European legislature made the pragmatic choice and granted these rights to consumers to overcome information asymmetries.

a) Granting Consumers the Right to Repay Loans Early

In Directive 2008/48/EU the European legislature granted consumers the right to repay loans early. The legislature sought to lay down rules for a world where there are consumers who obtain long-term loans but who, due to events that were unforeseen ex ante, no longer need to keep the loans until their natural maturity. In an ideal world with perfect information, both the consumer and the credit institution would be able to accurately determine in advance the optimal duration of the credit relationship. Then the need for an early repayment

provision, as mandated under Directive 2008/48/EC, would not be necessary in a contract. However, because it is often unfeasible to know ex ante the ideal duration of the credit relationship, European legislators sought to protect consumers entering into contracts with imperfect information.[13]

Imagining a world with these consumers, the European legislator sought to protect consumers by granting them the inalienable right to make early repayment of loans, allowing them to discharge fully or partially their obligations under the credit agreement at any time.[14] This mandatory rule may be considered efficient as it means that when a loan is repaid in advance by a consumer who no longer needs it, the money can be put to more economically productive uses. But if it were truly efficient, then one would expect this clause to be included in every lending contract.

In a transaction between a rational, perfectly informed consumer and an equally rational, perfectly informed lender, all the contractual terms will be efficient. Thus, a rule allowing the consumer to make early repayment would be instinctively included in the parties' contract, without being imposed as a binding requirement by the legal system. However, the need for a mandatory rule allowing early repayment of loans, as included in Art 16(1) of the Directive, arises from the fact that the consumer may not be perfectly informed. If the consumer is imperfectly informed, the lender could include an inefficient clause in the contract, unbeknown to the consumer, prohibiting the consumer from paying back the loan early. The appearance of inefficient clauses in consumer contracts, permitted by consumers' imperfect information, is commonplace. Thus, the legislature's choice to grant consumers the right to early repayment can be said to be efficient under such conditions, helping to protect consumers from information asymmetries.

b) Directive 2008/48/EC Interpreted to Allow for the Partial Reimbursement of Up-Front Costs of Loans

However, in the *Lexitor* judgment, the CJEU concluded that under Directive 2008/48/EC consumers also have the right to a partial reimbursement of up-front costs in the event of early repayment of loans. If granting consumers the right to early loan repayment, as laid down by Directive 2008/48/EC, can be regarded as efficient in a world where consumers take out loans without knowing whether unforeseeable events will counteract the advantage of keeping the loan until its natural maturity, it is worth verifying whether the interpretive rule laid down by the *Lexitor* judgment on the reimbursability of up-front costs is also efficient in a consumer world such as the one described above, specifically when there is an information asymmetry.[15]

[13] Information is imperfect when it is not complete, ie, full.

[14] European Parliament and of the Council Directive 2008/48/EC, n 1 above, Art 16(1).

[15] A situation of 'asymmetric information' arises when one party to a potential agreement has

It has to be said that the reimbursability of up-front costs produces an expected benefit for the consumer. In the event of early repayment of the loan, the consumer receives a sum of money that would not otherwise have been returned. However, this rule also imposes an expected cost on the lender who, when determining the fee for the service of providing the loan, must take into account all the costs that will be incurred, including the expected costs of fee reimbursement.[16] The consumer will then have to weigh his potential benefit against the fee increase to obtain the loan that will certainly be imposed by lenders under the *Lexitor* rule that up-front costs are partially reimbursable.

If the increase in fee imposed by lenders to provide a loan is higher than the value that the consumer attributes to an early repayment clause, the legislature can be said to have overprotected and ultimately damaged consumers. This would give rise to a heterogony of ends.[17] Suppose we have a lender who by definition is assumed to be risk neutral and a risk-averse consumer who decide to enter into a € 10,000 loan contract for ten years. Also assume that the loan origination costs are € 1,000. The lender incurs these costs, but immediately passes them on to the consumer. If the parties were able to freely negotiate the clauses concerning early repayment of the loan, the question is whether they would include a clause allowing for the pro-rated reimbursement of the up-front costs incurred to set up the loan. The up-front costs are sunk costs, meaning that they are incurred for the purpose of setting up that specific loan and have no potential use outside that relationship.

Imagine that the parties initially agreed on pro-rated reimbursement in the event of early repayment of the loan. Suppose that, given the expectation of reimbursement, there is a fifty per cent probability that the consumer will discharge the loan after five years and a fifty per cent probability that he will keep the loan until its natural maturity. In this case the lender faces a risk equal to ½ (–500) and ½ (0), with an expected monetary value equal to –€250. As the lender is risk neutral, the certainty equivalent of this risk is –€250.[18] Hence,

greater knowledge than the other party.

[16] Early repayment rules also seek to protect the credit institution, as it could enter into a credit relationship on the basis of asymmetric information in favor of the consumer. The information asymmetry arises from the fact that the credit institution does not know which consumers wish to have a sum of money at their disposal for a short period of time, repaying the loan early, and which for a long period of time. A consumer interested in having a sum of money at his disposal for a short period of time may prefer a long-term loan, with the intention of repaying it early, rather than obtaining a short term-loan. Because of the partial reimbursement of up-front costs, the long-term loan may ultimately be less expensive to the consumer but imposes an additional cost on the lender. The lender will raise the cost to obtain a loan to offset this potential loss.

[17] Heterogony of ends is the idea that seeking a certain end goal can cause experiences that modify the original motivation seeking out that goal. In other words, 'the end does not always produce the means, but the means oftener originate the end'. G. Villa, *Contemporary Psychology* (New York: MacMillan Co, 1903), 366-369.

[18] The certainty equivalent in this context is the value required to leave the lender indifferent between the consumer paying the loan after five years and the consumer paying the loan after its

the lender will request a fee increase of €250 from the consumer. For his part, the consumer faces an expected risk of ½ (500) and ½ (0), with an expected monetary value equal to €250. As the consumer is risk-averse, for him the certainty equivalent of this risk is lower than €250.[19] For the consumer, the right to obtain partial reimbursement of the up-front costs of the loan would be inefficient, as the consumer would be forced to pay more for the loan than the consumer would benefit due to the lender's risk that the consumer pays back the loan early. In this case, the consumer would not acquire the loan and the credit market would shrink. Thus, imposing the right to repay loans early and the rule of partial reimbursement, as under the *Lexitor* judgment, while seeking to benefit consumers, could actually end up harming consumers. But because of information asymmetries and possible evasive action by the credit institution, as detailed by the Advocate General and the CJEU, the *Lexitor* rule is the pragmatic choice.

While granting consumers the right to repay loans early and allowing for the partial reimbursement of up-front costs could end up harming consumers and thus be inefficient, so far it has been implicitly assumed that the rule that only recurring costs must be reimbursed is the 'ideal', or efficient, rule. However, only allowing for the reimbursement of recurring costs would give rise to evasive behavior by credit institutions. Thus, the legislature, as affirmed in the *Lexitor* judgment, made the pragmatic choice and granted consumers the right to partial reimbursement of both recurring costs and up-front costs to overcome information asymmetries in the market.

The European Court of Justice in its *Lexitor* judgment reminds that in practice a rule only allowing for the partial reimbursement of recurring costs (to the exclusion of the up-front costs) gives rise to evasive behavior, whereby lenders do not return the recurring costs in the event of early repayment of the loan because they artificially increase the up-front cost items and likewise artificially decrease the recurring cost items.[20]

Up-front costs are those costs that the lender incurs in processing the loan and preparing the contract. They cease to apply when the contract is finalized. In the language of the *Lexitor* judgment, up-front costs are those that 'do not depend objectively on the duration of the contract'.[21] Up-front costs are usually incurred by the lender but are then reimbursed by the consumer when the loan contract is concluded.[22] In contrast, recurring costs are defined as costs incurred by the lender while the loan is in progress, for example, management costs. In other

natural maturity, in this case €250.

[19] In other words, because the consumer is risk-averse, he is not indifferent to paying €250 to the lender to allow him to either pay the loan after five years or pay the loan after its natural maturity. The consumer would be indifferent only if he pays the lender less than that amount.

[20] Case C-383/18, n 6 above, para 31.

[21] ibid para 24.

[22] One form of reimbursement of the up-front costs may be paid by deducting these costs from the borrowed sum.

words, they are costs that 'depend objectively on the duration of the contract'.[23] As a rule, these costs are incurred periodically throughout the life of the loan at a constant amount.[24]

To establish whether allowing for the partial reimbursement of up-front costs would give rise to evasive behavior, it is first necessary to verify whether the result described by the European Court of Justice would actually happen in a competitive market and then to verify whether such a result would give rise to inefficiencies.[25]

Suppose that a ten-year loan contract entails up-front costs of €500 and recurring costs of €500. Following the reasoning developed by the European Court of Justice and taking it to its extreme, we can assume that lenders will charge up-front costs equal to €1,000 and recurring costs equal to zero.

Imagine that there are two competitor companies that do not collude, and that have adopted the same strategy. As in the previous example, suppose there is a ½ probability that the consumer will repay the loan early after five years.

So, the provision that the recurring costs are equal to zero (in other words that the recurring costs will not be reimbursed) will result in an expected benefit for the lender equal to €250. The same provision, however, will result in an expected loss for the consumer, as he will take a ½ risk of losing €500, when he repays early. The choice facing the consumer therefore has an expected monetary value equal to ½ (0) and ½ (–500), ie, – €250. However, it has been assumed that the consumer is risk averse (since he is unable to diversify the risk across a portfolio of loans); therefore, the expected loss will be higher than –

[23] Case C-383/18 n 6 above, para 24.

[24] Recurring costs are usually prepaid to the lender by the consumer by means of deduction from the borrowed sum.

[25] It could be argued that the European legislature was not concerned by the effect that establishing the reimbursability of the up-front costs might have, and namely that consumers interested in obtaining a short-term loan would request a long-term loan with the intention of repaying it early and therefore obtaining reimbursement of the up-front costs. The judges in the Court of Justice of the European Union point out that the European legislature's choice was pragmatic and not ideal, due to the various difficulties that would arise if a distinction was made between up-front and recurring costs. But if the European legislature's choice was pragmatic, this means that if it were not for the practical problems in distinguishing between up-front and recurring costs, its ideal choice would have been a different one. And that ideal choice would have been to make only the recurring costs reimbursable. This appears to be the most convincing reconstruction of the European legislature's intent. One might then ask why the ideal choice would be to make up-front costs non-reimbursable. And the answer inevitably lies in the need to prevent consumers interested in short-term loans from actually taking out long-term loans with the intention of repaying them early and obtaining reimbursement of the up-front costs. This conduct is harmful to credit institutions and indirectly to consumers who take out long-term loans with the intention of maintaining them until their due date, as the cost of the early reimbursements will be passed on through the installments these consumers are required to pay. Furthermore, the reimbursability of up-front costs gives rise to considerable inefficiencies, as will be seen in this article. So, if the European legislature's ideal choice would have been to make up-front costs non-reimbursable, it is necessary to go back to it when there are valid reasons, as this would discourage the taking out of long-term loans with the intention of repaying them early.

€250. Assume, for the sake of this example, it is – €400.

Where we have two competing lenders and perfectly informed consumers, the contractual provision setting up-front costs equal to €1,000 and recurring costs to zero will not survive. In fact, one of the two competitors could offer the consumer the correct breakdown of up-fronts costs equal to €500 and recurring costs equal to €500. With this contractual provision, the recurring costs will have to be partially reimbursed in the event of early repayment of the loan and therefore there will be an expected cost for the lender equal to – €250. The consumer on the other hand will obtain an expected benefit equal to €400.[26] If the lender were to offer the consumer an amendment of the contract in these terms in exchange for a fee increase equal to €300, for example, both parties would gain.[27] Therefore, a clause that would artificially qualify all the costs to be up-front, such that nothing would need to be reimbursed in the event of early repayment, would not be sustainable in a competitive market.

However, the efficient clause will only emerge if consumers are perfectly informed. In this specific case they would have to read the contracts accompanying the loan and understand their clauses. But acquiring knowledge of contracts involves a cost, and this cost may exceed the expected benefit. In this case, consumers will forgo reading and understanding the contract. There will be a form of rational apathy. In the presence of this consumer apathy, it will no longer be worthwhile for lenders to distinguish between up-front and recurring costs, but only to describe all costs as up-front costs. In that way, lenders will not have to reimburse any cost in case of early repayment of the loan, even if some costs are in fact depending on the duration of the contract.

Hence adverse selection will occur and the inefficient clause describing all the costs as up-fronts costs will prevail over the efficient clause making a truthful distinction between up-front costs and recurring costs. The inefficiency will lie in the fact that the clause describing all the costs as up-fronts costs will result in a benefit for the lender of €250 and a cost for the consumer of €400, with a loss of social welfare equal to €150 per contract.

In interpreting Directive 2008/48/EC, the CJEU concluded the European legislature opted for the pragmatic rather than ideal choice because of the practical difficulty in distinguishing between up-front and recurring costs. It may have also opted for this choice to avoid the risk that banks would artificially inflate up-front costs and reduce recurring costs in the presentation of their financial proposals to reduce their exposure to early repayment costs. However, this does not alter the fact that in the absence of these difficulties, the rule

[26] Because of the consumer's risk aversion, the benefit of receiving the loan with up-front and recurring costs separately, allowing for the potential of partial reimbursement, is greater than if the consumer were risk neutral.

[27] While the consumer must pay more for the loan, because of his risk aversion, he is willing to pay more to have the potential for reimbursement. In this example, the lender will benefit by € 50 and the consumer will benefit by €100.

affording greater protection to consumers and credit institutions should provide for proportional reimbursement of the recurring costs only, while keeping the up-front costs required for creating the debt intact. There would be no reason to set aside the ideal rule where the aforementioned pragmatic reasons no longer applied, but because of the information asymmetries, the legislature chose exactly that, opting for the pragmatic choice.

Because of information asymmetries, the European legislature established that consumers have the right to repay loans early and, as decided in *Lexitor*, the right to partial reimbursement of both up-front and recurring costs. While granting these rights helped overcome the information asymmetries present in lending, it creates several inefficiencies, changing the way consumers act in the credit market.

2. The Inefficiency of the Reimbursability of Up-Front Costs and Origination Fees

While Directive 2008/48/EC overcame information asymmetries in the right to repay loans early and obtain reimbursements of the costs of a loan, in allowing for the reimbursement of both recurring and up-front costs, the *Lexitor* judgment creates several inefficiencies in the market for credit. First, consumers may, instead of obtaining short-term credit, seek out long-term credit with the explicit intention to repay the loan early due to its lower final cost. Second, because of the lower final cost to obtain long-term credit, there may be an overconsumption of credit by consumers who value the loans at lower than their societal cost. While the Directive sought to provide protection to consumers, these inefficiencies may ultimately hurt them.

a) Consumers Fail to Internalize the Costs of a Loan Under Partial Reimbursement of Up-Front Costs

The *Lexitor* judgment granted consumers the right to a partial reimbursement of both recurring and up-front costs in the event of early repayment. But allowing the partial reimbursement of up-front costs in the event of early repayment of a loan means that a person entering into a long-term loan contract with the precise intention to repay it in the short term, and therefore recover the up-front costs, does not internalize (ie incorporate) all of the costs required to provide the specific service of granting the loan. As the costs of providing the service are not fully internalized, a person may well purchase this service even when the private benefit is lower than the social cost. This creates a negative externality, which the consumer fails to take into account, giving rise to an inefficiency.

In deciding whether to purchase a loan, a borrower will compare his own private benefit to his private cost, represented by the sum that will not be recovered when repaying the loan early. If the private benefit is higher than the

private cost, the borrower will purchase the loan. But in some instances, the private benefit is lower than the social cost, ie, the sum of the private cost that the borrower incurs and the private cost that the lender incurs (represented by the cost to be reimbursed). In such instances, when the private benefit is lower than the social cost, the rule providing partial reimbursement of up-front costs, as decided by the *Lexitor* judgment, produces an inefficiency.

To help determine how consumers would behave in a world where up-fronts costs must be partially reimbursed in the event of early repayment of loans, we can first examine the opposite scenario — a situation where there is no requirement to partially reimburse up-front costs in the event of early repayment. This world accurately reflects many European legal systems prior to the *Lexitor* judgment.[28]

Suppose that consumers can be divided into those seeking a short-term loan (one year) and those seeking a long-term loan (ten years). Suppose also that the up-front costs for a short-term loan are slightly lower than for a long-

[28] This includes the Italian legal system. For example, Art 125-*sexies* of the TUB (the Italian Banking Law) provides that consumers are entitled to a reduction in the total cost of the credit in the event of early repayment. Decreto Legislativo 1 September 1993, no 385, Art 125-*sexies* (It). But prior to the *Lexitor* judgment, this provision was consistently interpreted as meaning that up-front costs did not have to be partially reimbursed. This was the opinion expressed by both independent authorities and ordinary courts. After the *Lexitor* judgment, Italian law scholars have considered the issue whether that decision influences the interpretation of Italian law. Some scholars claimed that Art 125-*sexies* of the TUB should now be interpreted in the sense that all the costs must be reimbursed, both up-front and recurring. These scholars maintain that, since the formulation of the EU Directive and Italian law are quite identical, it's not possible to interpret Italian law in a way that is different from the interpretation given by CJEU to Art 16, of Directive 2008/48/EC. That opinion can be found in A. Dolmetta, 'Anticipata estinzione e "riduzione del costo totale del credito." Il caso della cessione del quinto' *Banca Borsa Titoli di Credito*, II, 639 (2019). The same idea is expressed by A. Tina, 'Il diritto del consumatore alla riduzione del costo totale del credito in caso di rimborso anticipato del finanziamento ex art. 125-sexies, primo comma, t.u.b. prime riflessioni a margine della sentenza della Corte di Giustizia dell'Unione europea' *Rivista di Diritto Bancario*, 155, 166 (2019).

A different idea is formulated by A. Zoppini, 'Gli effetti della sentenza Lexitor nell'ordinamento italiano' *Banca Borsa Titoli di Credito*, 1, 11 (2020), who states that, since the Italian regulatory system and the role of independent authority (Banca d'Italia) prevent lenders from manipulating upfront costs, Italian law should be interpreted so that only recurring costs must be reimbursed.

An important decision has been issued by Arbitro Bancario Finanziario, Collegio di Coordinamento, 17 December 2019 no 26525. It states that all costs must be reimbursed because the Italian law can be interpreted only in this way. Indeed, every Italian judge has a duty to interpret Italian law in a way that it results in a meaning that conforms to European law (obbligo di interpretazione conforme), and the formulation of Art 125-*sexies* of the TUB permits this interpretation.

At this moment the Supreme Court has not given a solution to the problem. Some judges have stated that Art 125-*sexies* of TUB can be interpreted in a way the is compatible with the *Lexitor* judgment while others have stated that it's not possible.

In the former sense, see Tribunale di Napoli 29 June 2020 no 4433, available at www.dejure.it; Tribunale di Torino 22 September 2020, available at www.dejure.it and Tribunale di Milano 3 November 2020, available at www.dejure.it.

In the latter sense, see Tribunale di Napoli 10 March 2020 no 2391, available at www.dejure.it; Tribunale di Mantova 30 June 2020, available at www.dejure.it and Tribunale di Roma 11 February 2021, available at www.dejure.it.

term loan. Let us assume that for a short-term loan the up-front costs are equal to €4,000, while for a long-term loan they are equal to €5,000. In the world that is being considered, that is, where the rule enshrined in the *Lexitor* judgment does not exist and up-front costs are not reimbursed, consumers seeking short-term financing will take out loans with a short-term maturity (one year), incurring up-front costs of €4,000. In contrast, consumers seeking long-term financing will take out loans with a ten-year duration, incurring up-front costs of €5,000. Both categories of consumers will internalize all of the up-front costs to determine which type of loan to take out, as no form of recovery is envisaged. This results in consumers self-separating.

But consumers will behave in a considerably different manner if it is established that up-front costs must be partially reimbursed in the event of early repayment of loans, as under the *Lexitor* rule. Suppose that the *pro rata temporis* principle[29] is applied to calculate the sum to be reimbursed for up-front costs. This means that a consumer who enters into a ten-year loan contract, initially paying €5,000 for the up-front costs incurred by the intermediary, will obtain a reimbursement of €4,500 if he repays the loan after one year.[30] In other words, the final cost of the loan would be equal to €500. Thus, operating under a rule imposing partial reimbursement of up-front costs means that all consumers seeking one year financing who, in the absence of this *Lexitor* reimbursement rule, would have taken out one-year loans, will instead find it advantageous to take out ten-year loans with the precise intention to repay them after one year, recovering a large portion of the up-front costs and significantly reducing their private costs of obtaining a loan. When taking out a long-term loan with the intention to repay it early reduces the cost of the loan for the consumer, every consumer would adopt this strategy, even though this would create higher overall costs of loan creation for the bank.[31]

The ability for consumers seeking one-year financing to take out a ten-year loan with the precise intention to repay it after one year arises from the information asymmetry between consumers and lenders. Indeed, when faced with a consumer requesting a ten-year loan, lenders have no way of knowing whether the consumer is someone who intends to repay the loan in advance or is someone who intends to

[29] The *pro rata temporis* principle means that the consumer is reimbursed at the proportional rate for the amount of time that the consumer keeps the loan for. For example, if the consumer repays a ten-year loan after one year, he is entitled to a reimbursement of nine-tenths of the costs.

[30] Since it is otherwise a ten-year loan and the consumer pays the loan back after one year, the consumer is entitled to a reimbursement of nine-tenths of the up-front costs of the loan, or €4,500.

[31] This switch between seeking short-term credit and long-term credit with the explicit intention to repay it in the short term is termed the 'switching effect'. The 'switching effect' will only occur if the up-front costs for ten-year loans are not much higher than those of one-year loans. For example, if the up-front costs for one-year loans are equal to €1000 while those for ten-year loans are equal to €12,000, this switching effect would not occur, as the cost to repay a ten-year loan early would cost €200 more than a one-year loan. Accordingly, consumers seeking one-year financing would enter into loan contracts with one-year maturity.

keep it until its natural maturity. If lenders could identify consumers intending to repay the loan after one year, they would refuse to enter into a long-term loan with them or would only offer them a loan with a one-year maturity.[32] However, given current legislative restrictions, there is no way for lenders to discern one type of consumer from another.

Thus, under the *Lexitor* rule, consumers may be incentivized to purchase long-term credit with the explicit intention to repay the loan early, despite only seeking short-term credit in the absence of such a rule. And because credit institutions cannot distinguish between consumers, they will need to increase the costs for consumers to obtain credit to compensate for the potential reimbursement of the up-front costs of loans. This creates an inefficiency in the market.

b) Partial Reimbursement of Up-Front Costs May Lead to Over-Consumption of Long-Term Credit

Until now the focus has been on the behavior of consumers who seek a one-year loan, but instead, under the *Lexitor* rule, take out a ten-year loan with the undisclosed intention to repay the loan at an earlier time, given the reimbursability of the fixed origination costs. However, requiring the partial reimbursement of the up-front costs can lead to an overconsumption of credit. The rule can incentivize consumers who otherwise would not obtain a loan (low-valuing consumers) to enter the market and purchase credit despite valuing it less than its social cost, creating a second inefficiency.

By creating a discrepancy between the private cost faced by the consumer and the actual total cost of loan origination faced by the bank, ie, the social cost, the *Lexitor* rule can encourage consumers who might not otherwise seek a loan to obtain one, leading to an overconsumption of lending. For example, let us again assume the up-front costs for a one-year loan equal €4,000. Consumers who choose a long-term loan instead of a short-term loan will accordingly value the loan to be at least €4,000.[33]

The introduction of the *Lexitor* rule allowing for the partial reimbursement of up-front costs will lead some consumers who value the loan less than €4,000 and might not otherwise take out a one-year loan, to purchase one. In a market without the rule of partial reimbursability, given the lower benefits obtained in the short term of a one-year loan, consumers would not justify incurring the high up-front costs. But under the *Lexitor* rule, because part of the high up-

[32] Due to asymmetric information, the credit institution cannot know whether the consumer is interested in having a sum of money at his disposal for a short period of time, an interest achieved by taking out a long-term loan with the intention of repaying it early and recovering the up-front costs, or whether the consumer is really interested in a long-term loan. If there was no asymmetric information, the credit institution would refuse the loan to the consumer intending to make early repayment, thus obliging him to take out a short-term loan.

[33] This assumes the long-term loan has higher up-front costs than the short-term loan.

front costs are reimbursable, consumers may justify taking out long-term loans with the intention to repay them early, lowering the private cost of obtaining a loan.[34] This solution will be adopted by consumers for whom the private benefit obtained from a loan is higher than the private cost they incur. In the example considered, this cost is equal to €500 (consumers pay €5,000 in up-front costs but recover €4,500). As a result, any consumer who obtains a benefit of more than €500 from a loan will request a long-term loan with the intention to repay it after one year. Accordingly, any new borrowers who value the one-year loan at between €500 and €4000 will enter into the market for long-term loans. But obtaining such a long-term loan imposes a social cost. This social cost is equal to the up-front costs of €5,000, which outweighs the private benefit of between €500 and €4,000. Here, a very clear inefficiency can be observed.

It must therefore be said that the introduction of the rule as enshrined in the *Lexitor* judgment into a world where the reimbursement of up-front costs is not required will create two negative consequences. First, consumers who would have otherwise taken out a one-year loan will now take out a ten-year loan with the intention to repay the loan early. Second, the rule will now incentivize consumers who would never have obtained a loan prior to this rule to enter the market because the private benefit will exceed the private costs after reimbursement, despite the high social cost. In other words, allowing for the partial reimbursement of up-front costs in addition to recurring costs, as under the *Lexitor* rule, will lower the costs to obtain a loan, leading to an overconsumption of credit by low-valuing consumers. Hence, a rule such as the one enshrined in the *Lexitor* judgment may produce at least these two inefficiencies in the credit market. However, while the *Lexitor* judgment may create these inefficiencies, the European legislature and the CJEU chose this rule because it was the pragmatic choice. While these inefficiencies may arise as part of consumer behavior, legislatures can choose to mitigate these inefficiencies by limiting what up-front costs are eligible for reimbursement. As described in the next section, by choosing to exclude from reimbursement the up-front costs for non-mandatory ancillary services, the legislatures can mitigate adverse consumer incentives in the credit market.

III. Reimbursability of Other Non-Mandatory Ancillary Services

As discussed in the previous sections, a rule imposing the mandatory partial reimbursement of up-front costs in addition to recurring costs can lead to several inefficiencies. Because consumers do not internalize the full costs when partially reimbursed in the event of early repayment, they may enter into long-term loans when only seeking short-term lending, or consumers may

[34] Again, this assumes that a long-term loan has higher up-front costs than a short-term loan.

obtain loans they might otherwise purchase, imposing a social cost on others. However, as detailed by the Advocate General and the CJEU, the European legislators chose the pragmatic option in allowing for reimbursement of the up-front costs.[35] One natural extension to the reimbursement of up-front costs is whether the European legislators also sought to allow for the reimbursement of up-front, non-mandatory third-party services, that is, those services voluntarily purchased by consumers in obtaining the loan. As detailed in this section, both the legislative intent and economic considerations lead to the conclusion that non-mandatory ancillary services should not be eligible for reimbursement in the event of early loan repayment. Excluding voluntarily incurred costs from eligibility for reimbursement in the event of early repayment helps to mitigate the inefficiencies in the credit market that arise under the *Lexitor* judgment. There might be some concern that this view cannot take into account the possibility that the formal 'non-compulsory' nature of additional costs may often times be a mere façade, over which the consumer has no control. Indeed, lenders could bundle these services. But it's worth noting that, according to Art 3 of the Directive 2008/48/EC, an ancillary service is 'non-compulsory' only if the conclusion of a service contract is not necessary in order to obtain the credit or '*to obtain it on the terms and conditions marketed*'.

1. Directive 2008/48/EC Excludes the Reimbursement of Non-Mandatory Services Offered by Third Parties in the Event of Early Repayment

The European legislature, in Directive 2008/48/EC, provided consumers with the right to repay a loan early and to receive a partial reimbursement of both the up-front and recurring costs of the loan. But there is an open question of whether the legislature intended this to include up-front costs of non-mandatory ancillary services, that is, services voluntarily purchased along with the loan, eg, insurance services. These costs should not be eligible for partial reimbursement in the event of early loan repayment. Excluding such costs from reimbursement carries out the intent of the legislature and mitigates the economic inefficiencies detailed previously, which would otherwise be present in the credit market.

Directive 2008/48/EC, as discussed earlier, states in part that consumers

> shall be entitled to a reduction in the total cost of the credit, such reduction consisting of the interest and the costs for the remaining duration of the contract.[36]

The total cost of the credit, as outlined in Art 3 *g)* includes

[35] Case C-383/18, n 6 above, paras 33-34, 53, 55, 63-65, 68.
[36] European Parliament and of the Council Directive 2008/48/EC, n 1 above, Art 16(1).

> interest, commissions, taxes and any other kind of fees' for ancillary services if 'the conclusion of (that) service contract is compulsory in order to obtain the credit or to obtain it on the terms and conditions marketed.[37]

A literal interpretation of the Directive 2008/48/EC suggests that, if the definition of reimbursable up-front costs includes the services that the credit institution prescribes as mandatory to obtain the credit, then it accordingly *does not* include the costs of other ancillary services offered by third parties that the credit institution does not prescribe as mandatory. According to the operative part of CJEU's judgment in *Lexitor*,

> Article 16(1) of Directive 2008/48/EC ... must be interpreted as meaning that the right of the consumer to a reduction in the total cost of the credit in the event of early repayment of the credit includes all the costs *imposed* on the consumer.[38]

An inattentive reading of the decision could lead to the view that the EU Court did not wish to introduce any distinction between costs, meaning that even the costs of non-mandatory services should be proportionally reimbursed along with all other up-front costs. However, the operative part of the judgment must be interpreted carefully and in light of the statement of reasons for the ruling.

Firstly, the operative part makes a clear reference to the costs 'imposed' on the consumer (ie, the costs for the services that the bank requires the consumer to purchase to obtain the credit), but not to those that the consumer bears voluntarily (ie, non-mandatory ancillary services purchased from third parties).[39] In describing which costs are reimbursable, the explicit usage of the 'total cost of the credit' in para 23 of the CJEU judgment makes express reference to Art 3(g) of the Directive, under which the costs borne by consumers are proportionally reimbursable 'if ... the conclusion of a service contract is *compulsory* in order to obtain the credit or to obtain it on the terms and conditions marketed'.[40] In light of Art 3 *g)* of the Directive, the operative part of the *Lexitor* judgment must therefore be read as meaning that the cost of non-mandatory ancillary services must be excluded from the definition of 'total cost of the credit' because they are not compulsory to obtain credit. Therefore, they must not be eligible for reimbursement in the event of early repayment.

It is also necessary to bear in mind that, as expressly ruled by the CJEU, the Court's interpretation of the term 'total cost of the credit' was established by taking into account the fact that the consumer protection system, set up by the Directive, 'is based on the idea that the consumer is in a weak position vis-à-vis

[37] European Parliament and of the Council Directive 2008/48/EC, n 1 above, Art 3 *g)*.
[38] Case C-383/18, n 6 above, para 36 (emphasis added).
[39] ibid para 36.
[40] ibid para 5 (emphasis added).

the seller or supplier, as regards both his bargaining power and his level of knowledge'.[41] If we consider the objective pursued by the European legislature, namely that of protecting the consumer with regards to the ancillary services imposed by the credit institution, then the costs that the consumer decides independently to incur must not be reimbursed in the event of early repayment.

Thus, interpreting the operative part of the *Lexitor* judgment in the light of its statement of reasons, the consumer's entitlement to the reimbursement of the proportional cost of the credit in the event of early repayment of the credit includes all of the costs that the consumer *must* incur, but excludes those for ancillary services purchased by the consumer voluntarily, ie, those services not imposed on him by the credit institution to obtain the credit or to obtain it on the terms and conditions marketed.

Non-mandatory ancillary services include, with greater reason, those purchased independently by consumers and offered by third parties rather than directly by the credit institution.[42] For example, these services include the cost of financial advice, intermediation, and additional insurance that are not required by the credit institution, but the consumer chooses to purchase. According to the aforesaid legislation and statement of reasons in the *Lexitor* judgment, the cost of non-mandatory services must be excluded from the costs eligible for reimbursement by the credit institution in the event of early repayment of the loan.

While the operative part of the *Lexitor* judgment, interpreted in light of its statement of reasons, leads to the conclusion that non-mandatory ancillary services are not reimbursable in the event of early credit repayment, this interpretation also falls in line with the pragmatic considerations made by the Advocate General and upheld by the CJEU. One objection raised by the Advocate General noted that in the event of a dispute, the distinction between up-front costs and costs dependent on the duration of the loan, ie, recurring costs, would require national courts to call on the services of accounting experts.[43] However, the need to obtain outside experts is not necessary for understanding the costs of non-mandatory services offered by third parties. Because these services are, in fact, services that are not offered by the credit institution itself, but by other entities, the costs are fully independent of those of the credit institution, and therefore are easily distinguishable ex post.[44] Therefore, allowing for the reimbursement

[41] Case C-383/18, n 6 above, para 29 (citing C-377/14, *Radlinger and Radlingerová*, EU:C:2016:283, para 63).

[42] The services offered by third parties cannot be inflated or manipulated to the advantage of the credit institution, the conclusion of a service contract is compulsory in order to obtain the credit or to obtain it on the terms and conditions marketed and forcing the lenders to reimburse the cost of those services would impose an even greater externality on borrowers who do not exercise early repayment.

[43] Case C-383/18, n 6 above, para 54.

[44] Because the consumer will receive a defined price when choosing to purchase a non-mandatory ancillary service from third parties, an estimation of the cost is not needed when

of these types of costs was not within the scope of the Advocate General's or CJEU's intent.

Similarly, the Advocate General's and CJEU's concern that credit institutions could use their discretion in invoicing costs to increase up-front costs to the detriment of recurring costs,[45] when applied to non-mandatory ancillary services, is not applicable for two obvious reasons. First, because the lending intermediary has no direct control over the prices of services offered by third parties, it does not benefit from any increase in the price of these services aimed at creating a nominal reduction in the reimbursable costs linked to the duration of the loan. Second, unlike the mandatory services governed by Directive 2008/48/EC, if the price of the non-mandatory services offered by third parties were artificially increased, the consumer demand to purchase such services would fall, resulting in the opposite effect to the one intended by the credit institution.[46] Accordingly, these costs were outside the consideration of the legislature and should not be eligible for partial reimbursement.

Interpreting the *Lexitor* judgment in light of its pragmatic considerations, the European legislature did not intent for the partial reimbursement of non-mandatory ancillary services in the event of early loan repayment as these incurred costs are voluntary. The European legislature only sought to allow for the partial reimbursement of *compulsory* costs. This conclusion is further supported by economic considerations. As detailed in the next section, even if it appears that the European legislature's decision was designed to protect credit institutions—by making the cost of non-mandatory services non-reimbursable in the event of early withdrawal—the decision is a legislative choice that, above all, protects consumer welfare and fosters economic efficiency. After all, as stated in Recital 7 of the Directive, the aim of the European legislation was to 'facilitate the emergence of a well-functioning internal market in consumer credit'.[47]

2. Economic Inefficiency of the Reimbursement of Up-Front Costs for Non-Mandatory Services Offered by Third Parties

The exclusion of up-front costs that are voluntarily purchased by consumers from eligibility for partial reimbursement in the event of early repayment also

determining any reimbursement of costs. Accordingly, these services fall outside the scope of the Advocate General's concern, ie, non-mandatory ancillary services are not meant to be included in any potential reimbursement of costs.

[45] See above Section II, 1, a).

[46] The credit institution, if it sought to maximize profits, would actually achieve the opposite effect to the one intended. Supposing that the credit institution arranges with the third party to artificially increase the fee for ancillary services offered by the third party, at the same time reducing the nominal amount of the recurring costs borne by the credit institution, this could lead to a decrease in the purchase of the ancillary services offered by the third party, leaving the credit institution unable to recover the recurring costs.

[47] European Parliament and of the Council Directive 2008/48/EC, n 1 above, pmbl 7.

follows from economic considerations. The Advocate General and the CJEU, both in their statement of reasons, detailed that the European legislature chose the pragmatic rule rather than the ideal rule. If non-mandatory ancillary service costs were reimbursable, it would create several negative externalities. First, allowing for the reimbursement of voluntarily incurred costs would impose higher costs onto other consumers. Second, because the costs of providing loans may rise, the long-term credit market could shrink. Finally, if, in response to the allowing partial reimbursement of non-mandatory up-front costs credit institutions chose to not enter into contracts with consumers purchasing such services, the market for third party services would become inefficient.

a) Inefficiency Caused by Early Paying Consumers Passing Costs onto Other Consumers

The conclusion that under Directive 2008/48/EC, the costs for non-mandatory ancillary services purchased from third parties should not be reimbursed follows from a specific logic of efficiency. Specifically, it helps mitigate the inefficiencies in consumer behavior when overconsuming long-term loans as described in Section 2. The efficiency of this rule can be illustrated by the following example. For the sake of simplicity, imagine the case of a twenty-year credit relationship with a total value of €1,000, repayable in twenty installments of €50. Suppose that the cost of the non-mandatory ancillary services offered by third parties is equal to €100. Consider a credit institution with a pool of consumers, thirty percent of whom intend to exercise the right of early withdrawal after payment of the first installment. Also, suppose that the ancillary service has a full value for consumers who intend to maintain the credit until its natural expiry, and it has a reduced value, lower than the price of the service itself, for consumers who intend to make an early repayment. Consider the two alternative rules.

Rule 1. Non-reimbursability of the costs relating to non-mandatory services offered by third parties.

Under this first rule, any non-mandatory ancillary service offered by third parties will only be purchased by consumers who are interested in it and who value it more than the requested price. The cost of the credit relationship remains equal for all consumers, €1,000, with installments of €50. Furthermore, presuming the service only has a special value for consumers who intend to maintain the loan until its natural expiry, for example, an additional insurance service, consumers will make different choices. Those who intend to maintain the loan for its full duration will purchase the ancillary service, while those who intend to repay it early will not purchase the service. Since early paying consumers know they will not receive a partial reimbursement if repaying the loan early, there will be no benefit to purchasing the ancillary service but not fully utilizing it over the duration of the loan. Hence, there will be no inefficiencies. Only those who

actually value the service at its consumer price will purchase the service.

Rule 2. Reimbursability of the costs relating to non-mandatory services offered by third parties.

In contrast, under this second rule, the pool of consumers will not make different choices, leading to inefficiencies. Consumers who intend to exercise the right of early repayment will purchase the non-mandatory ancillary service, regardless of the value they attach to it, expecting that the price paid for the service will be proportionally reimbursed upon repayment. The non-mandatory service offered by third parties would therefore be purchased by consumers who value it less than its original cost in addition to the purchasers under Rule 1.

So, for example, suppose that the non-mandatory service consists of mediation. The cost of mediation is €100. Even though the benefit is of little value to a consumer who intends to pay off the twenty-year loan after the first installment, he will still purchase it, trusting that the cost of the mediation will be reimbursed upon early repayment. Hence, there will be inefficiencies.

The decision to make the up-front cost of non-mandatory services reimbursable, which at first appears to be an extension of legislation to protect consumers, would instead harm consumers who use the credit facility for its original duration. Because all consumers will purchase the non-mandatory service offered by third parties, the consumer cost of a loan will rise. Indeed, the credit institution will have to consider the proportional reimbursement of services purchased by consumers when those consumers exercise the right of early repayment. Imagine a pool of 100 customers with 30 customers exercising the right of early repayment. In the example considered above, the installments would increase from €50 to €51.99 for all contracting parties.[48] Those who exercise the right of early repayment, receiving proportional reimbursement of the optional services they purchased, would impose an unwanted cost (ie, a negative externality) on those who instead remain bound by the credit relationship for its entire duration.

In social terms, there will also be a waste of resources because consumers will use the third-party, non-mandatory service even though they value the service less than its cost.[49] The costs of most of the non-mandatory services, even if not reimbursed to the consumer, cannot be reversed or recovered from a social point of view. Because these costs are incurred up-front, overconsumption of the service

[48] In calculating the reimbursement made to consumers who made early repayment, the *pro rata temporis* principle was applied, under which the credit intermediary reimbursed €95 to each one. In a pool of one hundred customers, the sum reimbursed to the thirty customers who made early repayments is equal to €2,850. When the cost of these reimbursements is passed on ex ante to all 1430 remaining installments, each installment increases by €1.99.

[49] Consumers who value the service less than the cost will purchase the service if they plan to repay the loan early if they value the benefit of the service more than the amount of costs not reimbursed upon early repayment (essentially, if the lowered 'cost' of the service is below the consumer's benefit). As discussed previously under Rule 1, if the consumer would not receive partial reimbursement, they would not purchase the service at the full 'cost' because the benefit is lower.

can lead to costs on others. By overconsuming, those consumers who value the service less than its cost can keep other, higher-valuing consumers from obtaining the service. In other words, low-valuing consumers impose a social cost, leading to inefficiencies in the credit market.

As discussed, when a consumer who intends to make early repayment purchases a non-mandatory service, the cost of the service passes on to all consumers through an increase in the consideration. In the scenario discussed previously, this cost is equal to €100. If the value assigned to the non-mandatory ancillary service is lower than €100, the consumer's overall welfare will decrease. So, if a consumer intending to make early repayment assigns the non-mandatory ancillary service a value of €10, on account of it being passed on to consumers through an increase in the cost of €100, the welfare (in other words their wealth) of the other consumers who maintain the loans for the full term will decrease by €90. In aggregate terms, allowing for the reimbursability of the costs of third-party, non-mandatory ancillary services will normally reduce the consumers' welfare. Paradoxically, in terms of distribution, the rule will benefit consumers who repay the loan early, providing them a partial reimbursement, by passing on the cost of the reimbursement to those who remain bound by the original terms of the credit agreement. This creates an inefficient market for third-party services.

In sum, if partial reimbursements of the up-front costs of non-mandatory services purchased from third parties were allowed, it would create an inefficient market. Consumers who repay the loan would impose an unwanted cost on those who retain a loan for its natural duration in the form of higher installment costs. It would also impose a social cost by reducing overall social welfare, as consumers would purchase the service even though they value it less than its cost. Thus, to avoid this inefficiency, and as supported by the intent of the European legislature, partial reimbursements of up-front costs of non-mandatory ancillary services should not be permitted. This aligns the incentives of consumers and also mitigates the inefficiencies described in Section 2 that arise by imposing the requirement of partial reimbursement of at least some up-front costs when repaying a loan early.

b) Long-Term Credit Market Could Shrink Under Market Inefficiencies

Second, allowing for the partial reimbursement of the costs of non-mandatory ancillary services in more extreme cases could raise the problem previously addressed by the economists Joseph Stiglitz in his work published with Andrew Weiss, 'Credit Rationing in Markets with Imperfect Information',[50] resulting in

[50] J.E. Stiglitz and A. Weiss, 'Credit Rationing in Markets with Imperfect Information' 71 *American Economic Review*, 393-410 (1981).

a contraction of the long-term credit market. By excluding the costs of these services from possible reimbursement in the event of early repayment, consumers' incentives to purchase a long-term loan when seeking short-term lending will be reduced.

Consider the example used above of a credit intermediary with a pool of customers who have taken out long-term loans. The intermediary has set the requested installment, bearing in mind that a percentage of the customers will pay the debt off early, requiring reimbursement of the service costs. The intermediary will set a higher installment cost to offset the costs created by customers who opt for early repayment. However, this higher installment cost will mostly be borne by consumers who intend to maintain the long-term loan agreement. It will only have a marginal impact on consumers who intend to pay the loan off in the short term.[51] The increased installment cost will therefore discourage a higher portion of consumers from requesting the loan who intend to maintain the long-term loan agreement until its natural expiry. The withdrawal of these consumers from the pool of consumers will cause the percentage of customers intending to make early repayment to increase. In response, the credit intermediary will have to deal with a higher percentage of early repayments than anticipated and will have to set an even higher installment price for subsequent agreements. This new installment price will compel even more potential customers intending to maintain the long-term credit until its natural expiry not to request the loan. In this way, the long-term consumer credit market will gradually shrink. In essence, long-term consumers would be penalized by externalities created by consumers who pay loans off early, creating a problematic imbalance in the capital market. This would create an inefficient market by shrinking the availability of long-term credit, all driven by consumers obtaining long-term credit to satisfy short-term credit needs.

But, if as proposed here, the costs of non-mandatory ancillary services are excluded from possible reimbursement in the event of early repayment, the inefficiency of reimbursing up-front costs as required under the *Lexitor* judgment can be mitigated. Only those consumers who value a given service at or above its social cost will purchase it. Accordingly, any increase in installment cost necessary to offset the possible reimbursement of costs to customers who opt to repay the loan early will be minimized.

As seen by this discussion, allowing for partial reimbursement of the costs of non-mandatory ancillary services can create several inefficiencies. It can harm those consumers who intend to keep long-term loans until their natural expiry by increasing the costs of obtaining a loan. It could also lead to the market for long-term loans to shrink, again hurting consumers. To avoid these

[51] Because the price of every installment for all consumers will increase, those who pay off the loan early only pay a fraction of the overall increase in cost, while those who pay off the loan over all the original installments face the full increase in cost.

negative externalities, and as intended by the European legislature, the partial reimbursement of non-mandatory services purchased from third parties in the event of early repayment should not be permitted.

c) Lender's Refusal to Enter into the Contract Can Create an Inefficient Market for Third Party Services

Despite the preceding discussion, there is the possibility that these supposed inefficiencies could disappear as the result of evasive action on the part of lenders. Under a rule allowing for the partial reimbursement of non-mandatory ancillary services, lenders may well refuse to enter into loan contracts with consumers who have purchased these services because they may have to partially reimburse the costs in the event of early repayment of the loan. As a consequence, consumers would be incentivized to no longer purchase that specific ancillary service (consider, for example, ancillary services such as insurance policies). However, this evasive action would not eliminate all the inefficiencies caused by the *Lexitor* rule.

In fact, if on the one hand the ancillary service would no longer be purchased by consumers who value it less than its cost – which would increase efficiency – on the other hand, it would also result in those consumers who value it more than its cost to decide not to purchase it. In this case an inefficiency would remain. So, even if the ancillary service were to disappear from the market because lenders refuse to enter into loans with persons who purchased this service, an inefficiency would still exist as consumers who value the ancillary service more than its cost would forgo the service.

Accordingly, if the *Lexitor* judgment were to allow for partial reimbursement of costs of non-mandatory services offered by third parties, the lenders, by refusing to contract with consumers who purchase third-party services, would create an efficiency. If, on the other hand, as is described earlier, the *Lexitor* judgment *excludes* partial reimbursement of costs of these services, then this inefficiency will not be present, as consumers will only purchase the service if they realize the full benefit, and lenders will not refuse to enter loan contracts due to fear of having to repay part of the costs.

Thus, the inefficiencies of allowing for the partial reimbursement of non-mandatory services offered by third parties under the *Lexitor* judgment show that such an interpretation is economically undesirable. This supports the legislature's intent as described by the pragmatic considerations discussed by the Advocate General and the CJEU. By granting the right to early repayment and a pro-rated reimbursement in the event of early repayment of both up-front and recurring costs under Directive 2008/48/EC and the *Lexitor* judgment, the legislature overcomes the information asymmetries present in the lending market. But as discussed in this section, that can lead to its own inefficiencies. To mitigate these inefficiencies, the partial reimbursement of voluntarily purchased

services from third parties under the *Lexitor* judgment should not be permitted.

IV. Concluding Thoughts on the Interpretation of Directive 2008/48/EC

The legal and economic points made above lead to a convergent conclusion on how best to the interpret European Directive 2008/48/CE, considering the fact that the European legislature sought to foster the 'emergence of a well-functioning internal market in consumer credit'.[52] With regard to the economic treatment of consumers who wish to 'discharge fully or partially their obligations arising from the credit agreement before the due date', the favored interpretation must follow the provisions set forth in Art 3 *g)* of the Directive.[53] More specifically, a consumer exercising the right of early repayment will be entitled to a pro-rated reimbursement of the costs that the intermediary prescribes as mandatory to obtain the credit from the credit institution. However, this right must not extend to the cost of ancillary services supplied by third parties that are not required by the credit institution, such as, the cost for financial advice and mediation or for non-mandatory additional insurance.

The inclusion of the costs of non-mandatory ancillary services supplied by third parties among those eligible for reimbursement in the event of early repayment would lead to misuse of the long-term credit facility. It would also lead to the excessive purchase of non-mandatory ancillary services by those intending to exercise the right to make early repayment provided for by European legislation, as described previously.

The long-term credit facility would be misused as consumers wishing to have short-term availability of a sum of money will find it advantageous to enter into long-term loan agreements, envisaging early repayment. This would allow them to proportionally recover all the costs, including the up-front costs (in the scenario examined, the costs for services purchased voluntarily from third parties), making long-term loans cheaper than short-term loans.

The excessive use of non-mandatory ancillary services, coupled with the rule of early repayment with reimbursement, would lead to the creation of negative externalities by those using the credit in the short term to the detriment of those using the credit for its natural duration, in other words, the long term. Whenever an activity creates negative externalities, it will be pursued more than the socially optimal level. Indeed, it will even be implemented when the private benefit is lower than the social cost. Hence, there will be excessive purchasing of non-mandatory ancillary services offered by third parties. To avoid

[52] European Parliament and of the Council Directive 2008/48/EC, n 1 above, pmbl 7. Even the *Lexitor* judgment reminds, in para 4, that the Directive seeks to create the 'emergence of a well-functioning internal market in consumer credit'. Case C-383/18, n 6 above, para 4.

[53] European Parliament and of the Council Directive 2008/48/EC, n 1 above, Art 3 *g)*.

this and adhere to the European legislature's intent, the partial reimbursement of non-mandatory ancillary services should not be permitted under the *Lexitor* judgment.

V. Conclusion

Upon examining the incentives under the *Lexitor* rule, we can say that the introduction of a rule under which up-front costs must be partially reimbursed in the event of early repayment of loans will produce inefficiencies compared to the ideal world where only recurring costs must be returned. But the existence of this ideal world necessitates the impossibility of evasive behavior on the part of lenders. If evasive behavior can be implemented because consumers are not informed, then the state of the world that emerges with the rule of reimbursement of only recurring costs will be affected by inefficiencies. These inefficiencies will arise from the fact that the prevailing market clause will effectively establish that not even recurring costs must be returned.

So, it is a matter of choosing the second-best solution. The *Lexitor* judgment opted for the rule that even up-front costs must be reimbursed. The inefficiencies produced by this rule could be limited by making the up-front costs items that cannot be artificially increased by lenders non-reimbursable. In other words, non-mandatory ancillary services, those voluntarily purchased by the consumer, should not be eligible for partial reimbursement in the event of early repayment. This conclusion follows from both the intent of the legislature as well as the economic considerations, namely the inefficiencies that arise if reimbursement of the costs of these services were mandatory.

Nothing New Under the Digital Platform Revolution? The First Italian Decision Declaring the Employment Status of a Rider

Maurizio Falsone*

Abstract

In 2020, an Italian tribunal classified a food-delivery rider working via a digital platform as an employee for the first time. Italian courts and scholars have struggled with new, ambiguous legal notions with the aim of (re)shaping the border between subordination and self-employment. In this case, a Sicilian judge ruled that the working relationship between the digital platform's owner and the rider using it to acquire work is characterized by hetero-direction – the basic element of subordination pursuant to Art 2094 of the Italian Civil Code. This article examines the arguments regarding the nature of both the platform and the rider's working relationship with the provider. The article analyses how the judge concretely subsumed the case into the legal framework and underlines the weaknesses in the reasoning regarding the role of continuity in qualifying a working relationship. The article further demonstrates a problematic surplus of arguments and exegetic techniques that risk weakening one another.

I. Introduction: 'Nothing New Under the Sun'?

On 24 November 2020, Palermo Tribunal ruled on a case concerning a rider who worked with a well-known food delivery service's digital platform. The decision comes after other similar cases, which are already abundantly disputed in Italy[1]. Since the management of staff and work performance using apps and algorithms represents an innovation that is already global and potentially inter-sectorial, the decision equally rests within heterogeneous case law and foreign national and supra-national decisions.[2]

* Assistant Professor of Employment and Labour law, Ca' Foscari University of Venice. This article is part of the funded research project PRIN 2017EC9CPX 'Dis/Connection: Labor and Rights in the Internet Revolution', with the participation of scholars from the Universities of Bologna, Milano Statale, Napoli Federico II, Udine, and Venezia Ca' Foscari.

[1] Cf Tribunale di Torino 5 July 2018, *Rivista giuridica del lavoro*, II, 317 (2018), Corte d'Appello di Torino 4 February 2019 no 26, *Rivista italiana di diritto del lavoro*, II, 340 (2019) and Corte di Cassazione 24 January 2020 no 16633, *Lavoro Diritti Europa*, 1 (2020); Tribunale di Milano 10 September 2018, *Labor*, 1, 112 (2019).

[2] Cf the report 'Taken for a Ride: Litigating the Digital Platform Model' by the International Lawyers Assisting Workers Network (ILAW) available at https://tinyurl.com/4c4usvfj (last visited 30 June 2021).

These new digital tools have challenged the traditional legal patterns governing work contracts and relationships worldwide,[3] because they can potentially obscure the entrepreneur's command and control power while simultaneously leaving room for greater flexibility and freedom for both parties in the working relationship. Analysis of the case in question is particularly merited in view of the rider's qualification as an employee pursuant to Art 2094 of the Italian Civil Code (Art 2094 hereinafter), a norm that was issued in 1942, when digitalization lay beyond the bounds of imagination.[4] The case is not a unique case in the international context,[5] but it represents the first of its kind in Italy, where – to date – the lawmaker and the courts have strived to distinguish between gig/crowd economy workers and traditional employees, creating and interpreting new legal notions, such as the hetero-organization, the negotiated coordination and the digital platform itself.[6] From a legal perspective, this court case raises doubts as to whether these recent legislative and exegetic efforts are useful or whether digital work can be managed through the legal labour patterns of the last century. In actuality, examination of the phenomenon raises the question of whether these new digital tools give rise to different (less hierarchical) models of human resource management or whether they simply obscure and enhance older vertical business models.[7] It is impossible to definitively answer these questions, which require legal and managerial knowledge; however, analysis of the strengths and weakness of the Tribunal's reasoning and arguments can contribute to the discussions about these contradictory hypotheses.

[3] The main case law concerns the legal qualification of the working relationship; however, there are decisions about other specific issues, such as, discriminations (Tribunale di Bologna 31 December 2020, *Rivista italiana di diritto del lavoro,* forthcoming (2021) and Tribunale di Palermo 12 April 2021, available at www.dejure.it) and unfair labour practices (Tribunale di Firenze 9 February 2021 and Tribunale di Milano 28 March 2021, available at www.dejure.it). For a criminal case about illegal digital hiring cf Tribunale di Milano 28 May 2020 no 9 available at www.GiurisprudenzaPenale.com, 7-8 (2020).

[4] Art 2094 describes an employee as a person 'who undertakes an obligation to *cooperate* in the business in exchange for a remuneration by performing his work manually or intellectually at the *dependence* and *under the direction* of the entrepreneur'.

[5] In France, Cour de Cassation 28 November 2018 no 1737 *Revue de droit du travail*, 12, 812 (2018); in Belgium, Commission Administrative de Règlement de la Relation de Travail (CRT) 9 March 2018 no 113; in Switzerland, Cour de Justice de Geneve, Chambre Administrative 29 May 2020 no 535; in Netherlands, Rechtbank Amsterdam 23 July 2018 no 6622665 and Rechtbank Amsterdam 15 January 2019 no 7044576. Outside Europe, cf 33a Vara do Trabalho de Belo Horizonte 14 February 2017 no 0011359- 34.2016.5.03.0112.

[6] Respectively, Art 2 decreto legislativo 15 June 2015 no 81, as revised by decreto legge 3 September 2019 no 101, Art 409 Code of Civil Procedure, as revised by legge 22 May 2017 no 81, and Art 47-*bis* decreto legislativo 15 June 2015 no 81.

[7] This is why the heated debate on digital revolution also pertains to sociology and management studies. Cf F. Miele and L. Tirabeni, 'Digital technologies and power dynamics in the organization: A conceptual review of remote working and wearable technologies at work' *Sociology Compass* 14 (2020), K.C. Kellogg, M.A. Valentine, and A. Christin, 'Algorithms at Work: The New Contested Terrain of Control' 14 *Annals,* 366 (2020).

First, this article lists and reorders the main facts of the case, highlighting the (many) analogies and the (few) differences between the parties' allegations (Section II). Second, the article considers the nature of the platform and its effective corporate purposes (Section III). Third, the article examines the reasoning about the qualification of the working relationship, scrutinizing the ambiguous role assigned to the working relationship and performance continuity/permanency within the concrete subsumption of the case into the relevant legal patterns (Section IV). Fourth, the article examines the Tribunal's interpretative approaches to the relevant law and notions implemented to avoid applying the new intermediate patterns between subordination and autonomy (the hetero-organization) in favour of qualifying the courier as an employee (Section V). The arguments' respective shortcomings (Section IV) and strengths (para V) will be highlighted, and general considerations on the issues at stake will be presented at the end of this article (para VI).

II. From the Uncontested Facts to the Controversial Legal Interpretations at Stake

It should first be noted that the facts presented by the rider are substantially uncontested. According to the judge, they do not essentially differ from the company's allegations regarding the events that occurred but only with respect to their interpretations and legal qualifications (cf Section III and IV). For this reason, the case was decided without any hearings and investigations, as is typical in matters concerning the qualification of working relationships.[8] At first glance, this may be surprising, since algorithm management and the use of digital platforms are potentially characterized by greater obscurity in terms of their functioning; thus, they can overshadow the exercises of the hierarchical powers of direction and control (or the freedoms allowed) on an unprecedented scale, as noticed above. In this instance, however, this is not the case for at least two main reasons: this case concerns a location-based platform, which means that the riders' tasks are ontologically physical (delivering food), and so the exercise of management and powers may be hidden but their effects can hardly be denied or disputed, since they imply material actions and reactions on the part of the workers. Rather, the algorithm obscurity represents a real threat to the judicial investigation of facts where online web-based platforms are concerned,[9]

[8] The point has been immediately highlighted in the literature. Cf F. Martelloni, 'Il ragazzo del secolo scorso. Quando il rider è lavoratore subordinato a tempo pieno e indeterminato' *Questione Giustizia*, 24 December 2020, 4. V.A. Poso, 'Qual è la natura giuridica dei rider? Sono subordinati, bellezza! Commento a prima lettura della prima sentenza-zibaldone che farà discutere' *rivistalabor.it*, 1 December 2020.

[9] Regarding the numerous kinds of digital platform from a labour viewpoint, cf the ILO report 'World Employment and Social Outlook 2021'. The role of digital labour platforms in transforming

in contexts wherein workers perform tasks online and remotely (common examples include translation and graphic design services) or where it matters of specific aspects of the employment relationships (the most obvious example is probably litigation regarding discrimination).[10]

Moreover, these new digital tools, as a matter of their ambiguity, seem to exclude only one or a few elements of working relationships patterns – and specifically of the employment relationship – while maintaining other characteristics of these legal patters. Digital revolution induces (bigger or smaller) shifts in an ideal *continuum* between subordination and autonomy and a concentration of work performances on the border between the two opposite poles.[11] Thus, the debates work to establish whether these adjustments are sufficient to change the legal qualifications of the working relationship facilitated by digital means. Ultimately, the parties' efforts to influence the qualification of the rider have concentrated on the interpretations of the relevant statutes – in a sort of a stress test – rather than on the interpretation of the mere facts.

Another aspect to highlight is that the judge carefully describes these undisputed relevant facts for each step of the rider's performance. Since the facts are acknowledged, this may seem inconsistent or an example of futile verbosity; after all, such detailed recounting of facts is not particularly common in Italian case law.[12] In general, it is an unescapable choice when it comes to wisely deciding the legal qualification of the working relationships. However, the meticulous attention to the facts of the case is indisputably appropriate because it concerns a new phenomenon that the judge brought under Art 2094 for the first time.

The facts go as following. First, at the beginning of the relationship, the courier signed a contract as a self-employed worker, without any negotiation. The platform owner asked him to obtain a VAT number for a subsequent freelancer contract.

Second, for almost two years, the claimant worked as a rider using the app provided by the company, which was an essential instrument in the service delivery. Specifically, every week, the digital platform allowed the rider to book available shifts (so-called *slots*) as a means of scheduling his own working

the world of work' 23 February 2021 available at https://tinyurl.com/8bbnpmc2 (last visited 30 June 2021).

[10] The anti-discrimination law covers workers and employers, without any distinction whatsoever: cf G. Gaudio, 'Algorithmic management, poteri datoriali e oneri della prova: alla ricerca della verità materiale che si cela dietro l'algoritmo' 6(2) *Labour & Law Issues*, 19 (2020), who scrutinizes possible solutions for a fair decision in cases where the algorithm's functioning is obscure.

[11] A. Perulli, Oltre la subordinazione. La nuova tendenza espansiva del diritto del lavoro (Torino: Giappichelli, 2021), 21 and 114.

[12] M. Barbieri, 'Il luminoso futuro di un concetto antico: la subordinazione nella sentenza di Palermo sui riders' 6(2) *Labour and Law Issues*, 63, 71 (2020).

hours. Thus, he could determine if, when, and where he would deliver food. However, the algorithm governing the app influenced the worker's prerogative through the ranking periodically assigned to him, which was based on his efficiency and experience, customers' and partners' feedback, and his previous bookings of high-demand slots. Shift cancellations, delays in logging into the app, and the rider's absence from the area where he was expected to work were treated as penalties in the ranking system. Each week, the best-scored riders could book their preferred shifts before the other low-ranking couriers; these latter workers were restricted to the remaining slots.

Third, during the execution of the work, the digital platform 1) offered (and, in case of acceptance, assigned) one request at a time, 2) proposed an efficient track (also used to establish the compensation), and 3) provided specific automatic procedures for communication in case of logistic/technological problems. The rider' compensation was delivered every two weeks; the riders were required to return the customers' cash (if used to pay for the service) to the platform's owner by bank transfer through a specific procedure unless individually authorized to retain it as partial compensation. The platform created an invoice for each service on the rider's behalf. The claimant worked more or less eight hours per day and/or forty hours per week.

Fourth, before the end of the working relationship, the courier protested in front of a manager at the lack of safety devices and about periods of suspension from the digital platform; after complaining publicly about riders' working conditions, he was suspended permanently. All things considered, the Sicilian rider demanded to be recognized as an employee of the digital platform's provider and sought a fair remuneration pursuant to the national collective agreement (CCNL) applied to the company employees or applicable in view of the productive sector along with compensation for the retaliatory and oral dismissal represented by the suspension of his account disconnection.

Other allegations by the company as cited in the decision concern two main aspects. First, the company ascribed the temporary disconnections to the claimant's delay in transferring cash received from customers and the definitive account blockage to a technical fault. However, the disconnections could not be classed as dismissal since he was not an employee. Second, the written contract was rightly qualified as a self-employment contract because the rider could choose when to work, since the ranking system should be considered a rewards system rather than a punitive arrangement. Each rider could choose from and work the available residual shifts even if they had lower scores or no scores, for example in the case of new couriers. Ultimately, both parties recognized the freedom with respect to scheduling working time while the disputed point concerned the capacity of the external constraints arranged for ranking purposes to nullify these freedoms, influencing the qualification of the working relationship.

Many of the uncontested facts are the same across Italian and foreign case law regarding platform-based work, at least in their main profiles,[13] and mount similar challenges to the corresponding patterns of working relationships worldwide.[14] This confirms the global relevance of the issue and unveils the oligopolistic framework – or at least the similar strategies – of the few multinational enterprises providing analogous services through similar models based on algorithms and digital platforms.[15] This background explains another peculiarity of the decision, which is the comparative approach employed in the ruling: the judge cited numerous decisions from European and non-European case law to build her reasoning and enhance her arguments, finding support in other national jurisdictions even if they operated within different legal frameworks. The comparative approach is rightly considered perilous or useless with respect to issues falling within national rules, but it is becoming increasingly helpful in view of the global context of various phenomena.[16] Digital work is probably the latest and most salient example of this trend, alongside labour law literature, which is more and more open to European and international scientific interactions in this field.

III. Digital Platforms as Intermediary or Transport/Delivering Services: An Issue no Longer at Stake

The first issue concerns the corporate purpose of the defendant company and consequently the real economic activity carried out through the digital platform and the algorithm that governs the app. This point is correctly discussed as a preliminary aspect, since the qualification of the working relationship demanded before the Tribunal clearly centres on the enterprise's aims and objectives. This issue is rarely disputed where the qualification of working relationships is concerned. However, digital platforms are now

[13] An overview of the digital work conditions is provided in a report by V. De Stefano and A. Aloisi, *European legal framework for "digital labour platforms"* (Luxembourg: Publications Office of the European Union, 2018), 16.

[14] J. Prassl, 'What if your boss was an algorithm? The rise of artificial intelligence at work' 41(1) *Comparative Labor Law & Policy Journal*, 123 (2019).

[15] D. Dazzi, 'Gig economy in Europe' 2(12) *Italian Labour Law e-Journal,* 68, 94 (2019).

[16] M. Biasi, 'Uno sguardo oltre confine: i "nuovi lavori" della gig economy. Potenzialità e limiti della comparazione' 4(2) *Labour & Law Issues*, 3 (2018), who points out the risk inherent in applying that method to very different legal frameworks. Cf M.A. Cherry and A. Aloisi, 'Dependent contractors in the gig economy: a comparative approach' 66 *American University Law Review*, 635 (2017); S. Giubboni, 'La subordinazione del rider' *Menabò di Etica ed Economia* 140, 14 December 2020, 1; M. Barbieri n 12 above, 87, who affirms that the legal frameworks of the decisions cited are similar or identical; M. Magnani, *Diritto sindacale europeo e comparato* (Torino: Giappichelli, 3th ed, 2020), 3 affirms that comparison is necessary in today's global economy. Cf P. Adam, M. Le Friant, Y. Tarasewicz eds, *Intelligence artificielle, gestion algorithmique du personnel et droit du travail* (Paris: Dalloz, 2020), A. Baylos Grau's contributions on his blog, https://tinyurl.com/4b4ssnt3 (last visited 30 June 2021).

affecting enterprises' organization so intensely that not only the nature of the working relationships involved but also the corporate purposes are ultimately overshadowed or transformed.[17] It was not by chance that the first decisions on digital platforms did not centre on working relationships but on corporate and competition law.[18]

At first glance, controversy surrounds whether the platform's purpose was to provide a virtual place to match supply and demand for transport and delivery services, allowing the company to play the role of an intermediary services provider (and sometimes that of booking agent for the partners – namely riders or drivers, food providers, customers) or whether the platform's main activity involved the direct supply of transport and delivery services for customers realized through the riders' work. Various national and supranational jurisdictions have already dealt with this dilemma with the aim of verifying the compliance with the regulations established to deliver particular services (such as the licenses needed to operate private vehicles). In doing so, they had to verify who actually delivers the services and who effectively organizes and influences them. Thus, these jurisdictions assigned particular relevance to who determines the fares, the quality, and other circumstances pertaining to the services, which indirectly describes the relationship between platform owners and workers. The case law of the Court of Justice of the European Union (CJEU) closely examined working relationships to make competition law rulings, establishing that the control and influence of off-line (and material) work functional to the main economic purpose of the digital platforms and their owners is crucial to solving the dispute.[19] Ultimately, the CJEU decided to treat these companies as supplying transport services in the case of digital platforms, which govern/influence drivers' work, while treating other companies as mere intermediaries in the case of digital platforms delivering different services, such as accommodation.[20]

[17]This is why different legal hypotheses aimed at qualifying the phenomenon have been proposed in the literature. Cf L. Ratti, 'Online platforms and crowdwork in Europe: A two-step approach to expanding agency work provisions' 38 *Comparative Labor Law and Policy Journal*, 477 (2017); A. Rosin, 'Applying the temporary agency work directive to platform workers: Mission impossible?' 36(2) *International Journal of Comparative Labour Law and Industrial Relations*, 141 (2020).

[18] Tribunale di Milano 9 July 2015, *Rivista Italiana di Diritto del Lavoro*, II, 46 (2016), Case 434/15 *Professional Elite Taxi* v *Uber Systems Spain SL* [2017] ECLI:EU:C:2017:981, Case 526/15, *Uber Belgium BVBA* v *Taxi Radio Bruxellois NV* [2016] ECLI:EU:C:2016:830, Case 320/2016 *Uber France SAS* [2018] ECLI:EU:C:2018:221. Cf V. Brino, 'Il caso Uber, tra diritto del lavoro e diritto della concorrenza', in M. Biasi and G. Zilio Grandi eds, *Commentario Breve allo Statuto del Lavoro Autonomo e del Lavoro Agile* (Milano: Wolters Kluwer, 2018), 135.

[19] A. Aloisi, 'Demystifying flexibility, exposing the algorithmic boss: A note on the first Italian case classifying a (food-delivery) platform worker as an employee' *Comparative Labor Law & Policy Journal*, forthcoming (2021), available at SSRN: https://tinyurl.com/3hkud4wd or https://tinyurl.com/ybp2vt57, 5-6 (last visited 30 June 2021).

[20] Case 390/18 *Airbnb Ireland* [2019] ECLI:EU:C:2019:1112. Cf J. Morais Carvalho, 'Airbnb

The Palermo Tribunal follows this homogeneous case-law. First, it ignores the official business purposes declared in the company's certificate and in the two contracts concluded with the claimant. These documents, in fact, refer to an alleged broking function of the digital platform, which the judge appears to consider formal and thus auto-referential. However, the tribunal does not really investigate the case by interpreting the facts and subsuming them into the proper legal framework, as an orthodox approach requires in a civil law system. Rather, it simply mentions the foreign and CJEU decisions on the issue, overlapping them with the specific case in question. Theoretically, this construction of the reasoning could be questionable because the conventional technique is replaced by a comparative approach.[21] As noted, this latter method is risky because of the different legal frameworks within which case law stands. However, it is increasingly considered beneficial (or necessary) when comparable multinational entrepreneur patterns are concerned (cf above fn no 16), since it facilitates the efficient employment of similar reasoning and arguments (avoiding redundancy) and promotes a homogeneous approach to global phenomena, such as digitalization. In this case, however, the risk lies not in the different legal frameworks, since EU and Italian jurisdictions are notoriously intertwined, but in the diverse characteristics of the services provided: driving services in the case of the CJEU decisions (cf above fn no 18) and food delivery in the Sicilian case. From this perspective, the decision merely affirms that the CJEU's approach is 'certainly referable to commodities transport, other than to people transport'. The conclusion is probably agreeable, but it surely deserved to be boosted by an examination of the similarities and differences between the cases already decided in other jurisdictions and in the case in question. For example, in European case law concerning drivers, the contractual relationships involve the digital platform owner, the drivers, and just one category of customers (the passengers), while in the case of food (and commodity) delivery services, the contractual relationships that must be considered are more fragmented, because they involve two categories of clients, the commodity providers (analogous to restaurants) and the buyers (analogous to diners). This difference could have confirmed the judge's stance, because the stronger contractual fragmentation likely renders the digital platform more crucial in the management of the overall service and riders' performances.

A closer look at the case at hand reveals that the hasty approach of this part of the decision also seems to rest on the litigation strategies of both parties. In

Ireland case: One more Piece in the Complex Puzzle Built by the CJEU Around Digital Platforms and the Concept of Information Society Service' *Italian Law Journal*, 463 (2020) for a critical review. J. Gil García, 'Las múltiples formas de trabajo en las economías colaborativas y su regulacíon: el caso de «Airbnb»', in A. Todolí Signes and M. Hernández Bejarano eds, *Trabajo en plateformas digitales: innovación, derecho y Mercado* (Madrid: Aranzadi, 2018), 359.

[21] F. Capponi, 'Lavoro tramite piattaforma digitale e subordinazione: il ruolo dell'algoritmo secondo il Tribunale di Palermo' *Bollettinoadapt.it*, 30 November 2020.

fact, the ruling does not consider the contractual documents and relationships between the rider and the company, on one side, and the two customers at stake (the providers and buyers of the commodity delivered), probably because the firm itself accepted the homogenous qualifications already devised by other courts worldwide, while focusing on the main – still disputable – issue concerning the qualification of the company's working relationship with the rider. Furthermore, this preliminary aspect is not momentous from the perspective of competition law, since the services provided are not regulated by public bodies; thus, the platform different qualification does not involve illegal behaviours. The Palermo Tribunal refers to this preliminary matter and corresponding case-law only to deny relevance to the formal documents provided and begin the reasoning from a point that is now acknowledged among courts and scholars. This assumption means that a staple on a previous heated debate has been fixed but only for deregulated services such as food delivery. A recent UK Supreme Court judgement confirms this hypothesis.[22] In that case, in fact, the company strongly asserted its broker function as well as the workers' autonomous status, with the aim of preserving a business model designed to deal with both competition and labour law. For this reason, the Supreme Court thoroughly analysed both written agreements and relationships between Uber and its drivers (paras 22-26) as well as Uber and its passengers (paras 27-29).

IV. The Predetermination of Working Time and the Role of Performance Continuity in the Qualification of the Working Relationship: An Issue to Probe Further

If the decision follows the mentioned case law on the nature of the digital platform and the purpose of tech companies, it represents a breakthrough vis-à-vis the legal qualification of the riders' status. As recognized by the Tribunal, the first Italian rulings on the riders focused on the new liberating features associated with working for organizations operated via digital platforms. Thus, they strongly rejected the workers' claims about their alleged employment status. First, they recognized self-employed relationships,[23] and later they applied the new *intermediate* pattern designed for the so-called hetero-organized work pursuant to Art 2 decreto legislativo 15 June 2015 no 81 (henceforth Art 2).[24] At first, the

[22] Uber BV and others v Aslam and others [2021] UKSC 5.

[23] Tribunale di Torino 7 May 2018, *Argomenti di Diritto del Lavoro*, 4-5, 1227 (2018); cf G.A. Recchia, 'Gig economy e dilemmi qualificatori: la prima sentenza italiana' *Lavoro nella Giurisprudenza*, 7, 726 (2018) and Tribunale di Milano 10 September 2018, *Labor*, 1, 112 (2019).

[24] Art 2 (as reformed in 2018) states that the employment regime shall be applied to collaboration that is mainly personal, continuative and whose performance is organized by the contract. Art 2 has given rise to considerable debate that cannot be comprehensively cited here. It stems from the idea that the norm has no effects since it formalizes the outcome of the case law

new legal pattern was considered by a court to be a *tertium genus* between subordination and autonomy, recognizing only selected employers' prerogatives to the riders;[25] eventually, Art 2 was applied more extensively, extending almost all these prerogatives to the digital workers with the exception of the employment disciplines that ontologically fit the hetero-direction exercised pursuant to Art 2094 upon subordinated workers.[26] In a climax of sorts, the Palermo Tribunal now qualifies the rider as an employee under Art 2094, challenging all the heated discussions and efforts put in place to emphasize the new characteristics of digital platform work.

The substance of the ruling is significant and problematic for various reasons. First, it openly criticizes previous Italian decisions for having excluded the riders' subordination by only observing the relationship during the time immediately preceding the working performance, when the couriers can actually exercise their contractual prerogative to choose if, when, where, and how much they work. According to relevant opinions, in fact, these recognized freedoms are inconsistent with the subordination pursuant to Art 2094, which requires not only cooperation and dependence but particularly hetero-direction.[27] The Palermo Tribunal, rather, elects to follow an approach aimed at investigating the effective performance of the work, thus focusing on the central phase of the

interpretation of Art 2094, or that the norm updates or enlarges the classical notion of subordination, to the idea that the norm represents a step toward a real third category (and regime) between subordination and autonomy. Cf A. Perulli, n 11 above; P. Tosi, 'Il diritto del lavoro all'epoca delle nuove flessibilità – le collaborazioni eterorganizzate' *Giurisprudenza italiana*, 737 (2016); Id, 'Autonomia, subordinazione e coordinazione' *Labor*, 245 (2017); O. Mazzotta, 'L'inafferrabile etero-direzione: a proposito di ciclofattorini e modelli contrattuali' *Labor*, 1 (2020); G. Santoro Passarelli, 'Ancora su eterodorezione, etero-organizzazione, su coloro che operano mediante piattaforme digitali, i *riders* e il ragionevole equilibrio della Cassazione n. 1663/2020' *Massimario di Giurisprudenza del Lavoro*, 203 (2020); L. Nogler, 'La subordinazione del d.lgs. n. 81 del 2015: alla ricerca dell'autorità dal punto di vista giuridico' *WP C.S.D.L.E. "Massimo D'Antona".IT*, no 267 (2015); O. Razzolini, 'I confini tra subordinazione, collaborazioni etero-organizzate e lavoro autonomo coordinato: una rilettura' *Diritto delle relazioni industriali*, 2, 360 (2020); G. Cavallini, 'Le nuove collaborazioni etero-organizzate: cosa cambia dopo la riscrittura dell'art. 2 d.lgs. n. 81/2015 (e la Cassazione sul caso Foodora)' *Giustiziacivile.com*, 2, 13 (2020); S. D'Ascola, 'La collaborazione organizzata cinque anni dopo' *Lavoro e diritto*, 1, 3 (2020).

[25] Corte d'Appello di Torino 11 January 2019, *Rivista Italiana di Diritto del Lavoro*, II, 350, 358 (2019); G.A. Recchia, 'Controrordine! I riders sono collaboratori eterorganizzati' *Lavoro nella Giurisprudenza*, 403 (2019).

[26] Corte di Cassazione 24 January 2020 no 1663, *Lavoro Diritti Europa*, 1 (2020); R. Romei, 'I riders in Cassazione: una sentenza ancora interlocutoria' *Rivista Italiana di Diritto del Lavoro*, I, 89 (2020); M.T. Carinci, 'Il lavoro etero-organizzato secondo Cass. n. 1663/2020: verso un nuovo di sistema di contratti in cui è dedotta un'attività di lavoro' *Diritto delle Relazioni industriali*, 2, 488 (2020); A. Lassandari, 'La Corte di Cassazione sui riders e l'art. 2, d.lgs. n. 81/2015' *Massimario di Giurisprudenza del Lavoro* 123 (2020); A. Perulli, 'La prima pronuncia della Corte di Cassazione sull'art. 2, co. 1, d. lgs. n. 81/2015: una sentenza interlocutoria' *Lavoro Diritti Europa*, 1 (2020).

[27] Hetero-direction has been considered the crucial criterion according to the case-law. Cf references in M. Pallini, 'Towards a new notion of subordination in Italian labor law?' 12(1) *Italian Labour Law E-Journal*, 1 (2019) and Corte di Cassazione 10 July 1991 no 7608, *Rivista Italiana di Diritto del Lavoro*, I, 103 (1992) about the so-called pony express.

working relationship.

An examination of the reasoning's construction, however, reveals that the decision is not innovative with respect to this abstract assumption: as a matter of fact, the criticized case-law – in its last rulings – also assigned relevance to the execution of the performance to recognize the hetero-organization pursuant to Art 2, and the Sicilian judge also identified the hetero-organization in this phase of the working relationship. The decision, furthermore, expressly confirms that the issue of the free predetermination of the working time remains a crucial point in qualifying a working relationship. Thus, it is at odds with the views of some authors, according to whom the qualification of the relationship shall depend *only* or *mostly* on the effectiveness of the performance delivered (when the issue of *how* to work is concerned).[28] Ultimately, it concretely assigns relevance to *both* the effective performance and the workers' exercise of their prerogatives prior to performing the tasks (at what time *if*, *when*, *where*, and *how much* to work are concerned).

Actually, the novelty of the decision regards the interpretation of the respective prerogatives in the times preceding the material tasks. The decision intensifies the investigation of this segment of the relationship by questioning how effectively the rider's liberties could be exercised. To support the approach, a recent CJEU order is cited, which excludes couriers by the scope of the EU law if national courts verify that their independence is merely notional – also in terms of predetermination of the working time – thus hiding bogus self-employed workers.[29] Even if the CJEU order is not really foundational, the relevant point is that, according to the judge, the riders' freedoms to determinate their working times are fictitious because of the algorithm's influence on the couriers' decision, exerted via awards and discouragements in the system of shift booking and cancellation.

The conclusion is agreeable in practice but problematic from the legal perspective.[30] On the one hand, the conclusion is correctly founded on the firm determination to transcend contractual formalism – such as assigning relevance only to written agreements – and to focus on the effective overall working relationships. The approach is appropriate and represents an 'established legal tradition'[31] in the specific sense that, in Italy, the need to adopt such a technique is clearly established by Art 1362 Civil Code. The latter affirms that 1) in the interpretation of a contract, the parties' common intentions shall be investigated

[28] M. Barbieri, n 12 above, 74-79.

[29] Case 692/19 *B* v *Yodel Delivery Network Ltd*, [2020] ECLI:EU:C:2020:288. Cf. A. Aloisi, 'Time is running out'. The Yodel order and its implications for platform work in the EU' 13(2) *Italian Labour Law E-Journal*, 67 (2020), J. Adams-Prassl et al, 'EU Court of Justice's decision on employment status does not leave platforms off the hook' *Regulation for Globalization*, 29 April 2020, available at https://tinyurl.com/yt22sc97 (last visited 30 June 2021).

[30] F. Martelloni, n 8 above, 7.

[31] A. Aloisi, n 29 above, 7.

and not only the literal meaning of the words and 2) to determine the parties' intentions, their comprehensive behaviour, even following the agreement, shall be considered. In the case of the working contract, the parties' comprehensive behaviour usually plays a more relevant role than that played in other kind of agreements, since it is a freeform contract and because of the unbalanced power in the bargaining process.[32] It is not by chance that these principles for interpreting contracts are common to other legal systems.[33] However, in interpreting the relevant behaviours, the decision regarding how the algorithm constraints became a legal exegetic element that allows the formal workers' freedoms to determine if, when, where, and how much they work to be ignored is not clear. In fact, it is affirmed that the freedoms were fictitious not because of their absolute ineffectiveness but because their exercise was altered by predicted negative effects, ultimately resulting in *constrained freedoms*.[34]

This knot is difficult to disentangle and scholars remain divided. Ultimately, the problem is linked to the function and meaning of work performance continuity within the employment relationship. If it is considered to be an essential element for the identification of an employee, continuity – as one of the consequences of the working time hetero-determination – has to be identified as an *a priori* obligation burdening somehow the riders even before the performance, not only as a *de facto* characteristic proved *a posteriori*.[35] In literature and case law, this critique is quashed in two different ways: first, by pointing out that subordination shall not necessarily focus on the stage that precedes the effective performance[36] – namely, saying that continuity is not essential but merely a possible indicator of subordination; and second, by highlighting that continuity does not concern the effective performance *per se*, but the *contractual relationship*, resting on the obligation to be available by a deadline after which the rider's account will be definitively disconnected.[37] The

[32] Cf L. Castelvetri, Perché discutere (ancora) di alternativa tra contratto e rapporto di lavoro?' *Diritto delle Relazioni Industriali* 467 (2002); C. Smuraglia, *Il comportamento concludente nel rapporto di lavoro* (Milano: Giuffrè, 1963); F. Benatti, 'Che ne è oggi del testo del contratto' *Banca Borsa Titoli di Credito*, 1 (2021). Corte di Cassazione 10 April 2000 no 4533, *Foro italiano*, I, 2196 (2000).

[33] Eg *Uber BV and others* v *Aslam and others* [2021] UKSC 5, paras 58-64.

[34] Cf A. Maresca, 'Brevi cenni sulle collaborazioni eterorganizzate' *Rivista Italiana di Diritto del Lavoro*, I, 73 on the distinction between directives and performances induced by the management.

[35] F. Carinci, 'Tribunale Palermo 24/11/2020. L'ultima parola sui rider: sono lavoratori subordinati' *Lavoro Diritti Europa*, 12 (2021); E. Puccetti, 'La subordinazione dei Riders. Il canto del cigno del tribunale di Palermo' *Lavoro Diritti Europa*, 12 (2021). Cf also A. Perulli, 'Il diritto del lavoro "oltre la subordinazione": le collaborazioni etero-organizzate e le tutele minime per i riders autonomi' 410 *WP CSDLE "Massimo D'Antona".IT*, 71 (2020).

[36] M. Barbieri, n 12 above, 76, M. Barbieri, 'Della subordinazione dei ciclofattorini' 5(2) *Labour & Law Issues*, 35 (2019), G. De Simone, 'Lavoro digitale e subordinazione. Prime riflessioni' *Rivista Giuridica del Lavoro*, 11 (2019). Cf also Razzolini, n 24 above, 371. For a different approach to the issue, see P. Ichino, 'Le conseguenze dell'innovazione tecnologica sul diritto del lavoro' *Rivista Italiana di Diritto del Lavoro*, I, 525 (2017).

[37] M. Barbieri, n 12 above, 81, V. Bavaro, 'Questioni in diritto su lavoro digitale, tempo e libertà'

decision does not take a clear stance regarding these different hypotheses because it ultimately considers continuity to be an essential aspect of the subordination, without establishing if and how the constraints on the workers' freedoms give rise to an *a priori* obligation to work or to be available. The vagueness (or the contradiction) became evident when the judge, at a different stage of the reasoning, returns to the issue of continuity and treats it also as a mere hint of subordination.

The issue is worth disentangling before the courts, which are now required to clarify whether an *a priori* obligation is required by law and, if so, what the content of such an obligation should be. The dilemma at stake is politically crucial since the given solution involves the possibility of applying the maximum level of labour protections – but also employers' prerogatives[38] – within the so-called gig economy. This latter, in fact, consists of the management of casual work – now easily reachable and exploitable due to the use of algorithms – and on the relinquishment of the continuous working relationships that have been the most efficient means of benefiting from work during the Fordism era.[39] One possible answer could be suggested by a recent decision of the German Federal Labour Court (BAG), which has recognized the contractual relevance of a *de facto* continuity of the working relationship because it has been 'induced' (rather than 'imposed') by the digital platform.[40]

V. From the Hetero-Organization to the Classic Subordination: A Surplus of Arguments

The dilemma posed by working time predetermination and performance continuity is relevant to the exclusion of a rider's status as purely self-employed. However, it is not sufficient to distinguish among hetero-directed workers pursuant to Art 2094 Civil Code and other *intermediate* patterns, such as the hetero-organization pursuant to Art 2 and the so-called 'co.co.co.' (continuative and coordinated collaborations) pursuant to Art 409 Procedural Civil Code.[41] It should be recalled that continuity (however it may be defined, cf section IV) is expressly provided for in Art 2 and Art 409 Procedural Civil Code, while Art 2094 Civil Code has traditionally been interpreted as inferring continuity even without literal references, owing to the implications of coordination, hetero-

Rivista Giuridica del Lavoro, 53 (2018).

[38] A. Aloisi, V. De Stefano, *Il tuo capo è un algoritmo* (Bari: Laterza, 2020), 141.

[39] V. De Stefano, A. Aloisi, n 13 above, 25, P. Ichino, n 36 above, 526.

[40] Cf the comment on the press release by L. Nogler, 'La Corte federale del lavoro tedesca risolve il rompicapo della qualificazione dei lavoratori delle piattaforme' *Giornale di Diritto del Lavoro e delle Relazioni Industriali*, 835 (2020). Cf the decision (in German language) https://tinyurl.com/m9p3c45p (last visited 30 June 2021).

[41] A. Perulli, n 11 above.

direction, and dependence.[42]

The judge leads towards the rider's subordination by also focusing on other aspects of the case and applying different legal interpretation techniques to them. From a methodological perspective, the decision simultaneously invokes two classic methods traditionally employed by Italian courts to identity subordination: the *subsumption* and *typological* techniques.[43] The first is aimed at connecting (ie *subsuming*) the worker and his/her relationship exactly (in)to the abstract provision of the law (Art 2094, in the case of employment status); the typological method consists in the study of analogies and differences between concrete circumstances and the relevant abstract provisions, through empirical indicators derived from experiences that represent indirect manifestations or presumptions of subordination.[44] This latter method is commonly used when the subsumption does not help in the specific case because of the difficulty in identifying specific and continuous command-and-control powers.[45] In this decision, rather, they are overlapped.[46]

Beginning with the subsumption method, the decision first recalls the necessity of updating the interpretation of Art 2094, since it was devised in 1942 when the archetypal workers were Fordist labourers.[47] This consideration has been condemned as a false myth by some scholars, who point out that in 1942, the number of agricultural workers was too significant to be ignored by lawmakers[48] and/or because Art 2094's wording was strongly influenced by Barassi's thoughts on employment relationships delivered during the first decades of the 1900s.[49] These critiques are significant because they aim to demonstrate that Art 2094 was written also for casual and daily workers, who are similar to present-day couriers with respect to their work performance

[42] O. Razzolini, n 24 above, 360, P. Digennaro, 'Subordination or subjection? A study about the dividing line between subordinate work and self-employment in six European legal systems' 6(1) *Labour & Law Issues,* 4, 33 (2020). Cf F. Ferraro, 'Continuatività e lavoro autonomo' *Labor*, 5, 583 (2020).

[43] L. Nogler, 'Metodo tipologico e qualificazione dei rapporti di lavoro' *Rivista Italiana di Diritto del Lavoro*, I, 182 (1990), Id, 'Ancora sul "tipo" e rapporto di lavoro subordinato nell'impresa' *Argomenti di Diritto del Lavoro*, 109 (2002), G. Proia, 'Metodo tipologico, contratto di lavoro subordinato e categorie definitorie', *Argomenti di Diritto del Lavoro*, 37 (2002).

[44] Cf V. Pietrogiovanni, 'Between Sein and Sollen of Labour Law: Civil (and Constitutional) Law Perspectives on Platform Workers' 31(2) *King's Law Journal*, 313, 317 (2020) for further references.

[45] Corte di Cassazione 5 March 2009 no 5314, available at www.dejure.it, Tribunale Genova 11 January 2016 no 5, available at www.dejure.it.

[46] As already noted (section IV), continuity is employed both as an essential element and as indicator of subordination.

[47] The decision wrongly cites (twice) as the historical framework of the Civil Code's entrance into force the first Industrial Revolution (which happened in 1700) instead of the Third Industrial Revolution.

[48] M. Barbieri, n 12 above, 84.

[49] F. Martelloni, n 8 above, 7. Cf L. Barassi, *Il contratto di lavoro nel diritto positivo italiano* (Milano: Società editrice libraria, 1st ed, 1901; reprint Milano: Vita & Pensiero: 2003).

(dis)continuity; thus, no significant exegetic efforts shall be invested in including digital work under the subordination pattern. The quarrel is probably theoretical: on the one hand, the 1942 lawmakers could not have ignored agricultural (and casual) work, which was still widespread in Italy; on the other hand, the industrial work model was equally significant for a lawmaker wishing to forge an updated legal framework. In fact, the report annexed to the Italian Civil Code for the king's ratification expressly explains that the phrasing chosen to describe the employee should necessarily be 'ample' and 'comprehensive'.[50]

Moreover, the judge revises the Art 2094 analysis by eliciting a purposive interpretation of it, in line with recent foreign courts' attempts.[51] She defends her purposive approach by highlighting that innovative interpretations of the subordination pattern have already been delivered within case law to cover employment relationships that do not clearly show the essential character of the hetero-direction. The first concerns high-skilled and low-skilled workers – such as managers and employees performing simple and recurrent tasks – for whom mitigated ('attenuated') subordination is sufficient, since they do not need continuous and relevant commands.[52] The second is based on the 'double alienness' (or alienity) theory, proposed by a minor portion of case law and literature.[53] It consists in the assumption that only in an employment relationship does the worker not own the product of the enterprise nor its organizational means, so that the subordination relies on a unique framework of different/oppositional interests, rather than a particular manner in which the activities should be executed.

The concrete handling of the subsumption method shows (again) the judges' (problematic) inclination to overlap and mix different approaches to the relevant legal pattern. In fact, while the mitigated/attenuated subordination theory seems to be cited simply to demonstrate the abstract possibility of innovating the subordination pattern without actually being applied to the case,

[50] Report to the Italian Civil Code, Libro V, 1942, available at https://tinyurl.com/3ejxwfcx (last visited 30 June 2021).

[51] M. Biasi, 'Tra fattispecie ed effetti: il "purposive approach" della Cassazione nel caso Foodora' *Lavoro Diritti Europa*, 2020. *Uber BV and others* v *Aslam and others* [2021] UKSC 5, paras 65-78; *Autoclenz Ltd* v *Belcher* [2011] UKSC 41; [2011] ICR 1157. Cf A. Bogg, 'For whom the bell tolls: "Contract" in the gig economy' OxHRH Blog, available at https://tinyurl.com/v6ebbsw (last visited 30 June 2021).

[52] Corte di Cassazione 28 October 2020 no 23768, available at www.dejure.it regarding a home delivering job and Corte di Cassazione 29 October 2020 no 23927, available at www.dejure.it regarding a manager.

[53] Corte costituzionale 12 February 1996 no 30, *Diritto del Lavoro*, II, 52 (1996), V. Pietrogiovanni, 'Redefining the Boundaries of Labour Law: Is "Double Alienness" a Useful Concept for Classifying Employees in Times of Fractal Work?', in A. Blackham, M. Kullmann, and A. Zbyszewska eds, *Theorising labour law in a changing world: new perspectives and approaches* (Oxford: Hart Publishing, 2019), 55, S. D'Ascola, 'Platform Work and "Double Alienness", in A. Perulli and T. Treu eds, *The future of work: labour law and labour market regulation in the digital era* (Alphen aan den Rijn: Kluwer Law International, 2020), 307.

the rider's working relationship has actually been declared as having satisfied both the *orthodox* requisite of subordination – namely the hetero-direction – and the requisite of the 'double alienness'.

As if the satisfaction of these two requisites was not enough, the typological technique has been employed (as mentioned above) as a subsidiary and complementary approach to confirm and enforce the working relationship subsumption into Art 2094. The decision, in fact, identifies several empirical indices of subordination: the *de facto* continuity (as already noted in the previous section), worker availability even in times of no orders, the exercised of unorthodox disciplinary power ended up to a disconnection which can be considered as an oral dismissal and, last but not least, the absolute dependence to the digital platform and the lack of room for autonomous initiative even in case of technical algorithm disfunction This last index is highlighted because, according to the judge, it reveals that the rider was toothless before the platform and wholly unaware of its functioning to the extent that it leaves the courier exactly at the same level of dependency or even at a lower level than a Fordist labourer of the last century, saying that he is hetero-directed in the classical meaning.[54]

Ultimately, the reasoning seems to be influenced by the urge to fortify a (hitherto) unique decision, but the surplus of approaches and interpretative techniques risks undermining the ruling by weakening arguments that cannot easily and efficiently coexist.

VI. Final Remarks: 'Mind the Gap' Between National Legal Orders and Global Digital Platforms

First, the case confirms that in the digital era, every human activity can be the object of employment relationships, self-employment, and other intermediate patterns.[55] Actually, Italian case law offers the maximum array of legal solutions and approaches to the issue, following this recent decision, which breaks the last taboo regarding the employment status of digital workers and imposes a discussion without restraints on any hypotheses. This ruling's influence on case law cannot be predicted, since each case is unique in a civil law system and considering that litigation strategies and the realities that emerged (or not) before the judges may differ significantly. However, the courts are dealing with business models that are similar worldwide owing to the oligopolistic framework

[54] Cf Opinion of Advocate General Szpunar delivered on 11 May 2017, case 343/15 *Asociación Profesional Elite Taxi* v *Uber Systems Spain* [2017] ECLI:EU:C:2017:364, according to whom 'indirect control such as that exercised by Uber, based on financial incentives and decentralised passenger-led ratings, with a scale effect, makes it possible to manage in a way that is just as – if not more – effective than management based on formal orders given by an employer to his employees and direct control over the carrying out of such orders'.

[55] Corte di Cassazione 15 June 2020 no 11539, available at www.dejure.it.

within which big tech companies operate;[56] thus, the heterogeneity of the case law comes at the expense of certainty and predictability as well as the possibility of the national legal system's ability to better influence global business models and work management. It is not by chance that several companies reacted to this first phase of the fight by preparing to hire riders as employees while other companies simply resisted or updated the algorithm in light of the worldwide jurisprudence stance.[57]

It is not easy to describe what relationship bounds facts (like business and management) and (case)law and to determine whether and to what extent the latter influences entrepreneurs' organizations and command-and-control power or whether and how far business and managerial trends affect lawmakers and interpreters of the law.[58] However, this case and the worldwide case-law on digital work and platforms represent an exemplar and updated illustration of the complex dialogue between economic actors (firms, workers, customers, and unions) and legal orders.[59] Clearly, the only way to exert greater influence on the economic and production system is to build a homogenous case-law at both the national and international levels – a goal that cannot be easily reached since the national statutes approach the gig economy and digital platforms in ways that are as heterogenous as the jurisprudence stances.[60] However, the next stages in this story will depend on the understanding of the digital platform's core nature: if it will be regarded as the manifestation of a new (less hierarchical) business model, the new freedoms (for all the actors involved) that it allows could be better recognized by courts. On the contrary, if digital platforms will be assumed as the new means for already-known underlying (vertical) business models, the judicial trend in favour of riders' subordination will not be easily stopped worldwide.

[56] N. Petit, *Big Tech and the Digital Economy: The Moligopoly Scenario* (Oxford: Oxford University Press, 2020).

[57] S. Sciorilli Borrelli and D. Ghiglione, 'Italy Emerges as Next Front in Gig Economy Labour Battle' *Financial Times*, available at https://tinyurl.com/9sd4v8x4 (last visited 30 June 2021).

[58] M. Barbieri, n 12 above, 69.

[59] M. Novella, 'Il rider non è lavoratore subordinato, ma è tutelato come se lo fosse' 5(1) *Labour & Law Issues*, 85 (2019), L. Mengoni, 'Diritto e tecnica' *Rivista trimestrale di diritto della procedura civile*, 1, 6 (2001).

[60] V. De Stefano and M. Wouters, 'Embedding Platforms in Contemporary Labour Law', in J. Drahokoupiland and K. Vandaele eds, *A Modern Guide to Labour and the Platform Economy* (Cheltenham: Edward Elgar Publishing, 2021), forthcoming.

Science at the Italian Bar: The Case of Hydroxychloroquine

Paola Monaco*

Abstract

Due to the increasing number of legal questions which cannot be answered without recourse to scientific knowledge, the issues surrounding the relation between science and the law have become a hot topic in legal debate. For this reason, it is not surprising that the tragedy of COVID-19 is raising many questions for lawyers to be debated in court. In the light of this, the paper aims to analyse one very interesting example of the use of scientific knowledge by an Italian (administrative) court: order no 9070 of the Italian Council of State (Consiglio di Stato) of 11 December 2020, through which the highest administrative court suspended the decision of the Italian Medicines Agency (AIFA) to forbid the off-label use of a drug (hydroxychloroquine – HCQ) in the treatment of COVID-19. After having analysed the main points on which the Italian Council of State decision is based, the essay will offer some considerations on how legal scholarship, across both common and civil law jurisdictions, has tried to offer some solutions to the problem of courts dealing with scientific or technical knowledge. In the conclusions it will be verified whether these principles might be useful and applicable before administrative courts as well.

I. Science and the Law: The Never-Ending Story

There is nothing new in science entering Western courtrooms.[1] Due to the increasing number of legal questions which cannot be answered without recourse to scientific knowledge, the issues surrounding the relation between science and the law have become a hot topic in legal debate, both in common and civil law.[2] This means that courts have the difficult task of devising legal methods for determining the proper evidentiary place to be given to science in judicial disputes. In some jurisdictions – such as the US – the courts have mostly handled the problem by trying to establish the standards for the

* PhD and Contract Researcher at Bocconi University (Milano, Italy).

[1] S. Jasanoff, *Science at the Bar: Law, Science and Technology in America* (Cambridge: Harvard University Press, 1995); D.L. Faigman, *Legal Alchemy: The Use and Misuse of Science in the Law* (New York: W.H. Freeman and Co, 1999); Id, *Laboratory of Justice: The Supreme Court's 200-Year Struggle to Integrate Science and the Law* (New York: Times Books/Henry Holt, 2004).

[2] P. Monaco, 'Scientific Evidence in Civil Courtrooms: A Comparative Perspective', in F. Fiorentini and M. Infantino eds, *Mentoring Comparative Lawyers: Methods, Times, and Places.* Liber Discipulorum *for Professor Mauro Bussani* (Cham: Springer, 2020), 95-110.

admissibility of scientific evidence in (civil) proceedings,[3] while in other jurisdictions – such as on the European continent – the courts have largely focused on the rules for choosing the experts who will assist judges in decisions involving scientific matters.[4]

It is not surprising that the tragedy of the COVID-19 pandemic has given rise to (and, unfortunately, will give rise to) numerous scientific questions for lawyers to be debated in court. And it not surprising that judges are finding more difficulty than ever in trying to deal with these new and delicate issues, because of the fact that this new field raises questions in legal proceedings to which science has not provided many answers yet.

One very interesting example of the use of scientific knowledge before the Italian (administrative) courts stems from decision no 9070 of 11 December 2020, by the Italian Council of State (Consiglio di Stato), the highest Italian administrative court.[5] The case stemmed from a proceeding for interim relief against the decision of the Italian Medicines Agency (AIFA) to forbid the off-label use of a drug (hydroxychloroquine – HCQ) in the treatment of COVID-19. As we will see in this paper, the Consiglio di Stato overruled the decision of AIFA.

After a brief introduction of the framework in which the order of Consiglio di Stato was reached (section II), we will summarise the main points on which the Consiglio di Stato decision is based (section III), focusing our attention on the use of scientific principles in legal reasoning. Then, some considerations will be presented on how legal scholarship, across both common and civil law jurisdictions, has tried to offer some solutions to the problem of courts dealing with scientific or technical knowledge, especially in civil proceedings (section IV). We will then verify whether these principles might be useful and applicable before administrative courts as well (section V).

II. The Context of the Order 9070/2020

In the aftermath of the explosion of the COVID-19 pandemic, hydroxychloroquine, an antimalarial drug used to treat systemic lupus erythematosus and rheumatoid arthritis, was suggested as a possible method of prevention or treatment for the new illness thanks to the evidence of its in-vitro inhibition of severe acute respiratory syndrome.[6] This is why, lacking an effective therapy to treat the illness, many national medicine agencies,[7] including the

[3] See below section IV.

[4] ibid.

[5] The order is available – in Italian – at www.giustizia-amministrative.it.

[6] F. Turone, 'Ruling Gives Green Light for Controversial COVID-19 Therapy. Administrative Judges Overrule Regulator to Authorize Hydroxychloroquine' *nature.com*, 18 December 2020.

[7] See for example the authorisation of March 2020 (now revoked) of the Food and Drug Administration (FDA) in the USA: https://tinyurl.com/3zzm3yfp (last visited 30 June 2021), and

Italian Medicines Agency (AIFA), allowed an emergency authorisation for its off-label use.[8]

However, soon after the availability of new studies and data showed the lack of efficacy of the drug and even an increase in adverse events in patients. Under such conditions, hydroxychloroquine was no longer recommended in COVID-19 patients by the medicine agency regulators.[9] So AIFA, with two notes (the first on 26 May 2020 and the second on 22 July 2020) suspended the authorisation for the off-label use of hydroxychloroquine outside of clinical trials.[10]

As we will explain in the next section, a group of doctors did not agree with the decision of AIFA. On the contrary, they maintained they had observed that HCQ was able to provide certain benefits in early-stage patients and for this reason they presented a claim to attempt to have this provision suspended.

III. The Claim

As we said above, a group of specialist physicians presented a claim to the administrative court of first instance – the Regional Administrative Court of Lazio (TAR Lazio) – to ask for an interim suspension of the decision of AIFA of 22 July 2020,[11] which allowed the off-label use of hydroxychloroquine in the treatment of COVID-19 patients only in randomised clinical trials. Since TAR Lazio rejected the doctors' claim, the decision was appealed in front of the Council of State.[12] The Council of State reformed the order of the TAR Lazio and suspended the decision of AIFA. The case is now being debated in front of the TAR on its merits.

the Société Française de Pharmacologie et de Thérapeutique (SFPT - France).

[8] AIFA, 'COVID-19 - AIFA autorizza nuovo studio clinico sull'idrossiclorochina', 9 April 2020, available at https://tinyurl.com/y23y4m9h (last visited 30 June 2021).

[9] See for example the decisions of the European Medicines Agency (EMA), 'COVID-19: chloroquine and hydroxychloroquine only to be used in clinical trials or emergency use programmes', 1 April 2020, available at https://tinyurl.com/e7j96uxy (last visited 30 June 2021); Medicines and Healthcare products Regulatory Agency (MHRA – UK), 'MHRA suspends recruitment to COVID-19 hydroxychloroquine trials', 16 June 2020, available at https://tinyurl.com/m86wnm68 (last visited 30 June 2021); FDA, 'Letter revoking EUA for chloroquine phosphate and hydroxychloroquine', 15 June 2020, available at https://tinyurl.com/p9nvjxes (last visited 30 June 2021); German Bundesinstitut für Arzneimittel und Medizinprodukte (BFARM), 'Hydroxychloroquin: Risiko für schwerwiegende Nebenwirkungen bei Anwendung zur Behandlung von COVID-19', 29 April 2020, available at https://tinyurl.com/3f9jjm83 (last visited 30 June 2021).

[10] AIFA, 'AIFA sospende l'autorizzazione all'utilizzo di idrossiclorochina per il trattamento del COVID-19 al di fuori degli studi clinici', 26 May 2020, available at https://tinyurl.com/2ptvstry (last visited 30 June 2021).

[11] The same physicians had already presented a claim to ask for an interim suspension of the decision of AIFA of 26 May 2020, but this was rejected by the Tribunale amministrativo regionale Lazio Roma 14 September 2020 no 5911: Consiglio di Stato 24 November 2020 order no 9070, point 2.3, available at www.giustizia-amministrativa.it.

[12] Tribunale amministrativo regionale Lazio Roma 16 November 2020 no 7069, available at www.giustizia-amministrativa.it.

1. The Theses of the Appellants and Respondents

The reasoning of the Council of State begins by giving some scientific details about hydroxychloroquine (HCQ). For many years – the court explains – HCQ has been used as an antimalarial drug, and also as a treatment for systemic lupus erythematosus and rheumatoid arthritis, used by about 60,000 patients in Italy.[13] After the development of the SARS-CoV-2 pandemic, HCQ use was suggested as a possible method of prevention or treatment for COVID-19.[14] Even if its efficacy was demonstrated by evidence of in-vitro inhibition of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the in-vivo studies benefits were much debated in the scientific community.

After having offered a brief overview of the scientific studies about the drug,[15] the Council of State illustrates that at the outbreak of the pandemic AIFA consented to the off-label use of HCQ, but on 26 May 2020 it modified the product information label for these medicines (also according to the recommendations of the European Medicines Agency, EMA) suspending its off-label use. This decision was based on two reasons. First, the available data were not consistent and did not demonstrate a clear clinical benefit; second, there was a risk of cardiac toxicity from high doses usage.[16]

The Council of State then goes on to describe the position of the claimants/appellants. According to the doctors who brought the request for interim relief, the drug was effective and the decision of AIFA was lacking in a proper investigation of the data. Furthermore, in the appellant doctors' opinion, the decision of AIFA violated their autonomy – as guaranteed by the Italian Constitution and by the law – in prescribing under their own responsibility the drug to non-hospitalised subjects who have given their informed consent.[17]

These arguments were rebutted by AIFA. AIFA maintained that its decisions, far from being taken without a profound study of the evidence as the appellant tried to demonstrate, were based on the last and best available evidence in the light of the safest guarantee for patients. For this reason, according to AIFA, its decision on the matter represented the fruit of their very technical discretion the merits of which the court cannot evaluate (especially during an interim proceeding like the one in front of the Council of State).[18]

In response to this, highlighting the undisputed role of AIFA in the protection of public health, and the equally undisputed scientific basis of the determinations of AIFA, the Council of State stresses that there are no decisions – no matter how

[13] Consiglio di Stato, n 11 above, point 1.2.

[14] J. April, 'Hydroxychloroquine in the prevention of COVID-19 mortality' 3 *The Lancet Rheumatology* (2020).

[15] ibid points 1.5 to 1.6.

[16] ibid point 1.9. See the first AIFA communication on the use of hydroxychloroquine: https://tinyurl.com/26pa83pe (last visited 30 June 2021).

[17] ibid points 2-2.1.

[18] ibid point 8.1.

delicate the issue at stake is, such as the off-label use of drugs against COVID-19 – which could not be judged by the administrative court to control the correct use of the technical discretionary powers. In fact, under Art 113 of the Italian Constitution, the judicial safeguarding of rights before the bodies of administrative justice is always permitted against acts of the public administration, and cannot be excluded for or limited regarding particular categories of acts, such as the elements of the information sheet on the use of HCQ forbidding its off-label use.[19] From the perspective of the court, the off-label prescriptions could be restricted by AIFA only if two conditions are met: when according to scientific knowledge and experimental evidence the use of the HCQ proves to be ineffective and unsafe.[20]

As we will discuss below,[21] given the fact that judicial review is admitted on this type of decision, what kind of control is the court entitled to carry out? The answer to this question represents the core of the decision, and the crucial point of our analysis. The reasoning of the court on this matter is discussed in the next section.

2. Form or Substance Under Review?

Once the Council of State recognised the admissibility of the judicial review of a decision of a public administration in a case that involves technical discretion, the next hot topic to discuss for the highest administrative court is if this control could be exercised only on the *extrinsic* side of the act (ie, only on the form), or also on the *intrinsic* sphere of the act (ie, on the reasoning underpinning the decision).

For a long time, the control of the administrative court on the technical decisions of public administrations was interpreted as a merely formal control on the logical reasoning followed in reaching a decision.[22] However, the most recent interpretations[23] allow the administrative court to perform a truly direct control of the coherence and correctness of the technical criteria used by the administrative authority to reach the decision.[24]

As underlined by the Council of State, this does not imply that the administrative court is charged with a jurisdictional control on the merits of debatable choices; rather, the task of the administrative court is to verify the rational credibility of the scientific knowledge underlying these choices.[25] In this light, the judicial review is not a mere extrinsic control, but an intrinsic one, involving the use of the same scientific knowledge applied by the public

[19] ibid points 8.2-8.3.
[20] ibid point 7.
[21] See below section V.
[22] See below section V.
[23] See, for example, Consiglio di Stato 6 July 2020 no 4322, available at www.giustizia-amministrativa.it.
[24] M. Clarich, *Manuale di Diritto Amministrativo* (Bologna: Il Mulino, 2019), 131.
[25] Consiglio di Stato, n 11 above, point 9.1

administration in its acts in order to assess the reliability, coherence, correctness of the methodology and the conclusions adopted.[26]

In this context, the Council of State interpreted its task as an entitlement and an obligation to check whether the decision of AIFA to forbid the off-label use of HCQ for COVID-19 patients had a solid scientific base, in the light of the limited available evidence.[27] In other words, the Council of State had to verify the latest scientific knowledge to understand whether the ban of the off-label use of HDQ was reasonable or not.

3. Scientific Criteria for the Solution

The Consiglio di Stato continues by illustrating that medical science has to indicate the most appropriate and safe treatment for a disease, and the dominant approach in accomplishing this task is the so-called *evidence based medicine* (EBM). Under EBM, the choice of treatment should be based on the best evidence of efficacy and on randomised controlled trials (RCT) which represent the so-called gold standard of medical research.[28]

Applying this scientific methodology in the COVID-19 pandemic, however, was (and is) not straightforward. The urgency of the situation did not permit the collection of significant and final findings on the best type of treatment, and specifically on the benefits of the use of hydroxychloroquine as a treatment for COVID-19 patients with a non-serious condition and in the early stages of the disease.

The Council of State therefore had to establish whether the technical discretion of AIFA in the application of the scientific laws made the suspension of the use of HCQ logical and proportionate. In other words, the Consiglio di Stato had to analyse whether the temporarily suspension of the administration of the drug in patients with non-serious symptoms responded to the necessity to ensure the most (1) *appropriate* and (2) *safe* therapy in the interest of public health.[29]

As to the *appropriateness* of the cure, the Consiglio di Stato observes that there were no medical studies demonstrating the *unquestionable inefficacy* of HCQ in the early treatment of COVID-19 patients (even if the results tend towards its inefficacy), and that most of the studies lack internal and external validity because of the urgent conditions in which the clinical trials were conducted.[30]

[26] ibid point 9.2.
[27] ibid point 10.
[28] ibid point 11.4.
[29] ibid point 12.
[30] ibid points 13-13.1. The court illustrates that the problem concerned, for example, the so called 'endpoint' of the studies, ie, the measure to consider the success or otherwise of the therapy: the most indicative endpoint for the use of HCQ in the early stage of the illness should have been 'how many of the treated patients needs to be treated in the hospital after the treatment', and not for example – as considered in many trials – the percentage of mortality or the number of days spent in hospital. In addition, the problem also lays in the difficulty of collecting reliable evidence on the

Against this background of uncertainty regarding the usefulness of HCQ, which was admitted – as the court stresses – by AIFA itself,[31] the appellants submitted to the courts an expert opinion based on several randomised studies which, on the contrary, suggested the efficacy of the treatment in a non-advanced stage of the illness.[32]

The Consiglio di Stato makes clear that if, on the one hand, it is beyond its competence to decide on the efficacy or otherwise of HCQ against COVID-19 in the early stage of the illness, on the other hand, it is up to the administrative court to point out that this uncertainty regarding the efficacy of HCQ is not sufficient from a legal point of view to justify the suspension of its possible use.[33] In the face of the limited experimental evidence available and of divergent medical opinions, the decision to suspend the use of the drug does not even allow testing the slightest efficacy of HCQ in patients in the early stage of the illness, and delays its experimentation to a point in the future in which it would be probably not be useful.[34] The strict and severe scientific methodology has to deal with the emergency of the pandemic situation and, lacking an alternative therapy, the use of a drug which could be even slightly useful could not be denied, unless its risks clearly outweigh its benefits.[35]

In this situation of uncertainty, the Consiglio di Stato maintains that the choice to use or not to use a drug should lie with the autonomy of the physician, with the informed consent of the patient, and not on a generalised and aprioristic ban on using it, based on a principle expressed 'in the name of a pure scientific model'.[36]

As to the requisite of the *safeness* of the treatment, the Consiglio di Stato illustrates that AIFA itself recognised that the last clinical trials seemed not to demonstrate a higher risk of toxicity (especially an increased risk of heart problems) and did not show any difference between patients who use it or not in terms of mortality.[37] As to the psychiatric disorders pointed out by the safety committee of EMA, these are related to higher doses of HCQ.[38]

On these premises, the Council of State could reach only one decision. Given the abovementioned considerations regarding the (limited) efficacy and the (apparent) safeness of the treatment, the court states that, on the basis of the available scientific knowledge and considering the possibility of any further

usefulness or not of HCQ in patients who were at home and at an early stage of the disease, because for example much of the data was collected by telephone or online.

[31] Consiglio di Stato, n 11 above, point 14.5.

[32] ibid point 14.5.

[33] ibid point 15.

[34] ibid point 16.

[35] ibid point 17.1.

[36] ibid point 17.2.

[37] ibid point 19. See also AIFA, 'Hydroxychloroquine in the treatment of adult patients with COVID-19', available at https://tinyurl.com/7fcb25se (last visited 30 June 2021).

[38] Consiglio di Stato, n 11 above, point 19.3.

investigation by the trial court and AIFA itself, the evaluation of the risks/benefits of the drug demonstrated that the suspension of the off-label use of HCQ and of its prescription by doctors, under their responsibility, in the treatment at home of COVID-19, was not reasonable.

On these grounds the claim of the doctors was accepted. On 23 December 2020, AIFA updated its recommendations on the use of hydroxycloroquine (HCQ) in patients with COVID-19.[39] The case is now being discussed in the trial court.

IV. The Intersection Between Science and Law from a Comparative Perspective

The reason why the order of the Council of State is interesting for lawyers is of course not related to the usefulness or not of HCQ as a treatment for COVID-19. What is interesting to analyse is the use by the judge of the scientific reasoning to decide the case. Even if the decision we are commenting on was adopted by an administrative court, the power and position of which vis-à-vis the parties are in many respects different from the ones of ordinary courts in civil proceedings, it is reasonable to assess the decision in question against the legal debate on science and law emerging in the context of civil procedure. Since these problems are shared by both the common law and civil law traditions, it is also useful to analyse how they are viewed and approached on the two sides of the Atlantic.

The legal debate on science and law focuses mostly on doubts about the ability of a judge to arrive at sound inferences from scientific or technical data. This is why it becomes crucial for legal scholars to study and understand the admissibility standards of access of evidence in judicial proceedings.[40]

Since the principal (albeit not the only) means through which scientific evidence enters the courtrooms is the expert testimony, the majority of criteria for the evaluation of scientific data before courts has been developed in relation to the standard for admitting scientific expert testimony in trials.

In this light, the first and most important example comes from the common law experience. The reference is to the US leading case *Daubert* v *Merrell Dow Pharmaceuticals, Inc.* 509 US 579 (1993) and the so-called Daubert Trilogy.[41]

[39] AIFA, 'AIFA recommendations on hydroxychloroquine', available at https://tinyurl.com/f64ksnfb (last visited 30 June 2021).

[40] From a comparative perspective see, for example, P. Monaco, *Sostanze tossiche e danni. Profili di diritto globale, europeo e nazionale* (Napoli: Editoriale Scientifica, 2020). For common law see, among many, D. Faigman et al, *Modern Scientific evidence: the law and science of expert testimony* (St. Paul: West Pub Co, 2010), III; P.C. Giannelli and E.L. Imwinkelried jr, *Scientic Evidence* (Newark: LexisNexis, 2007); in Italian: M. Taruffo ed, *La prova nel processo civile* (Milano: Giuffrè 2012); Id, *La semplice verità. Il giudice e la costruzione dei fatti* (Bari: Editori Laterza, 2009).

[41] *Daubert* v *Merrell Dow Pharmaceuticals*, Inc 509 US 579 (1993).

Daubert is a toxic tort case regarding the tragedy of Bendectin, a drug prescribed during pregnancy to reduce 'morning sickness'. Here the plaintiffs argued that the drug caused deformities

In Daubert, the Supreme Court of the United States proffered new standards for the admissibility of scientific evidence. In particular, the Supreme Court stated that judges had to consider

> (1) whether the theory or technique (…) has been tested; (2) whether it has been subjected to peer review or publication; (3) its known or potential error of rate (…); and (4) whether it has attracted widespread acceptance within a relevant scientific community.[42]

But perhaps the most important statement of the Daubert decision is the one in which the Supreme Court vested the judge with the role of gatekeeper of the trial, entrusted with ensuring that 'all scientific testimony or evidence admitted is not only relevant, but reliable'.[43] To stress again the importance of the topic of science and law, it is worth emphasising how revolutionary this judicial task is in an adversarial system like that of the US, where the proceedings are usually controlled by the parties and their attorneys, while the judges acts as a passive umpire, basing their decision on the evidence presented by the parties.[44]

As for the civil law tradition, the field of science in courtrooms is treated from a slightly different perspective. Despite the fact that the rules on the participation of experts in litigation differ across civil law jurisdictions, European legal systems converge on the fact that is the judge, and not the parties, as in the US adversarial model, who exercise the control on the evidence phase of the proceedings, and consequently also on the choice of and the tasks given to the expert, the main instrument through whom scientific evidence enters the courtroom. The main worry in continental Europe about experts' reports concern the evaluation of their competence.[45] That is why the gatekeeping role could belong to the judge

in children exposed to it while in utero. The problem was with the admissibility of expert testimonies sustaining the existence of a causation between the use of the drug and the deformities. On the Daubert decision see, among the many, M.A. Berger, 'The Supreme Court's Trilogy on the Admissibility of Expert Testimony', in M.A. Berger et al eds, *Reference Manual on Scientific Evidence* (Federal Justice Center, 2nd ed, 2000), 9; D. Bernstein, 'The Misbegotten Judicial Resistance to the Daubert Revolution' *Notre Dame Law Review*, 29 (2013); M.A. Berger, 'Upsetting the Balance Between Adverse Interests: The Impact of the Supreme Court's Trilogy on Expert Testimony in Toxic Tort Litigation' 64 *Law & Contemporary Problems*, 289-326 (2001). In Italian, P. Monaco, *Sostanze tossiche e danni* n 40 above, 84.

[42] *Daubert* v *Merrell Dow Pharmaceuticals*, Inc. 509 US 579, 592 (1993).

[43] ibid 589. After Daubert, two other decisions completed the framework of the discipline. The first, *General Electric Co.* v *Joiner* 522 US 136 (1997), regarded the standard for federal appellate courts to review the evidentiary determination of the lower court. The second decision, *Kumho Tire Company Ltd.* v *Carmichael* 526 US 137 (1999) considered the question of whether Daubert's reliability test was extended also to non-scientific expert testimony, in this case the plaintiffs' expert witness – an expert in tire industry, who testified that the defective car tire has caused the accident where one passenger died and others were injured. These three cases are commonly known in the legal debate as the Daubert Trilogy: M.A. Berger, 'The Supreme Court's Trilogy' n 41 above.

[44] P. Monaco, *Sostanze tossiche e danni* n 40 above, 92.

[45] P. Monaco, 'Scientific Evidence' n 2 above; P. Monaco, *Sostanze tossiche e danni* n 40 above, 271.

while appointing the expert and defining their duties,[46] as well as while checking the notions and methodology employed by the expert. In their role, it is said, the judge need not be an expert. But at this point, as noted by legal scholars,[47] a paradox appears. If judges need to be assisted by an expert because they lack the required specific knowledge, how can they have the ability to evaluate the soundness of the final technical report? The answer to this problem is that (perhaps) they should only deal with the knowledge of the necessary conditions under which information could be considered to possess scientific validity.[48]

Could these principles born and developed in the civil proceedings be valid when applied in front of an administrative court? We will try to answer this question in the next section.

V. The Administrative Court Dealing with Science

The administrative trial also represents a special laboratory to test the intersection between law and science, particularly because it is closely linked to the judicial review of the administrative action. In fact, an extensive debate between administrative courts and scholars has developed around the justiciability of so-called 'technical discretion'.[49] As is well known, administrations enjoy the technical discretion when decisions have to be taken on the basis of specific technical expertise which, when applied, presents profiles of uncertainty or questionability because it depends on divergent scientific opinions.[50] This is the case, for example, with AIFA forbidding the off-label use of HCQ.

As the litigation herein commented on makes clear, the judicial review of administrative technical acts raises a number of questions, such as: how far can

[46] See for all, M. Taruffo, 'La prova scientifica nel processo civile' *Rivista trimestrale di diritto e procedura civile*, 1110-1111 (2005).

[47] ibid 1079-1111.

[48] ibid 1110-1111.

[49] For M. Clarich, *Manuale* n 24 above, 130-131; S. Cognetti, 'Potere amministrativo e principio di precauzione fra discrezionalità tecnica e discrezionalità pura', in S. Cognetti et al eds, *Percorsi di diritto amministrativo* (Torino: Giappichelli, 2014), 131. In English: See G. della Cananea, 'Judicial Review of Administrative Action in Italy: Beyond Deference?', in G. Zhu ed, *Deference to the Administration in Judicial Review* (Cham: Springer, 2019), 271. On judicial review of technical discretion of Independent Administrative Authorities see G. Sigismondi, 'Il sindacato sulle valutazioni tecniche nella pratica delle Corti' *Rivista Trimestrale di Diritto Pubblico*, 705 (2015); G. De Rosa, 'La discrezionalità tecnica: natura e sindacabilità da parte del giudice amministrativo' *Diritto e Processo Amministrativo*, 513 (2013); A. Travi, 'Il giudice amministrativo e le questioni tecnico-scientifiche: formule nuove e vecchie soluzioni' *Diritto pubblico*, 439 (2004); P. Lazzara, *Autorità indipendenti e discrezionalità* (Padova: CEDAM, 2001).

[50] On the contrary, 'administrative discretion' involves choices of political nature. On this point seminal are the studies of M.S. Giannini, *Il Potere Discrezionale della Pubblica Amministrazione, Concetto e Problemi* (Milano: Giuffrè, 1939). More recently: M. Clarich, *Manuale* n 24 above, 130. On the technical discretion see D. de Pretis, *Valutazione amministrativa e discrezionalità tecnica* (Padova: CEDAM, 1995).

the technical discretionary power be scrutinised by the court? Is the judicial review limited to verifying compliance with the procedural rules or can the judge scrutinise more profoundly the decision of the authority? As discussed above, one of the defences of AIFA was that its decision to forbid the use of HCQ was the fruit of the expression of its technical discretion and therefore could not be subjected to judicial review, even more so in an ad interim proceeding.[51]

For a long time, the tendency of the Italian administrative judiciary was to deny the judicial review of choices stemming from a technical discretion.[52] But recently things have changed, as the Council of State pointed out in the decision we are commenting on.[53] The path changed in particular in 1999, with the leading case on the matter by the Council of State, section IV, of 9 April 1999, no 601.[54] In this decision, the Consiglio di Stato made it clear that it is true that judicial review is not possible if related to the direct evaluation of the public interest underlying the merits of the choice made by the authority (ie, administrative discretion properly intended), but on the contrary the control of the court is not excluded for assessments based on technical standard.[55]

From this perspective, the judicial review by the administrative judge is no longer limited simply to the existence of formal errors of assessment in the decision of the authority, but also includes the evaluation of whether technical assessments have been made following a rational credibility supported by scientific and technical argumentations correctly applied in the specific context.[56] This is the reason why the Consiglio di Stato, in its order 9070/2020, engaged in a profound analysis of the available scientific data supporting the decision of AIFA as regards HCQ. The court is entitled to check if the assessment adopted by the administrative authority is correct, in the light of the technical and scientific rules which have been applied.[57]

Once it is admitted that the administrative court has the power to evaluate scientific knowledge, we return to the problem we discussed above with regard to judges in civil courts: how can the judge have the ability to control a very

[51] See above, section 1. Point 8.1 of the Consiglio di Stato, n 1 above.

[52] M. Clarich, *Manuale* n 24 above, 130-131.

[53] See above, section 2.

[54] Consiglio di Stato 9 April 1999 no 601, available at www.giustizia-amministrativa.it. The case concerned the claim of a civil servant who maintained that his illness was caused by his administrative activity. The lower administrative court endorsed the respondent's argument that the claim was not substantiated on the basis of objective standards and that technical assessments made by the administration escaped judicial review. The Council of State reversed the decision of the lower court.

[55] See G. della Cananea, 'Judicial Review' n 49 above, 271.

[56] The importance of this decision is most evident in some very specific areas, such as for example the one of antitrust: see R. Chieppa, 'Il differente controllo del Giudice amministrativo sulle attività di regolazione e giustiziali delle Autorità amministrative indipendenti', in R. Chieppa et al eds, *Il controllo del giudice amministrativo sulla discrezionalità tecnica e, in particolare, sugli atti delle autorità indipendenti* (Milano: Giuffrè, 2009), 47.

[57] Consiglio di Stato, n 11 above, point 9.3.

specialised act, adopted by a highly qualified authority? An easy answer is that the administrative judge can and should do so with the support of experts, whose appointment is possible, when necessary, in administrative proceedings as well.[58] But, as we mentioned above, allowing the participation of experts in judicial proceedings means opening a Pandora's box of other problems, related to the ability of judges to govern and control such experts and the knowledge they provide in the litigation. What is certain is that this topic is also far from being settled in administrative proceedings.

VI. Conclusion

As we observed at the beginning of our analysis, the relationship between science and law has never been one of simple understanding. Owing to the increasing number of legal questions which can (or even have to) be resolved by resorting to scientific knowledge, in recent years the challenges posed by the use of scientific evidence became a hot topic in legal debates. Problems arising from scientific evidence are shared by the legal traditions and by the different areas of law.

As to the common law and civil law traditions, despite the divergence of attitudes and solutions proposed by the US and continental European legal systems, when we look at the operational results we find a surprising convergence. On the US side, Daubert not only fixed the standards for admitting scientific evidence, but also attacked one of the pillars of the US adversarial system, the traditionally passive role of the judge. After Daubert, both federal and state judges no longer neutrally umpire the proceedings, but actively intervene in the development of evidence and in the whole process itself. As to the situation in Europe, the situation seems, at the first glance, to be very different, since it is the judge, and not the parties, who exercises control over the evidentiary phase of the proceedings. However, if we dig a little deeper under this surface, we will discover that these divergences are less clear. Continental European judges do not only experience the same difficulties as their US colleagues in dealing with scientific problems, but, in the end, they also behave in the same way: they are the ultimate gatekeepers of the submission of scientific evidence in the courtroom.[59]

The same observations also apply to the context of the decision we are commenting on. The scientific and medical issues which exploded with the

[58] Art 63(5) of the Code of Administrative Procedure foresees the possibility for the judge to order any of the means of evidence provided by the Civil Procedure Code, included the assistance of an expert. On the expert in administrative proceedings see, for example, F. De Luca, 'I differenti tipi di misure cautelari', in F. Freni ed, *La tutela cautelare e sommaria nel nuovo processo amministrativo* (Milano: Giuffrè, 2011), 74; F. Cintioli, 'Consulenza tecnica d'ufficio e sindacato giurisdizionale della discrezionalità tecnica' *Consiglio di Stato*, 2371 (2001).

[59] For more extensive comment see P. Monaco, 'Scientific Evidence' n 2 above, 108.

COVID-19 pandemic have given rise to many related legal aspects which will end up on the desks of legal scholars and judges. Order no 9070 of 11 December 2020 of the Council of State represents a good example in this context.

In this decision the problems stemming from the uncertainty of medical reasoning are reflected in the difficult challenge for the judge who had to deal with complex scientific data. This was an occasion for the Consiglio di Stato to stress the boundaries of its role and make it clear that judicial review in the administrative context is not merely limited in the verification of the logic and consistency of the conclusions reached by the administration. On the contrary, when the assessment under the scrutiny of the court concerns the technical discretion of an administration (ie, discretion involving decisions based on technical and scientific expertise), the administrative judge is allowed to go further: they are entitled to verify whether the technical choices are based on a logical argumentation and a valid scientific knowledge. In other words, they act as the gatekeeper of scientific knowledge debated in front of them.

In conclusion, to secure the promise of effectively allowing only 'good' science to enter the courtrooms, it seems that the solution found in common law and civil law jurisdictions, as well as in front of civil and administrative courts, is converging: the gatekeeper of the scientific and technical knowledge is everywhere the judge. Whether these standards built and followed by courts actually work represents another big question, one that only time might answer.

Hard Cases

State Immunity and European Civil Procedural Law – Remarks on the Judgment of the CJEU of 7 May 2020, C-641/18, LG v Rina SpA and Ente Registro Italiano Navale

Bartosz Wołodkiewicz*

Abstract

In European procedural law, the existence of jurisdiction implies that a case must be heard by a court, which may be in collision with the obligation to decline jurisdiction when the defendant relies on state immunity. In its recent judgment of 7 May 2020, C-641/18, the Court of Justice of the European Union ruled on the relationship between state immunity and the exercise of jurisdiction resulting from the Brussels I Regulation. The ruling is noteworthy for a number of reasons. Its significance for the development of international law in the sphere of state immunity has already been noted in the literature. This paper analyses the consequences of the judgment for European civil procedural law by way of addressing two specific issues. The first one is a question about the relationship between state immunity and the concept of 'civil and commercial matters' which sets out the scope of the Brussels I Regulation. The second one is a question about the influence of state immunity on the exercise of jurisdiction granted by the Brussels I Regulation. Answering these questions will make it possible to determine the relationship between state immunity and the European civil procedural law.

I. Introduction

In its judgement of 7 May 2020, C-641/18 (*Rina*),[1] the Court of Justice of the European Union (CJEU) answered a question on the relationship between state immunity and the exercise of jurisdiction resulting from Brussels I Regulation.[2] The *Rina* case is thus an interesting example of the interaction between international law and European civil procedural law. The impact of this judgement on the ongoing international discussion on the scope of state immunity has already been addressed in literature.[3] Consequently, this article

* Assistant Professor of Civil Procedure, University of Warsaw. Proofreading financed under DIALOG no 0227/2018 a Program of the Polish Ministry of Science and Higher Education.

[1] Case C-641/18 *LG* v *Rina SpA and Ente Registro Italiano Navale*, Judgment of 7 May 2020, available at www.eur-lex.europa.eu.

[2] Council Regulation (EC) 44/2001 of 22 December 2000 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters [2001] OJ L12/1.

[3] A. Spagnolo, 'A European Way to Approach (and Limit) the Law on State Immunity? The Court of Justice in the RINA Case' 5 *European Papers – A Journal on Law and Integration*, 645

focuses on the significance of the judgement for European civil procedural law. Following an outline of the facts and the fundamental grounds of the ruling, this article will address what *Rina* means for the functioning of state immunity in European civil proceedings. This analysis of the CJEU judgement will seek answers to two specific questions. The first question concerns the relationship between state immunity and the concept of 'civil and commercial matters' which sets out the material scope of the Brussels I Regulation. The second question concerns the influence of state immunity on the exercise of jurisdiction granted by the Brussels I Regulation. The resulting answers will help to clarify the relationship between state immunity and European civil procedural law.

II. Facts of the Case and Ruling

1. Factual Background

The issues resolved by the ruling which is the focus of this paper emerged from an action brought by relatives of the victims and survivors of the sinking of the Al Salam Boccaccio '98 vessel before the Tribunale di Genova against *Rina SpA* and *Ente Registro Italiano Navale* (Rina Group companies) with their seats in Genoa. The claimants, who sought compensation, argued that liability for the sinking of MS Al Salam Boccaccio'98 could be attributed to the Rina companies, which completed the certification and classification of the vessel in question as delegates of the Republic of Panama for the purposes of obtaining the Panamanian flag for that vessel. The defendants pleaded immunity from jurisdiction, citing the principle of state immunity. The defendants argued that the vessel certification and classification completed by delegation of the Republic of Panama, as a manifestation of the sovereign powers of the delegating state, was covered by state immunity.

Throughout the course of this action both parties raised a number of arguments for and against its admissibility. The defendants invoked Art 94 of the UN Convention on the Law of the Sea,[4] which stipulates:

> Every State shall effectively exercise its jurisdiction and control in administrative, technical and social matters over ships flying its flag.

This is additionally supplemented by the International Convention for the Safety of Life at Sea,[5] which obliges states to carry out inspections of ships flying their flags, and at the same time authorises them to entrust inspections and surveys to organisations recognised by them. Therefore, the vessel classification

(2020).

[4] United Nations Convention of 10 December 1982 on the Law of the Sea (UNCLOS).

[5] The International Convention for the Safety of Life at Sea (SOLAS) concluded in London on 1 November 1974, Chapter I, Regulation 6.

and certification was carried out by the defendants in the exercise of obligations arising under international law. In this regard, the defendants acted on behalf of the Republic of Panama as a recognised organisation, although, at the same time, the classification and certification activities were carried out against payment and under a contract concluded with the shipowner. However, in respect of the operations carried out by Rina companies, the claimants submitted that state immunity does not cover technical activities, which are not connected to the sovereign powers of a state and, in consequence, are not acts undertaken in the exercise of public powers.

In light of these conflicting views, the Tribunale di Genova decided to stay the proceedings and to refer the matter to the Court of Justice for a preliminary ruling on the interpretation of Art 1(1) and Art 2(1) of the Brussels I Regulation in light of Art 47 of the Charter of Fundamental Rights (CFR)[6] and Art 6(1) of the European Convention on Human Rights (ECHR). The court sought to establish if an action for compensation brought against a private law entity which carries out classification and certification activities on behalf of a third state falls into the scope of the concept of 'civil and commercial matters' and, consequently, into the material scope of application of Brussels I Regulation.

2. Grounds for the Judgment

Having found the request for a preliminary ruling admissible, the CJEU began by distinguishing between the question of whether the matter falls into the scope of Brussels I Regulation, and the issue of whether, due to the jurisdiction immunity possibly enjoyed by the defendants, the jurisdiction resulting from Art 2(1) of the Brussels I Regulation may be exercised.

As regards the first question, the CJEU referred to its case-law and indicated that in order to determine whether a given matter is a civil and commercial matter, it is necessary to identify the legal relationship between the parties to the dispute, the basis of the action, and the rules governing the bringing of the action.[7] Within this context, the Court pointed out that matters in which one of the parties exercises public power (*acta iure imperii*) are excluded from the scope of 'civil and commercial matters' in the Brussels I Regulation.[8] Although in the case at hand vessel classification and certification were carried out upon delegation, on behalf of and in the interest of the Republic of Panama, and their purpose was to ensure the safety of ship passengers, these circumstances did not mean that the defendants enjoyed powers falling outside the scope of the ordinary legal rules applicable to

[6] Charter of the Fundamental Rights of the European Union [2012] *OJ C326/ 391.*

[7] Case C-302/13 *flyLAL-Lithuanian Airlines AS* v *Starptautiskā lidosta Rīga VAS and Air Baltic Corporation AS*, Judgment of 23 October 2014, paras 24, 25 and the case-law cited; available at www.eur-lex.europa.eu.

[8] n 2 above, paras 33-34.

relationships between private individuals, and only in the case that such powers were indeed possessed might it be possible to claim that they exercised public powers.[9] The Court attributed importance to the evaluation of competences of recognised organisations, such as the defendants, in carrying out vessel classification and certification.

Within the context of classification and certification, the CJEU noted that the activities of the Rina companies consist of checking the condition of the ship as per the relevant provisions of the law, which may result in the withdrawal of the certificate in cases in which the ship does not comply with those requirements. The ineligibility of a ship sail following the withdrawal of a certificate is the result of a sanction imposed by the law, and not of the decision-making power of those recognised organisations, whose is limited to verification activities and is technical in nature.[10] This position has been already expressed in an earlier ruling on the exercise by the Rina companies of public powers within the meaning of Art 51 of the Treaty on the Functioning of the European Union (TFEU). At the time, the CJEU found that recognised organisations are commercial undertakings performing their activities under competitive conditions of and do not have any power to make decisions connected with the exercise of public powers.[11] Nothing in the circumstances of the case at hand justified a departure from this conclusion within the interpretation of Art 1(1) of the Brussels I Regulation.

Ultimately, subject to the determination of certain matters by the referring court, the CJEU found that Art 1(1) of the Brussels I Regulation

> must be interpreted as meaning that an action for damages, brought against private-law corporations engaged in the classification and certification of ships on behalf of and upon delegation from a third State, falls within the concept of 'civil and commercial matters', within the meaning of that provision, and, therefore, within the scope of that regulation, provided that that classification and certification activity is not exercised under public powers, within the meaning of EU law.[12]

In light of the distinction made by the CJEU, at the stage of evaluating the scope of Brussels I Regulation, there is no need for an examination as to whether or not the party has immunity from jurisdiction. The case is qualified from the perspective of the 'civil and commercial matters' concept, which is subject to independent interpretation. Only in the event of finding that provisions of European civil law procedure apply to this dispute does it become necessary to examine the plea of immunity from jurisdiction. Having found that the dispute

[9] ibid paras 39-42.

[10] ibid para 47.

[11] Case C-593/13 *Presidenza del Consiglio dei Ministri and Others* v *Rina Services SpA and Others*, Judgment of 16 June 2015, para 16-21, available at www.eur-lex.europa.eu.

[12] n 2 above, para 60.

falls within the scope of the Brussels I Regulation, the Court then moved on to an examination of whether the defendant Rina companies, as private law entities, could invoke state immunity in regard to the vessel classification and certification. This issue, although addressed with caution in the judgment, has been broadly discussed in the opinion of the Advocate General, to which the grounds for the CJEU judgment refer approvingly.[13]

The point of departure for the deliberations was the concept of relative state immunity, which is based upon the distinction, pursuant to the international law provisions concerning state immunity, between *acta iure imperii* and *acta iure gestionis*. Within this context, the CJEU found, in accordance with the opinion of the Advocate General, that the principle of state immunity is governed by customary international law.[14] This justified the establishment by the CJEU of whether or not there exists a rule of international customary law which admits the invocation of immunity from jurisdiction by recognised organisations carrying out classification and certification.

The existence of a rule of customary international law is only possible where a given practice actually exists and is accepted as a law (*opinio iuris*). Findings in this respect are to be found primarily in the opinion of the Advocate General,[15] based on which the CJEU held that immunity from jurisdiction of private law entities is not universally recognised in relation to vessel classification and certification.[16] This interpretation was also reinforced by the interpretation of recital 16 of Directive 2009/15,[17] the wording of which indicates that recognised organisations, such as the defendants, do not enjoy immunity from jurisdiction, which only states may invoke. In the Court's view, this recital confirms the intention of the EU legislator to give a limited scope to its interpretation of the customary international law principle of immunity from jurisdiction with regard to the classification and certification of ships. This set of circumstances had led the Court to conclude that

> the principle of customary international law concerning immunity from jurisdiction does not preclude the national court seised from exercising the jurisdiction provided for by that regulation in a dispute relating to such an action, where that court finds that such corporations have not had recourse to public powers within the meaning of international law.[18]

[13] ibid para 57.

[14] Case C-641/18 *LG* v *Rina SpA and Ente Registro Italiano Navale*, Opinion of Advocate General Szpunar of 14 January 2020, paras 37-39, available at www.eur-lex.europa.eu.

[15] ibid paras 108-128.

[16] n 2 above, para 57.

[17] Directive of the European Parliament and of the Council of 23 April 2009 on common rules and standards for ship inspection and survey organisations and for the relevant activities of maritime administrations [2009] OJ L131/47.

[18] n 2 above, para 60.

III. Consequences for European Civil Procedural Law

1. General Remarks

The scope of application of European civil procedural law regulations within the context of interpretation of the 'civil and commercial matters' concept is a recurring issue in the Court's rulings. While the problem with qualifying a matter as 'civil and commercial' usually occurs in cases in which one of the parties to proceedings is a state or a state authority, situations in which the CJEU must assess the influence of immunity from jurisdiction on this qualification are rare.

When ruling in the *Lechouritou* case,[19] which was considered in light of the Brussels Convention, the CJEU found that an action brought by a natural person for compensation in respect of loss or damage against a foreign state as bearing civil liability for acts and omissions of its armed forces does not fall within the scope of 'civil and commercial matters', because acts perpetrated by armed forces, even when illegal, are acts carried out in the exercise of public powers. An exclusion of *acta iure imperii*, earlier present in other European civil procedural law, was added to the effective Brussels I-bis Regulation as a result of this ruling.[20] In the later *Mahamdia*[21] case, the CJEU held that immunity from jurisdiction of a state does not preclude the application of Brussels I in a case in which an employee demands compensation from a foreign state and questions termination of the employment contract when his duties did not form part of the exercise of public powers. In both cases, the consequences of invoking state immunity were assessed from the perspective of the scope of its application and exclusion from it of matters in which one of the parties exercised public powers.

In contrast, the interpretation made in *Rina* extends beyond the understanding of the concept of 'civil and commercial matters'. The Court focused on the controversial issue of the position of state immunity within the regime of Brussels I Regulation. With the exception of recital 14 of Regulation no 2201/2003,[22] immunity from jurisdiction is neither regulated nor cited in the provisions of European civil procedural law. Thus, the establishment of the relationship between immunity and jurisdiction is left primarily to case-law and legal scholars. Any subsequent CJEU judgment would be of great importance for the ongoing discussion in this regard, but the considerations presented in *Rina* make this judgment crucial for shaping the approach of the European civil

[19] Case C-292/05 *Eirini Lechouritou and Others* v *Dimosio tis Omospondiakis Dimokratias tis Germanias*, Judgment of 15 February 2007, available at www.eur-lex.europa.eu.

[20] B. Hess, 'The Application of the Brussels Ibis Regulation in the EU Member States', in G. Van Calster and J. Falconis eds, *European Private International Law at 50. Celebrating and Contemplating the 1968 Brussels Convention and Its Successors* (Cambridge: Intersentia, 2018), 35.

[21] Case C-154/11 *Ahmed Mahamdia* v *People's Democratic Republic of Algeria*, Judgment of 19 July 2012, available at www.eur-lex.europa.eu.

[22] Council Regulation (EC) 2201/2003 of 27 November 2003 concerning jurisdiction and the recognition and enforcement of judgments in matrimonial matters and the matters of parental responsibility, repealing Regulation (EC) 1347/2000 [2003] OJ L338/1.

procedural law to state immunity.

2. Civil and Commercial Matters

Before moving on to discussing the issues that make *Rina* of key significance for the placement of state immunity within European civil procedural law, I would first like to address the part of the judgment in which the CJEU only applied an existing concept, in order to provide a background against which to assess the ruling on qualifying the dispute as a 'civil and commercial matter' within the meaning of Art 1(1) of the Brussels I Regulation.

The Court's settled case-law provides that this concept is subject to independent interpretation, which means that it should be understood without references to *lex fori* or *lex causae*.[23] In practice, qualification of a given dispute as a civil and commercial matter is not always unambiguous, to which the wealth of case-law concerning this concept offers abundant testimony.[24] Of significance to its understanding is also the exclusion of tax, customs and administrative matters from civil and commercial matters; cases concerning state liability for acts and omissions committed in the exercise of public powers (*acta iure imperii*) are also excluded. The CJEU confirmed the existence of the latter exclusion in light of both the Brussels I Regulation and of the earlier Brussels Convention.[25] It has been expressed in the currently effective Art 1(1) of the Brussels I-bis Regulation.

The understanding of the *acta iure imperii* exclusion gives rise to doubts, especially in borderline cases. This is well illustrated by *Rüffer*,[26] which hinged on the qualification of a claim in connection with the costs of removing a shipwreck that had sunk on an international waterway. The obligation to remove the wreck arose out of an international agreement in which Germany conferred upon the Netherlands the exercise of river police functions. In the Court's view, the Netherlands entered this dispute in connection with the exercise of public authority, as the costs of removal of the wreck had been incurred in the performance of obligations under international law. Based on the foregoing, the Court held that the case did not fall into the scope of civil and commercial matters, even though its subject was the payment of a given amount of money, and the claim had been examined by a court. Of deciding significance was the fact that the claim had arisen out of the exercise of a public power.

[23] Case C-29/76 *LTU Lufttransportunternehmen GmbH & Co. KG* v *Eurocontrol*, Judgment of 14 October 1976, available at www.eur-lex.europa.eu.

[24] For an overview of the CJEU's case-law on the concept of "civil and commercial matter" see eg M. Illmer, in A. Dickinson and E. Lein eds, *The Brussels I Regulation Recast* (Oxford: Oxford University Press, 2015), 60-61; P. Rogerson, in U. Magnus and P. Mankowski eds, *Brussels I bis Regulation. Commentary* (Köln: Verlag Dr. Otto Schmidt KG, 2016), 63-70.

[25] B. Hess, 'The Application' n 21 above, 35.

[26] Case C-814/79 *Netherlands State* v *Reinhold Rüffer*, Judgment of 16 December 1980, available at www.eur-lex.europa.eu.

In this context, *Pula Parking*[27] provides an insight. The facts of this case speak for a practice of using the Brussels regime for the assertion of claims against private persons by public authorities.[28] The case concerned an enforcement proceeding brought by a publicly-owned company in Pula (Croatia), against a German resident, for the recovery of an unpaid 'parking debt'. The administration of public parking and the collection of parking fees had been delegated to the company by that public authority.[29] In *Pula Parking*, the Court observed that the conferral or delegation of powers by public authority does not imply that those powers are exercised *iure imperii*.[30] Key to this concept is the nature of the acts and how the powers were exercised.[31] In this regard, the Court held that the 'parking debt' was a 'consideration for a service provided' by the company, and that the relationship between the parties was contractual.[32] Based on this, the case fell into the scope of civil and commercial matters. In consistency with the case law of the CJEU,[33] *Pula Parking* may be viewed as an example of a rather broad interpretation of the scope of application of Brussels I-bis regulation.[34]

In light of this, doubts may arise as to the qualification adopted in *Rina*, as the defendants carried out acts in order for a third state to fulfil its obligations of international character, as was the case in *Rüffer*. The difference, as the Advocate General observed, lies in the exercise of public powers. While in *Rüffer*, the public authority that brought the action exercised the functions of river police and acted *iure imperii* with regard to the vessels, it was held in *Rina* that recognised organisations, such as the *Rina* companies, had no decision-making powers. The weight of the *Rina* case, then, rests not so much on the interpretation of the concept of 'civil and commercial matters' and of the delineation of the boundaries of the *acta iure imperii* exclusion, but rather on qualifying vessel classification and certification as the exercise of public powers.

This issue was previously examined by CJEU within the interpretation of Art 51 of TFEU,[35] which employs the concept of the 'exercise of official authority'. In *Rina Services,* it was held that as regards certification activities, the defendant companies perform their activities under competitive conditions and do not

[27] Case C-551/15 *Pula Parking d.o.o.* v *Sven Klaus Tederahn*, Judgment of 9 March 2017, available at www.eur-lex.europa.eu.

[28] Case C-172/91 *Volker Sonntag* v *Hans Waidmann*, Judgment of 21 April 1993; Case C-266/01 *Préservatrice foncière TIARD SA* v *Staat der Nederlanden*, Judgment of 15 May 2003; Case C-433/01 *Freistaat Bayern v Jan Blijdenstein*, Judgment of 15 January 2004; available at www.eur-lex.europa.eu.

[29] n 28 above, para 29.

[30] ibid para 35; see also Advocate General Szpunar at 79.

[31] P. Rogerson, in U. Magnus and P. Mankowski eds, *Brussels I Regulation. 2nd Revised Edition* (Münich: Sellier European Law Publishers, 2012), 55.

[32] n 28 above paras 35, 39.

[33] H. Roth, 'Vollstreckungsbefehle kroatischer Notare und der Begriff „Gericht" in der EuGVVO und der EuTVTO' *Praxis des Internationalen Privat- und Verfahrensrechts*, 41, 42 (2018).

[34] B. Hess, *The Application* n 21 above, 36.

[35] Treaty on the Functioning of the European Union [2012] OJ C326/47.

have any power to make decisions connected with the exercise of public powers.[36] Regardless of the qualification in the perspective of Art 1(1) of the Brussels I Regulation, ensuring the coherence of the system was an argument made in favour of sustaining this view in *Rina*.

Within the discussed scope, the *Rina* case does not contribute a great deal to the interpretation of the concept of 'civil and commercial matters'. In this part, it boils down to the application of existing concepts to unusual factual circumstances. The judgment, however, remains compatible with the interpretation of the *acta iure imperii* exclusion in the CJEU case-law.

3. Immunity from Jurisdiction and the Concept of 'Civil and Commercial Matters'

In the European procedural law, the existence of jurisdiction implies that a case must be heard by a court.[37] A member state's obligation to grant legal protection may be waived, but only in cases explicitly provided for in individual EU regulations.[38] At the same time, state immunity precludes a ruling on the merits of the case in an action brought to a foreign court, owing to observance of international law obligations. This may lead to a conflict between the obligation to exercise jurisdiction, resulting from the provisions of European civil procedural law, and the obligation to decline jurisdiction when the defendant enjoys immunity, as resulting from international law.

Thus far, the existing case-law of the Court has not provided clear answers as to how such a conflict can be solved. In *Lechouritou*, the Court found that

> a legal action brought by natural persons in a Contracting State against another Contracting State for compensation in respect of the loss or damage suffered by the successors of the victims of acts perpetrated by armed forces in the course of warfare in the territory of the first State[39]

is not a civil matter. By holding that acts perpetrated by armed forces are a manifestation of state sovereignty, the Court found that that the case did not fall into the scope of the Brussels Convention, which concerns civil matters. In this way, the conflict between immunity from jurisdiction and the exercise of jurisdiction was avoided by finding that the case was excluded from the scope of the Brussels Convention.

[36] n 12, paras 16-21.

[37] A. Layton, H. Mercer, *European Civil Practice* (London: Thomson/Sweet & Maxwell, 2nd ed, 2004), I, 373.

[38] CJEU ruled that that the Brussels Convention precludes a court of a Contracting State from declining the jurisdiction on grounds of *forum non conveniens* doctrine, see Case C-281/02 *Andrew Owusu* v *N.B. Jackson, trading as "Villa Holidays Bal-Inn Villas" and Others*, Judgment of 1 March 2005, available at www.eur-lex.europa.eu.

[39] Case C-292/05 *Eirini Lechouritou and Others* v *Dimosio tis Omospondiakis Dimokratias tis Germanias*, Judgment of 15 February 2007, available at www.eur-lex.europa.eu.

The ruling in *Lechouritou* was not tantamount to acknowledging that state immunity affects the qualification of the case, leading *ipso iure* to its exclusion from the scope of the Brussels Convention. However, in a theoretical approach, this is one of the possible solutions of the conflict at hand. It has been noted both in the case-law of member states[40] and in the literature,[41] that, in general, the reliance by a defendant on immunity from jurisdiction affects the qualification of the case as civil and commercial, to the effect that it is excluded from the scope of a given regulation. An alternative opinion has also emerged, according to which immunity from jurisdiction affects the exercise of jurisdiction resulting from the provisions of a given regulation, but not the scope of the regulation. The difference between these two approaches boils down to how justify the conclusion that the provisions of European civil procedural law, while granting national jurisdiction to the courts of the Member States, do not preclude the option of declining the examination of the case on merit on the grounds of immunity from jurisdiction.[42]

The later *Mahamdia* case, in which it was concluded that state immunity is not excluded in disputes concerning acts carried out *iure gestionis*,[43] did resolve doubts. As the Advocate General observed, in *Mahamdia* it was only ruled that

> once it has been established that immunity from jurisdiction does not preclude the application of that regulation, the latter must, *a fortiori*, apply in the dispute.

Leaving aside evaluations of this judgement, it must be noted that the observation made by the Advocate General indicated that the CJEU was not bound in *Rina* by its earlier ruling as to the influence of state immunity on the scope of Brussels I Regulation. This was of great significance, given that the Advocate General proposed a clear break with the position connecting state immunity with the scope of application. The opinion of the Advocate General emphasized that legislators may adopt rules governing jurisdiction with regard to disputes in which one of the parties relies on immunity from jurisdiction, and that international law only requires that jurisdiction should not be exercised towards such a party against its will.[44] If the European civil procedural law adheres to this position, it is not necessary to define the scope specifically taking into account the issue of state immunity.

The concept presented by the Advocate General was acknowledged in the

[40] *Grovit* v *De Nederlandsche Bank NV and Others* [2005] EWHC 2944 (QB), [2006] 1 WLR 3323.

[41] P. Rogerson, in U. Magnus and P. Mankowski eds, *Brussels I Regulation. Commentary* (Köln: Sellier, 2007), 51.

[42] P. Grzegorczyk, *Immunitet państwa w postępowaniu cywilnym* (Warszawa: Wolters Kluwer, 2010), 598.

[43] n 22 above, para 55.

[44] n 15 above, paras 41.

distinction between two issues made in *Rina*. The first issue was the interpretation of the concept of 'civil and commercial matters' within the context of vessel classification and certification in order to establish whether the dispute falls into the scope of the Brussels I Regulation. The second issue was an evaluation of the influence of immunity from jurisdiction on the exercise of jurisdiction conferred by this regulation. Therefore, already at the onset of its considerations, the CJEU rejected the position that the concept of 'civil and commercial matters' coincides with the negative scope of immunity from jurisdiction. In this respect, *Rina* presents a new approach to the relationship between state immunity and jurisdiction in European civil procedural law.

In light of *Rina*, the facts that lend state immunity to a party are examined during the determination of whether or not the dispute is a 'civil and commercial' matters, and may lead to the conclusion that the case does not fall within the scope of Brussels I Regulation. In such event, the case will be excluded from the application of the regulation due to the *acta iure imperii* exception. This, however, does not result from state immunity, but rather from a qualification made within the framework of European civil procedural law for the purpose of deciding whether the matter falls within the application scope of a given regulation.

In practice, the *acta iure imperii* exception excludes most cases in which state immunity must be accounted for from the scope of application of a given regulation.[45] This results from the distinction between *acta iure imperii* and *acta iure gestionis* made in respect of state immunity. The CJEU settled case-law presents the view that state immunity, in the present legal circumstances of international law, is not absolute, and only applies to acts carried out in exercise of public authority. In areas in which a state acts *iure gestionis*, it is subject to being sued. As a result, both for the purposes of determining the scope of state immunity and the scope of application of the provisions of European civil procedural law, a similar criterion is applied to the exercise of public authority. This criterion, however, serves different purposes and refers to other legal orders, so its application may lead to different results.[46] The CJEU stressed this difference for the first time in *Rina*, stating that, with regard to immunity from jurisdiction, it is necessary to examine whether the defendant acts *iure imperii* in the light of the provisions of international law and, with regard to the scope of application, in the light of the independently interpreted concept of 'civil and commercial matters'.

The CJEU judgment correctly defines the relationship between state immunity and the scope of European civil procedural law. The mere possibility of a case falling within the scope of this regulation does not in itself breach the limitations as to the potential outcomes of the case resulting from state

[45] M. Stürner, 'Staatenimmunität und Brüssel I-Verordnung – Die zivilprozessuale Behandlung von Entschädigungsklagen wegen Kriegsverbrechen im Europäischen Justizraum' *IPRax*, 197, 203 (2008).

[46] ibid 203; P. Grzegorczyk, n 43 above, 602.

immunity. In some cases the defendant may enjoy state immunity, but the case will be qualified as 'civil and commercial' within the meaning of European civil procedural law nonetheless. Within this context, some authors have pointed to disputes concerning employment contracts with persons employed in diplomatic posts, which meet the criteria of a civil case and in which, pursuant to Art 11 of the New York Convention, state immunity may be at play, depending on the plea.[47] In such event, the regulation would apply to the subject matter of the dispute, but at the same time the court should refuse to rule on the merits of the case because of the state immunity. Thus, the approach is both theoretically sound and practically applicable.

In conclusion, the resolution of doubts regarding the impact of immunity from jurisdiction on the scope of regulations of the Brussels-Lugano system, provided by the ruling in *Rina*, was a significant step in placing state immunity within the European civil procedure. The position adopted in *Rina* was confirmed by CJEU in its judgment of 3 September 2020 C-186/19 (*SHAPE*),[48] the Court applied the same test to define civil and commercial matters and accepted that the material scope of European civil procedural law is not defined with the concept of immunity from jurisdiction. SHAPE suggests the emergence of a clear trend in the case law.[49]

4. Declining of Jurisdiction by Reason of State Immunity

In *Rina*, the issue of the impact of state immunity on the exercise of jurisdiction conferred by European civil procedural law was merely outlined. The CJEU held that defendants could not invoke state immunity if the national court found that the defendants did not exercise the prerogatives of public authority under international law, as unambiguously evidenced by criteria set out in the grounds. Thus, in the *Rina* case, immunity from jurisdiction did not preclude exercise of jurisdiction. Nevertheless, the CJEU admitted such a possibility in its considerations, with two important remarks.

Firstly, the CJEU adopted the position that the principles that are a manifestation of customary international law form a part of the EU legal order. It reflects not only a general assumption that the EU is bound by general international law, but also a position based on the 'fundamental rules of customary international law' above EU secondary legislation.[50] This confirms

[47] P. Grzegorczyk, n 43 above, 603.

[48] Case C-186/19 *Supreme Site Services GmbH and Others* v *Supreme Headquarters Allied Powers Europe*, Judgment of 3 September 2020, available at www.eur-lex.europa.eu.

[49] See also G. Cuniberti, 'Sovereign Immunities and the Scope of the Brussels Ibis Regulation after Rina and SHAPE', available at https://tinyurl.com/23rfcd93 (last visited 30 June 2021).

[50] Case C-162/96 A. Racke GmbH & Co. v Hauptzollamt Mainz, Judgment of 16 June 1998, paras 45-51, Case C-366/10 Air Transport Association of America and Others v Secretary of State for Energy and Climate Change, Judgment of 21 December 2011, paras 101, 107-110; Case C-266/16 Western Sahara Campaign UK v Commissioners for Her Majesty's Revenue and Customs and

that it is necessary to decline the exercise of jurisdiction conferred by provisions of European civil procedural law when this obligation results from immunity from jurisdiction granted by international law. Within this context, it seems surprising that the CJEU did not address Art 71 of the Brussels I Regulation, which stipulates that this regulation does not affect any conventions to which the Member States are parties and which, in relation to particular matters, govern jurisdiction or the recognition or enforcement of judgments.

It has been noted in the literature that Art 71 of the Brussels I Regulation governs relations between this regulation and, among others, the Basel Convention on state immunity and Vienna conventions on diplomatic and consular immunity.[51] A classification of these treaties as conventions that take precedence over the Regulation is one of the ways in which the position of immunity from jurisdiction within European civil procedural law could be determined.[52] However, it was not possible to apply the Basel Convention to the *Rina* case, as neither Italy nor the Republic of Panama are parties to it.

The source of possible immunity for defendants was to be international custom, which was difficult to fit under Art 71 of the Brussels I Regulation, both in terms of literal and purposive interpretation. As the Advocate General observed, this would lead to a 'freezing' of customary international law in the state it was in when that Brussels I Regulation was adopted. For this reason, the Advocate General proposed that the relationship between immunity from jurisdiction and jurisdiction should be examined in the light of the relationship between EU law and international law.

The approval that the Court expressed for the Advocate General's approach is not tantamount to rejection of the assertion of Art 71 of the Brussels I Regulation that the Regulation does not affect any international treaties to which a state is a contracting party, with reference to immunity from jurisdiction. In any case, to attribute this opinion to the CJEU based on *Rina* would be premature. For this reason, this judgment will not resolve the dispute between those who argue that the issue of immunity from jurisdiction remains outside the scope of the Brussels I Regulation altogether,[53] and those who maintain that

Secretary of State for Environment, Food and Rural Affairs, Judgment of 27 February 2018, paras 47, 58 available at www.eur-lex.europa.eu.

[51] See ie B. Hess, 'European civil procedure and public international law', in U. Fastenrath et al eds, *From Bilateralism to Community Interest: Essays in Honour of Bruno Simma* (Oxford: Oxford University Press, 2011), 936; J.J. Fawcett et al, *Cheshire, North, Fawcett, Private International Law* (Oxford: Oxford University Press, 2008), 510; P. Mankowski, in U. Magnus and P. Mankowski eds, *Brussels I bis Regulation. Commentary* (Köln: Verlag Dr. Otto Schmidt KG, 2016), 1058.

[52] For an overview of this approach, see S. Rinke, Schadensersatzklagen gegen Staaten wegen schwerer Menschenrechtsverletzungen im Europäischen Zivilprozessrecht. Zugleich ein Beitrag zum Verhältnis der EuGVVO zur Staatenimmunität (Berlin: Berliner Wissenschafts-Verlag, 2016), 182-200.

[53] J. Kropholler, *Europäisches Zivilprozessrecht. Kommentar zu EuGVO, Lugano-Übereinkommen und Europäischem Vollstreckungstitel* (Frankfurt am Main: Verlag Recht und

the issue of the exercise of jurisdiction in the event of plea of immunity from jurisdiction can be resolved under Art 71 of the Brussels I Regulation.[54]

Secondly, the CJEU stated that in the event of the referring court upholding the plea relating to immunity from jurisdiction, it should ensure that a refusal to exercise jurisdiction would not deprive the claimants of access to a court. This is one of the guarantees under Art 47 of the Charter of Fundamental Rights. In this way, by allowing the refusal of jurisdiction for reasons of observance of international law obligations, the Court defined the boundary by noting the necessity of accounting for the upholding of fundamental rights. This issue arose in connection with the court's reference in its request for a preliminary ruling to Art 47 of the CFR and Art 6(1) of the ECHR, which was an expression of a doubt as to whether the possible recognition of immunity does not violate the right to a court. This was broadly addressed by the Advocate General, who stressed that immunity from jurisdiction, which is a manifestation of international law obligations, and which the European Union must observe, may be in conflict with the obligation to observe fundamental rights. It was not examined in *Rina*, where it was found that immunity from jurisdiction does not apply to vessel classification and certification carried out by recognised organisations, so far as these activities are not performed in the exercise of public powers. Nevertheless, the CJEU *obiter dictum* expressed the view that upholding the plea of immunity from jurisdiction, leading to refusal of jurisdiction, must remain in compliance with European fundamental procedural rights.[55]

The standard of European fundamental rights operates within the framework of application of the EU law, which means that it is also in line with the interpretation and application of European civil procedural law.[56] The CJEU case-law includes references to guarantees stipulated under Art 47 of the CFR both within the context of jurisdictional regulations, recognition and enforcement of rulings in relations between Member States, and in the interpretation and application of national civil procedural law, when they supplement the regulations of the EU in the necessary scope.[57] The *Rina* case links this standard to the

Wirtschaft, 9th ed, 2009), 120-121; R. Geimer, in. R. Geimer, R.A. Schütze eds, *Europäisches Zivilverfahrensrecht* (München: C.H. Beck, 3rd ed, 2010), 122; P. Grzegorczyk, n. 43 above, 607-608; B. Wołodkiewicz, *Ustanowienie jurysdykcji krajowej przez wdanie się w spór na podstawie rozporządzenia Bruksela I bis* (Warszawa: Wolters Kluwer, 2020), 38. See also Oberste Gerichtshof (OGH) 14 May 2001, 4 Ob 97/01w, ECLI:AT:OGH0002:2001:RS0115353, with approbation of W. Obwexer, 'Staatenimmunität innerhalb der EU', *ecolex* (2002), 57-59.

[54] See works cited in n 51 above.

[55] n 2 above, para 55.

[56] B. Hess, *Europäisches Zivilprozessrecht* (Heidelberg: C.F. Müller, 2010), 136; K. Weitz, 'Wpływ prawa Unii Europejskiej na krajowe prawo procesowe cywilne' *Kwartalnik Prawa Prywatnego*, 297, 305-207 (2019).

[57] Case C-7/98 *Dieter Krombach* v *André Bamberski*, Judgment of 28 March 2000; Case C-327/10 *Hypoteční banka a.s.* v *Udo Mike Lindner*, Judgment of 17 November 2011; Case C-292/10 *G* v *Cornelius de Visser*, Judgment of 15 March 2012; Case C-112/13 *A* v *B and Others*, Judgment of 11 September 2014, available at www.eur-lex.europa.eu.

application of international law provisions in conjunction with European civil procedural law.

The CJEU has held that the principle of state immunity and the attendant necessity to decline jurisdiction may collide with the right to a fair trial stipulated by Art 47 of the CFR. The point here is not to restrict the boundaries of state immunity by deciding on the admissibility of the invocation of immunity in cases involving serious human right abuses perpetrated by state functionaries in the exercise of public powers, an issue that the CJEU confronted in *Lechouritou*. The position of the CJEU addressed the tendency to view the refusal of jurisdiction due to state immunity as a breach of the right to a fair trial, which is demonstrated in the case-law of some Member States and of the European Court of Human Rights.[58]

This position embodies the view that the granting of immunity to a foreign state, even though a 'legitimate means of complying with international law to promote comity and good relations between states',[59] may nonetheless be found to be a disproportionate restriction of the right of access to a court.[60] Within this approach, access to a court requires that the court exercise its jurisdiction unless international law requires that immunity be granted to the foreign state. The declining of jurisdiction on the basis of state immunity is, therefore, an exceptional circumstance, and so the granting of immunity must take place within the strict limits of the requirements of international law.[61] Against this background, the reference of the CJEU to the premise that a national court applying EU law in the form of Regulation 44/2001 must comply with the requirements of Art 47 of the CFR is an undeniable simplification. This premise can be successfully applied to national procedural law, but with regard to obligations of international law, it must be considered in the context of the European Union's obligation under Art 3(5) TEU to contribute to the strict observance and development of international law. If a conflict between fundamental rights and international obligations in the field of state immunity should arise, its resolution may require an adaptation of the rules relating to the relationship between national law and European law.

In the light of EctHR case law, one of the available solutions is to adopt a requirement that the exercise of state immunity should respect the limits set by international law. In one case between an individual and an international

[58] A Sanger, 'State Immunity and the Right of Access to a Court under the EU Charter of Fundamental Rights' 65 International and Comparative Law Quarterly, 213 passim (2016).

[59] See, in particular, Eur. Court H.R, *Al-Adsani* v *United Kingdom*, Judgment of 21 November 2001, para 56; *Jones and Others* v *United Kingdom*, Judgment of 14 January 2014, paras 186-189, available at www.hudoc.echr.coe.int.

[60] R. Pavoni, 'The Myth of the Customary Nature of the United Nations Convention on State Immunity: Does the End Justify the Means?', in A. Van Aaken and Iluia Motoc eds, *The European Convention on Human Rights and General International Law* (Oxford: Oxford University Press, 2018), 265.

[61] n 59 above, 213, 222-223.

organisation, the ECtHR also considered whether claimants would have access to an alternative court if immunity from jurisdiction was recognised.[62] It seems that in formulating the requirement for the national court to ensure that an alternative forum is available, the CJEU made a reference to this concept as mentioned in the opinion of the Advocate General.[63] In establishing this limit for state immunity, the CJEU has not stipulated the conditions for its application.

It seems appropriate to at least consider whether access to the courts of a third state, which is not a Member State of the European Union, is sufficient. The circumstances of *Rina* indicate that the CJEU takes this possibility into account.[64] However, the question arises as to whether there are conditions that a third state court should meet in order for it to be sensible to require the claimant to use a reasonable alternative forum. In particular, this is relevant to whether a court of a Member State should be satisfied that the alternative forum will ensure the implementation of other standards under Art 47 of the CFR. Additionally, this may concern the reasonable length of proceedings or the independence of the judiciary. Undoubtedly, the ordinary burdens of seeking legal protection abroad should not preclude the requirement to initiate proceedings before the courts of a foreign country. This leads to the conclusion that a court of a Member State should at the same time make sure that referring claimants to an alternative forum will not lead to a denial of justice.[65]

It is difficult to escape the impression that the impact of permitting the examination of compliance of the declining of jurisdiction with Art 47 of the CFR was not thoroughly considered by the CJEU. The *Rina* case merely presents this problem, but stops short of solving it. To a certain extent, the failure of the CJEU to explain how it assessed whether a declining of jurisdiction interferes with the right of access to a court is explained by the fact that immunity from jurisdiction was not an obstacle to the exercise of jurisdiction in *Rina*. In any case, in stating that the court should make sure that upholding the plea of immunity from jurisdiction does not deprive claimants of access to court, the CJEU pointed to another limitation to which immunity from jurisdiction is subject. However, the limits of this exception have been left undefined.

In closing, it remains to be noted that *Rina* does not give space to addressing an issue that is raised in the doctrine of some Member States, namely, the question of the order in which state immunity and jurisdiction immunity is to be examined. Some commentators have expressed the opinion that priority of examination should be given to state immunity, as it precludes any action in the case. From the perspective of European civil procedural law, these questions

[62] Eur. Court H.R *Waite and Kennedy* v *Germany*, Judgment of 18 February 1999, para 68, available at www.hudoc.echr.coe.int.

[63] n 15 above, fn 106.

[64] ibid para 153.

[65] In the literature, this problem is considered in relation to negative conflicts of jurisdiction, among other things, which can be a point of reference for the problem that arises from *Rina*.

must give way to determination of whether the case falls within the scope of the regulation in question at all. A negative answer to this question will mean that the question of immunity will be assessed on the basis of the forum's internal procedural law. Only a positive determination will bring the issue into the sphere of European civil procedural law. The *Rina* case does not provide for a resolution of the priority-issue, because the court of a Member State had no doubt that it had jurisdiction, since the seat of the defendants was located in this state. Thus, in this case immunity was examined as a secondary concern. It is relevant to mention that in the later *SHAPE* case, the priority of examination was clearly given to jurisdiction.[66]

IV. Conclusion

The search for a legal basis for jurisdictional immunity and the assessment of its scope is carried out regardless of whether the concern is with the internal forum's procedural law or with European civil procedural law. In the latter, however, there are some differences, as evidenced by *Rina*.

First of all, the court of a Member State must first determine whether the rules of European civil procedural law apply to the case, which in the context of state immunity requires in particular an assessment of the scope of the regulation in question.

Second, the scope of application of the concept of 'civil and commercial matters' and the *acta iure imperii* exception have not been defined using state immunity. In classifying a case as 'civil and commercial', the court should identify the legal relationship existing between the parties to the dispute, the basis of the claim and the conditions under which it was brought, and take into account whether one of the parties exercises public authority. This means that the existence of state immunity does not lead *ipso iure* to the exclusion of a case from the scope of 'civil and commercial matters'.

Third, if it is established that the rules of European civil procedural law apply to the case, state immunity may constitute an obstacle to the exercise of the jurisdiction conferred by it. This will be the case even if the source of immunity from jurisdiction is international custom, since the rules that are the expression of customary international law are part of the EU legal order.

Fourth, the possible declining of jurisdiction must be in compliance with European fundamental procedural rights, including the right of access to a court as enshrined in Art 47 of the Charter of Fundamental Rights. How this is to be assessed remains to be further defined. It can only be postulated that, while making sure that claimants are entitled to an alternative forum, the court should take into account whether the proceedings before it will meet the

[66] n 49 above, para 74.

requirements set out in Art 47 of the CFR.

The above remarks prove that the considerations presented in *Rina* may provide the courts of Member States with a useful guideline in cases involving state immunity. This is not only because of the importance of the case for international law, including its method of identifying international custom in the field of state immunity, but also because of the position of state immunity in European civil procedural law. By providing an example of the interaction between international law and European civil procedural law, *Rina* contributes to the development of both areas.

Malebolge

'From Paris with Love': Transnational Public Policy and the Romantic Approach to International Arbitration

Giovanni Zarra*

Abstract

This article discusses the concept of imperative norms (either public policy or mandatory rules) in the context of international commercial arbitration. It demonstrates that, as of today, arbitrators are perfectly suited to apply domestic imperative norms and that they have to carry out the difficult task of applying – or at least taking into account – all the imperative norms that may affect the enforceability of the award. The arbitrators' task is, in this regard, to carry out a balancing process between the need to respect party autonomy and the duty to issue an award which is worthy of enforcement. In this regard, the author criticizes and demonstrates the lack of conceptual autonomy of the concept of transnational public policy, a non-identified set of rules grounded in the practice of international commerce which is allegedly binding for international arbitrators.

I. Introduction

This article considers a phenomenon which has attracted increasing attention from scholars and has also found some followers among domestic judges: the existence of a transnational form of imperative norms, ie a set of mandatory principles and rules developed in the practice of transnational commerce and detached from any country of origin. These principles and rules allegedly emerge within the context of a spontaneous legal system shaped by trade practice, which – in deference to the wording used to define the rules developed by merchants in transnational commerce in the Middle Age – is today referred to as 'new *lex mercatoria*'.[1] It is not by chance that transnational public policy and mandatory

* Assistant Professor of International Law and Transnational Litigation, Federico II – University of Naples.

[1] In general terms, on the concept of lex mercatoria, see G. Cordero Moss, *International Commercial Arbitration. Party Autonomy and Mandatory Rules* (Oslo: Tano Aschehoug, 1999), 261; J.H. Moitry, 'Arbitrage international et droit de la concurrence: vers un ordre public de la lex mercatoria?' *Revue de l'arbitrage*, 3 (1989); J. Paulsson, 'La lex mercatoria dans l'arbitrage C.C.I.' *Revue de l'arbitrage*, 55 (1990); E. Gaillard, 'Thirty Years of Lex Mercatoria: Towards the Selective Application of Transnational Rules' 10 *ICSID Review – Foreign Investment Law Journal*, 208 (1995); F. Marrella, *La nuova lex mercatoria* (Padova: CEDAM, 2000), passim; O. Lando, 'Choice of lex mercatoria', in J.P. Ancel et al eds, *Vers de nouveaux équilibres entre ordres juridiques. Melanges en l'honneur de Hélène Gaudamet-Tallon* (Paris: Dalloz, 2008), 747. According to the arbitral decision *Petroleum Development (Trucial Coast) Ltd* v *The Sheikh of Abu Dhabi* (1951) lex mercatoria is the ensemble of 'principles rooted in the good sense and common practice of the

rules have been, indeed, named as the '*jus cogens de la lex mercatoria*'.[2]

According to some scholars, mainly originating and/or teaching in French speaking countries,[3] this conception of public policy and mandatory rules has a precise scope of application: international commercial relationships and arbitration. Indeed, the fact that the latter form of dispute resolution is grounded in a manifestation of party autonomy would, at least in theory, authorize arbitrators to take into account, in their decisions, transnational imperative norms alone, considering that they supposedly operate within a delocalised legal context and they are not bound by any domestic legal order. Hence, within the context of international commercial arbitration, transnational public policy and mandatory rules should (at least according to some scholars) completely replace domestic imperative norms.[4] However, according to others, the two bodies of law must

generality of civilized nations'. Generally speaking, examples of sources of lex mercatoria are found in the UNIDROIT principles on international commercial contracts and in the INCOTERMS developed by the International Chamber of Commerce. These are, however, sources of soft law whose application to a contractual relationship is based on a manifestation of party autonomy. See, in this regard, M.J. Bonell, 'Soft law and party autonomy: the case of the UNIDROIT principles' 51 *Loyola Law Review*, 229 (2005). As far as this author is concerned, there are very few cases of domestic courts applying lex mercatoria. Among these cases, it is worth mentioning the decision of Corte di Cassazione 8 February 1982, *Ditta Fratelli Damiano s.n.c.* v *Ditta August Topfer & Co. GmbH*. The concept of lex mercatoria is related to the tendency which is discussed of contrat sans loi, ie a form of agreement only based on party autonomy and completely detached from domestic legal systems. See, inter alia, S.M. Carbone, 'Il 'contratto senza legge' e la Convenzione di Roma del 1980' *Rivista di diritto internazionale privato e processuale*, 279 (1983); N. Boschiero, Il coordinamento delle norme in materia di vendita internazionale (Padova: CEDAM, 1990), 129; and F. Sbordone, *Contratti internazionali e lex mercatoria* (Napoli: Edizioni Scientifiche Italiane, 2008), 11, 18, where the author talks about party autonomy as a possible source of an a-national system of rules applicable to transnational relationships (a possibility that the same author seems to deny at 41-42). The idea of law beyond the state, however, does not convince and is significantly and explicatively rebutted, ex multis, by H.L.A. Hart, *Contributi all'analisi del diritto* (Milano: Giuffrè, 1964) 126-127.

[2] P. Mayer, 'L'application par l'arbitre des conventions internationales de droit prive', in B. Ancel et al eds, *L'internationalisation du droit. Mélanges en l'honneur de Yvon Loussouarn* (Paris: LGDJ, 1994), 285. See also J.H. Moitry, n 1 above. On the grounding of transnational public policy in lex mercatoria see also S.L. Brekoulakis, 'Transnational Public Policy in International Arbitration', in T. Schultz and F. Ortino eds, *The Oxford Handbook of International Arbitration* (Oxford: Oxford University Press, 2020), 122.

[3] See, eg, B. Goldman, 'The Applicable Law: General Principles of Law', in J.D.M. Lew ed, *Contemporary Problems in International Commercial Arbitration* (Dordrecht: Springer Science+Business Media, 1987), 113; O. Lando, 'The law applicable to the merits of the dispute', in J.D.M. Lew ed, ibid, 101 and 104; N. Boschiero, n 1 above, 124. For some criticisms see C. Seraglini, *Lois de police et justice arbitrale internationale* (Paris: Dalloz, 2001), 296, claiming that lex mercatoria is not a complete system of law; F. Sbordone, n 1 above, 102; and G. Zarra, 'Arbitrato commerciale internazionale, principio di autonomia delle parti e legge applicabile' *Il Foro Napoletano*, 419 (2020), contesting the very existence of lex mercatoria; G. Cordero Moss, n 1 above, 270. Also observes the inadequacy of lex mercatoria as a system of law on which arbitral decisions may be based, in particular in light of its vagueness. In general terms see H. Grigera Naón, *Choice-of-law Problems in International Commercial Arbitration* (Tubingen: J.C.B. Mohr - Paul Siebeck, 1992), 26-39.

[4] E. Loquin, 'Les manifestations de l'illicite', in P. Kahn and C. Kessedjian eds, *L'illicite dans le*

be jointly applied but in case of conflicts between domestic public policy and truly international public policy, the latter shall prevail.[5]

International commercial relationships might, therefore, escape from the national conception of imperative norms and would (mainly) see the application of transnational principles and rules. This is, in the opinion of some scholars, the direct consequence of states' inability to adequately regulate transnational commercial relationships.[6] Furthermore, the reference to a hard core of spontaneous commonly accepted principles and rules might represent the only way of avoiding the cultural clashes which are otherwise commonplace in a world – the one of transnational commerce – where parties with very different cultural and legal backgrounds meet for their businesses.[7]

This approach is, however, misleading (even if fascinating) and, as provocatively asserted in the title of this article, might be considered as a form of 'romantic' view of international arbitration, not by chance originating in France, imagining arbitration as a purely transnational (and autonomous) form of justice detached from any domestic system of law. As demonstrated in the present paper, such an approach might have had greater purchase in a context such as that of the Sixties, when it was argued that arbitrators were incapable of taking into account and/or applying domestic public policies. However, if we recognize – as is today commonly accepted – that arbitrators are well-suited to apply domestic public policy and mandatory rules, it is not necessary to make reference to other imperatives applicable only in the context of international arbitration.

In this article, after having clarified the suitability of arbitrators to apply domestic imperative norms and demonstrated how imperativeness operates in the context of international commercial arbitration, it will be argued that arbitrators are only bound by the relevant domestic imperative provisions. As we will see, on the one hand, they will have to assume a case-based approach and take into account *all* the imperative provisions that they consider relevant for the efficacy and enforceability of their decisions, and, on the other hand, their analysis is not to be limited to substantive norms but shall be extended also to procedural principles and rules. In this regard, the article shows that arbitrators carry out a balancing process between the principle of party autonomy – possibly requesting to disregard some domestic imperative norms – and the need to issue an enforceable award (to be considered both in light of the sources

commerce international (Paris: Litec, 1996), 278, affirming that 'l'ordre public véritablement international est l'assurance d'une execution universelle de la sentence'.

[5] C. Seraglini, n 3 above, 296.

[6] J.B. Racine, *L'arbitrage commercial international et l'ordre public* (Paris: LGDJ, 1999), 3.

[7] J.B. Racine, n 6 above, 359. On the cultural clashes in international arbitration see O. Sandrock, 'How Much Freeedom Should an International Arbitrator Enjoy? The Desire for Freedom from Law v The Promotion of International Arbitration' 3 *The American Review of International Arbitration*, 55 (1992).

of the arbitrators' power and of the *ex post* judicial controls to which the award may be submitted). Against this background, the article argues that transnational public policy (originated in *lex mercatoria*) is actually a category deprived of any autonomous content. This result is apparent when one recognizes that all the provisions composing this alleged set of norms may be grounded in other legal systems (either public international law or domestic law).[8] Similarly, should we place transnational public policy within the usages of transnational commerce, the reference to a set of principles and rules, whose boundaries and content are not defined in any source of positive law, may be misleading and lead to confusion. Finally, the article highlights the particularities of the application of imperative norms by domestic courts at the post-award stage – viz challenge proceedings before courts of the country that is the seat of the arbitration as well as proceedings for the recognition and enforcement of the arbitral award – where domestic judges usually have to balance the application of imperative norms, on the one hand, with the diffused policy favouring the recourse to arbitration, on the other.

II. The Suitability of Arbitrators to Take into Account Public Policy and Mandatory Rules in Their Decisions: The Necessity to Respect Party Autonomy Versus the Duty to Issue an Enforceable Award

At first glance, the relationship between international arbitration and domestic imperative norms might seem a complicated one.[9] Arbitration is based on party autonomy and, at least in theory, arbitration proceedings do not depend on any legal system. Hence, they would seem to escape the application of domestic imperative norms, which are by nature the strongest manifestation of the legal identity of a national community.[10] This observation has certainly a

[8] For a contrary approach see J.B. Racine, n 6 above, 355, arguing that the application of truly international public policy is useful because it allows an escape from the unforeseeable effects of the conflict of laws mechanism and, moreover, due to transnational recognition of the values enshrined by transnational public policy, ensures a wider possibility of recognition and enforcement of the award.

[9] On this topic see P. Mayer, 'Mandatory Rules of Law in International Arbitration' 2 *Arbitration International*, 274 (1986); I. Fadlallah, 'L'ordre public dans le sentences arbitrales' *Collected Courses of the Hague Academy of International Law*, 377 (The Hague: Brill, 1994), passim; J.B. Racine, n 6 above, 19; J. Kleinheisterkamp, 'Overriding Mandatory Laws in International Arbitration' 67 *The International and Comparative Law Quarterly*, 903 (2018). For a general analysis see S.L. Brekoulakis, n 2 above, passim.

[10] G. Bermann, 'The Origin and Operation of Mandatory Rules', in G. Bermann and L.A. Mistelis eds, *Mandatory Rules in International Arbitration* (Huntington (NY): Juris Publishing, 2011), 3; H. Smit, 'Mandatory Law in Arbitration', in G. Bermann and L.A. Mistelis eds, ibid, 207; G. Zarra, 'Arbitrato internazionale e ordine pubblico' *Il giusto processo civile*, 539 (2018). A significant precedent towards the overcoming of this approach is the US Supreme Court Decision *Scherk* v *Alberto Culver* 417 US 506 (1974), in which it was affirmed that securities transactions were not exclusive jurisdiction of domestic courts. This decision, which significantly influenced other

basis in fact: when two (or more) parties choose to make recourse to arbitration they do so, *inter alia*, in order to avoid the particularities and shortcomings of domestic legal systems.

However, the above assertion can be rebutted on the basis of two considerations. First, arbitration is not detached from any legal system. To the contrary, all commercial arbitrations have a 'seat', which is the legal system supporting the proceedings throughout their duration and whose courts will have jurisdiction over possible challenges to the award in cases where one of the parties is not satisfied with the regularity of the arbitral decision.[11] Secondly, and relatedly, while national legal orders – differently from the past[12] – tend to look with favor on the use of international arbitration as a way to solve disputes arising from international transactions[13] and thus allow the arbitrability of almost all subject matters,[14] they still want to ensure that in *all* arbitration proceedings

jurisdictions, overcome the idea of inarbitrability of matters of public relevance affirmed in the US Supreme Court decision *Wilko* v *Swan* 346 US 427 (1953).

[11] On the essential role of the seat of arbitration within all arbitration proceedings see F. Mann, 'The UNCITRAL Model Law. Lex facit Arbitrum' 2 *Arbitration international*, 241 (1986). For a significant number of authorities, refer to G. Zarra, 'La lex arbitri e la lex loci arbitri: tra verità normative e incertezze dottrinali' *Diritto del commercio internazionale*, 533 (2019). For the sake of the present work, it is to be highlighted that the seat does not necessarily correspond to the venue where arbitral proceedings take place.

[12] Arbitration was historically considered with hesitation even in the UK. See Lord Thomas of Cwmgiedd, 'Developing Commercial Law Through the Courts: Rebalancing the Relationship Between the Courts and Arbitration, Text of the Bailii Lecture 2016', available at https://tinyurl.com/tr6u5a77, 2 (2016) (last visited 30 June 2021); for a partially different opinion, see S.L. Brekoulakis, 'The Historical Treatment of Arbitration Under English Law and the Development of the Policy Favouring Arbitration' 39 *Oxford Journal of Legal Studies*, 124 (2019). On the increasing role of arbitration as an effective tool of dispute settlement see W.W. Park, 'Private Adjudicators and the Public Interest: The Expanding Scope of International Arbitration' 12 *Brooklin Journal of International Law*, 629 (1986).

[13] J.B. Racine, n 6 above, 4; F. Salerno, 'Il coordinamento tra arbitrato e giustizia civile nel regolamento (UE) 1215/2012 (Bruxelles I-bis)' *Rivista di diritto internazionale*, 1146 (2013); P. Perlingieri, 'Sulle cause della scarsa diffusione dell'arbitrato in Italia' *Il giusto processo civile*, 657 (2014); T. Rossi, *Arbitrabilità e controllo di confomità all'ordine pubblico* (Napoli: Edizioni Scientifiche Italiane, 2018), 27. In this regard, A. Leandro, 'Qualche riflessione sul rinvio nell'arbitrato commerciale internazionale', in G. Contaldi et al eds, *Liber Amicorum Angelo Davì* (Napoli: Editoriale scientifica, 2019), 1881, where it is agreeably affirmed that we can today look as arbitration and domestic jurisdiction as perfectly equivalent means of conflict resolution.

[14] S.L. Brekoulakis, 'On Arbitrability: Persisting Misconceptions and New Areas of Concern', in L.A. Mistelis and S.L. Brekoulakis eds, *Arbitrability: International and Comparative Perspectives* (London – The Hague: Kluwer Law International, 2009) 31, who points out that arbitrability is no longer related to public policy or to the inadequacy of arbitrators as decision makers but, on the contrary, it depends on the inherent characteristics of arbitration as an instrument based on consent: '(i)narbitrability should be examined in light of the inherent limitations of arbitration as a dispute resolution mechanism of contractual origins. Based on consent, arbitration has intrinsic difficulties to affect a circle of persons other than the contractual parties to an arbitration agreement. This conceptual limitation of arbitration has repercussions on the scope of arbitrability' (31-32). For example, Brekoulakis argues, even if nothing precludes this possibility in theory, it is very difficult to arbitrate bankruptcy proceedings because this would require a manifestation of consent by the entire plethora of involved parties and this is very unlikely to occur. For a similar approach favoring

related to their own legal system the imperative norms expressing the fundamental values of the forum are nevertheless respected.[15] This finds further confirmation, on the one hand, in the text of the relevant treaties (such as the 1958 New York Convention on the Recognition and Enforcement of Foreign Arbitral Awards (New York Convention))[16] or domestic legislation concerning challenge and enforcement of arbitral awards (eg the English Arbitration Act and the Italian Code of Civil Procedure)[17] and, on the other hand, in the circumstance that domestic courts have regularly annulled or refused to enforce awards running against mandatory rules or public policy of the forum.[18] In other words, while (most) States have significantly enlarged the scope of arbitrable matters, including also particularly sensitive areas,[19] they still want to have a so-called 'second look' at the award in order to ensure its compliance with the imperative norms of the forum.[20] Such a second look can both be exercised when domestic courts at the seat evaluate a challenge to the award or when a domestic court is called upon to enforce an arbitral award in accordance with the New York Convention, whose Art V(2)(b) allows non-enforcement for contrariness to the public policy of the forum.[21]

arbitrability see P. Perlingieri, 'La sfera di operatività della giustizia arbitrale' *Rassegna di diritto civile*, 583 (2015); G. Zarra, n 10 above, 550-555. For scholarly essays still connecting arbitrability and public policy see K.H. Bockstiegel, 'Public Policy and Arbitrability', in P. Sanders ed, *Comparative Arbitration Practice and Public Policy in Arbitration*, (The Hague: ICCA Congress Series, Volume 3, Kluwer Law International, 1986), 177; M. Hunter, G. Conde e Silva, 'Transnational Public Policy and Its Application in Investment Arbitrations' 4 *Journal of World Investment*, 367 (2003).

[15] As a confirmation of this idea, see, explicatively, J.D.M. Lew et al, Comparative International Commercial Arbitration (London – The Hague: Wolters Kluwer Law and Business, 2003), 83. Indeed, generally speaking, in international commercial arbitration the relevant imperative norms are those of international public policy (opposite to national public policy) and overriding mandatory rules. See A. Atteritano, L'enforcement delle sentenze arbitrali del commercio internazionale (Milano: Giuffrè, 2009), 335.

[16] See Art V, para 2, stating that '(r)ecognition and enforcement of an arbitral award may also be refused if the competent authority in the country where recognition and enforcement is sought finds that: (a) The subject matter of the difference is not capable of settlement by arbitration under the law of that country; or (b) The recognition or enforcement of the award would be contrary to the public policy of that country'.

[17] See, eg, s. 67 of the 1996 English Arbitration Act or Art 829 of the Italian Code of Civil Procedure. A separated analysis is deserved by the French law and case law, for which see the next Section.

[18] See the last Section of this article.

[19] See the significant example of succession law (in Italian law) on which see G. Perlingieri, 'La disposizione testamentaria di arbitrato. Riflessioni in tema di tipicità e atipicità del testamento' *Rassegna di diritto civile*, *L'autonomia negoziale nella giustizia arbitrale* (Napoli: Edizioni Scientifiche Italiane, 2016), 411.

[20] See Mitsubishi Motors Corp. v Soler Chrysler-Pymouth Inc., 473 US 614 (1985); Case C-126/97 Eco Swiss China Time Ltd. v Benetton International NV, Judgment of 1 June 1999. See H. Grigera Naón, n 3 above, 221-270; L.G. Radicati di Brozolo, 'Arbitrage commercial international et lois de police: Considerations sur le conflits de jurisdictions dans le commerce international' *Collected Courses of the Hague Academy of International Law*, 265 (The Hague: Brill, 2005), 471; and, also for other references, G. Zarra, n 10 above, 553.

[21] A.J. van den Berg, *The New York Arbitration Convention of 1958* (Deventer, Antwerp,

In light of the above, it can be argued that arbitrators' task in relation to imperative norms is to undertake a process of balancing. On the one hand, they have to ensure, as far as possible, respect for party autonomy (the cornerstone of international arbitration)[22] and this might require them to depart from the application of certain imperative norms, for example in cases where a contractual provision runs against the public policy of the country of the seat. On the other hand, however, they will have to issue an award that is compliant with the requirements of the relevant legal systems as expressed by its imperative norms. This is mainly due to both the necessity to respect the provisions of the legal system which is the source of arbitral power – the law of the seat – and the legal provisions of all the systems which will have 'a second look' on the final award or where the decision will have to produce its effects – and here again the law of the seat, as well as the law of the likely places of enforcement and all the legal systems which are somehow related to the case.

Having made these clarifications, it is to be noted that, usually, judicial decisions applying the New York Convention, supported by scholars, tend to simply discuss the effects of the contrariness of arbitral awards to public policy, but it is commonly accepted that the reference applies to both public policy and overriding mandatory rules.[23] As a confirmation of this statement, it is possible to mention Recommendation 1(d) of the Final Report on public policy as a bar to enforcement of arbitral awards issued by the International Arbitration

Boston, London, Frankfurt: Kluwer Law and Taxation Publishers, 1981), 359. In this regard, it is to be noted that the public policy exception set forth by the New York Convention may be extended also to mandatory rules (as discussed later in this Section) and this is the reason why we can consider that – for the purpose of this article – all forms of domestic imperatives may be protected through Art V(2)(b) of the Convention.

[22] See, inter alia, J.D.M. Lew, L.A. Mistelis and S. Kroll, n 15 above, 86 and 413, where it is affirmed that it 'is now recognised that party autonomy operates as a right in itself. The rule has a special transnational or universal character and has binding effect because it has been agreed to and adopted by the parties. Unquestionably, party autonomy is the most prominent and widely accepted international conflict of laws rule. These national conflict of laws systems recognise that contracting parties do express their view as to the law to govern their contractual relations, and the national laws have no reason to ignore and very limited rights to interfere with the expressed will of the parties'.

[23] A.N. Zhilsov, 'Mandatory and Public Policy Rules in International Commercial Arbitration' 42 *Netherlands International Law Review*, 81 (1995); C. Seraglini, n 3 above, 157; L.G. Radicati di Brozolo, 'Controllo del lodo internazionale e ordine pubblico' *Rivista dell'arbitrato*, 629 (2006); G. Bermann, n 10 above, 6; A. Sheppard, 'Mandatory Rules in International Commercial Arbitration: An English Law Perspective', in G. Bermann and L.A. Mistelis eds, n 10 above, 173; L. Villiers, 'Breaking in the Unruly Horse: The Status of Mandatory Rules of Law as a Public Policy Basis for the Non-Enforcement of Arbitral Awards' 18 *Australian International Law Journal*, 155 (2011). Contra, see E. Gaillard, 'Droit applicable au fond du litige' *Journal Clunet du Droit International*, 27 (1991), arguing that in arbitration there is space for public policy only, because lois de police preclude an analysis of the content of the applicable law. This problem is, however, overcome if one accepts that, from the point of view of their substantive content, there is generally no difference in functioning between public policy and mandatory rules. On the same vein see A. Atteritano, n 15 above, 327.

Committee of the International Law Association after the 2002 New Delhi Conference, according to which:

> (t)he international public policy of any State includes: (i) fundamental principles, pertaining to justice or morality, that the State wishes to protect even when it is not directly concerned; (ii) rules designed to serve the essential political, social or economic interests of the State, these being known as 'lois de police' or 'public policy rules'; (…)

This is not surprising if we accept that, as confirmed by Art 9 of the Rome I Regulation, overriding mandatory rules express the fundamental principles of a country and, therefore, from the substantive point of view they are usually not different (or at least not far) from public policy, both representing the imperative norms of a legal system.

On the other hand, Recommendation 3(a) asserts that simple mandatory rules are irrelevant for the purpose of international arbitration, considering that they do not express fundamental values of a state which shall be applied to all cases (viz also to arbitration proceedings which are not entirely related to the forum) and without exceptions. It affirms that:

> (a)n award's violation of a mere 'mandatory rule' (ie a rule that is mandatory but does not form part of the State's international public policy so as to compel its application in the case under consideration) should not bar its recognition or enforcement, even when said rule forms part of the law of the forum, the law governing the contract, the law of the place of performance of the contract or the law of the seat of the arbitration.

However, nothing seems to preclude that, where the Rome I Regulation is considered applicable by arbitrators (being part of the law of the seat) or is at least taken into account as a point of reference for determining the applicable law,[24] if an arbitration is entirely located within a domestic legal order and the parties have chosen to apply a foreign system of law, Art 3, para 3, of the Regulation still allows the application of simple mandatory rules of the law of the country where the dispute is located.[25]

Having explained that the application of imperative norms in arbitration is usually required by domestic legal orders in order to give effect to arbitral awards, it is important to point out that this conclusion is also dictated by considerations pertaining to the intrinsic characteristics of international

[24] In this regard, it is to be noted that, in the unlikely case that arbitrators decide to autonomously determine the applicable law without referring to any domestic conflict of laws (something that is possible for whoever considers arbitration as detached from any country of origin), it is nevertheless possible that they make reference to domestic conflict of laws systems in order to justify their decision. See J.D.M. Lew et al, n 15 above, 425.

[25] See G. Zarra, n 10 above, 563.

commercial arbitration as an effective means of dispute settlement.

The first and foremost reason pushing arbitrators to ensure respect for the mandatory rules lies in the (today universally accepted) obligation according to which they should do their best to issue an award that is enforceable.[26] In this regard, Art 42 of the Arbitration Rules of the International Chamber of Commerce (as amended in 2020 and applicable since 1 January 2021) is emblematic. It is named 'General Rule' and affirms that:

> In all matters not expressly provided for in the Rules, the Court and the arbitral tribunal shall act in the spirit of the Rules and shall make every effort to make sure that the award is enforceable at law.

Similar provisions may be found in Art 32.2 of the 2020 Arbitration Rules of the London Court of International Arbitration[27] and in Art 2.2 of the Stockholm Chamber of Commerce Rules.[28] It is evident, in this regard, that an award not respecting the imperative norms of the country where the arbitration is seated, or of the place where the award will be enforced, risks not being recognized (and then enforced).

Hence – apart from the obligations imposed by the law of the seat – the reason why arbitrators tend to respect domestic mandatory rules may be, at least implicitly, founded on a different legal obligation, pertaining to the arbitrators' mandate.[29] This is another factor of the balancing process to be carried out by arbitrators pointing towards a possible compression of party autonomy in favor of the application of the relevant imperative norms.

The question, therefore, arises regarding which country's imperative norms have to be respected by arbitrators.[30] In this regard it is commonly accepted

[26] On this obligation (and its fundamental role in arbitration) see P. Mayer, n 9 above, 284; H. El Talhouny, 'The Respect by the Arbitrator of Rules of Public Policy in International Commercial Disputes'1 International Journal of Arab Arbitration, 27 (2009).

[27] 'For all matters not expressly provided in the Arbitration Agreement, the LCIA, the LCIA Court, the Registrar, the Arbitral Tribunal, any tribunal secretary and each of the parties shall act at all times in good faith, respecting the spirit of the Arbitration Agreement, and shall make every reasonable effort to ensure that any award is legally recognised and enforceable at the arbitral seat'.

[28] 'In all matters not expressly provided for in these Rules, the SCC, the Arbitral Tribunal and the parties shall act in the spirit of these Rules and shall make every reasonable effort to ensure that any award is legally enforceable'.

[29] A.S. Rau, 'The Arbitrator and 'Mandatory Rules of Law'', in G. Bermann and L.A. Mistelis eds, n 10 above, 90; H. Smit, n 10 above, 212. In this regard, J.B. Racine, n 6 above, 20, also talks about an approach by arbitrators based on utilitarism and this is not a plausible argument considering that they may consider worth applying the relevant imperative norms in order to issue an enforceable award and ensuring their reputation within the 'community' of international arbitration. On the application by arbitrators of private international law see A. Leandro, n 13 above, 1882.

[30] Some authors have tried to make a decalogue of circumstances which may induce arbitrators to apply a certain country's mandatory rules. See M. Blessing, 'Mandatory Rules of Law Versus Party Autonomy in International Arbitration' *Journal of International Arbitration*, 14, 23

that the reference primarily applies to public policy and mandatory provisions of the state of the seat of arbitration, whose courts will certainly annul the award in the case that it does not comply with imperative norms.[31] The imperative norms of the seat, therefore, shall never be ignored by arbitrators. However, in order to respect the abovementioned obligation to issue an award which is enforceable, other legal systems may come into play and, among them, particular relevance is to be conferred to the country or countries where – in light of the concrete circumstances of the case – it is likely that enforcement proceedings will be started.[32] Some authors also refer to the law of the country of performance of the obligation in question.[33] In this regard, however, it is not possible to give a universal solution: arbitrators will have to make an analysis of the case and, every time imperative norms of the place of performance may risk causing the non-enforceability of the award, they will certainly have to take these rules into account.[34] For this reason, it is also arguable that – in adopting a pragmatic approach – arbitrators have to take into account, on a case-by-case basis, the imperative norms of all the other relevant domestic legal systems.[35] In this regard, however, due to the essential role of the law of the seat in relation to any arbitration proceedings, it can be asserted that in all cases where the imperative norms of the seat are in conflict with other imperative norms that are considered relevant to the case, the latter shall prevail; in the alternative, the final award

(1997); A. Barraclough and J. Waincymer, 'Mandatory Rules of Law in International Commercial Arbitration' 6 *Melbourne Journal of International Law*, 205 (2005).

[31] This reasoning is reinforced if one accepts that annulled awards may not be enforced abroad or that, generally speaking, the annulment decision taken at the place of the seat is to be significantly taken into account before foreign enforcing courts. On this issue see, also for the relevant bibliography and case law, G. Zarra, 'L'esecuzione dei lodi arbitrali annullati presso lo stato della sede e la Convenzione di New York: verso un'uniformità di vedute?' *Rivista dell'arbitrato*, 574 (2015).

[32] H. Fazilaftar, Overriding Mandatory Rules in International Commercial Arbitration (Cheltenham – Northampton: Edward Elgar, 2019), 5 and 76.; A.S. Rau, n 29 above, 77; L. Shore, 'Applying Mandatory Rules of Law in International Commercial Arbitration', in G. Bermann and L.A. Mistelis eds, n 10 above, 131.

[33] A. Sheppard, n 23 above, 172.

[34] A.K.A. Greenewalt, 'Does International Arbitration Need a Mandatory Rules Method?', in G. Bermann and L.A. Mistelis eds, n 10 above, 148, notes that the parties might contractually rule out the application of imperative norms. It is my opinion, in this regard, that this reasoning may not apply with regard to mandatory rules of the seat of arbitration – the non-respect of which will reasonably lead to the annulment of the award – but might be taken into consideration when the parties agree to exclude the application of imperative norms of other legal systems, which do not have a direct influence on the efficacy of the decision (provided that one accepts that annulled awards may not circulate or that decisions annulling arbitral awards are at least a significant factor to be considered when enforcing foreign awards).

[35] C. Seraglini, n 3 above, 188. See also N. Voser, 'Mandatory Rules of Law as a Limitation on the Law Applicable in International Commercial Arbitration' 7 *American Review of International Arbitration*, 338 (1997); G. Zarra, n 10 above, 563. With express reference to the application of this pragmatic approach to lois de police see L.G. Radicati di Brozolo, 'L'arbitrato come sistema transnazionale di soluzione delle controversie: caratteristiche e rapporto con il diritto interno' *Rivista dell'arbitrato*, 25 (2020).

may risk being annulled.

According to some scholars, moreover, respect for the relevant imperative norms is also dictated by the parties' expectation that such norms will be respected.[36] This consideration is certainly based in fact – we could talk, in this regard, of a form of *social* legitimacy of awards meeting the parties' expectations[37] – and has a significant *psychological* relevance in arbitral decisions[38] even if, in this author's opinion, it does not play a significant role in the balancing process to be carried out by arbitrators as the *legal* obligation to do their best to issue an enforceable award.[39]

Such a duty also has an impact on (and it is to be valorized in light of) the principle of legal certainty, which requires stability of the arbitral *res judicata* and promotes the idea according to which, once a case is decided, the decision should put an end to disputes. In this respect, considering that an award which is possible of being annulled or not enforced risks not ensuring a final dispute resolution, this is another reason supporting the idea that arbitrators may sacrifice, when necessary, the respect of party autonomy in favor of the relevant imperative norms.[40]

In conclusion, it can be argued that arbitrators have a duty to take into account the relevant national imperative norms. In this regard, it has been suggested that respect for this duty is the main reason why states continue to give credit to international arbitration as a valuable (arguably the most valuable) way of solving disputes pertaining to transnational commerce.[41] Should arbitrators carry out their function without respecting the legal provisions of the countries which are related to the proceedings, indeed, this might lead to a diminishing of the role of arbitration within domestic legal systems and, possibly, to a restraint of the attitude of *favor arbitrati*.

1. Procedural and Substantive Imperatives in Arbitration: Some Brief Examples

[36] L.G. Radicati di Brozolo, n 20 above, 464; R.S. Rau, n 29 above, 78.

[37] The argument is, mutatis mutandis, applicable with regard to the reasons why arbitrators apply precedents. See A. Rigo Sureda, 'Precedent in Investment Treaty Arbitration', in C. Binder at al eds, *International Investment Law for the 21st Century: Essays in Honour of Christoph Schreuer* (Oxford: Oxford University Press, 2009), 831-832.

[38] In this regard, note that C. Seraglini, n 3 above, 189, affirmed that arbitrators shall apply imperative norms in order to enhance the credibility of arbitration as the natural forum of transnational commerce.

[39] On the identification of the relevant imperative norms, see B. Audit, 'How do Mandatory Rules of Law Function in International Civil Litigation?', in G. Bermann and L.A. Mistelis eds, n 10 above, 58. There is no space in arbitration, like in domestic litigation, for considerations based on subjective morals of arbitrators. See C. Seraglini, n 3 above, 271.

[40] On the value of legal certainty and finality see G. Zarra, *Parallel Proceedings in Investment Arbitration* (Torino: Giappichelli; The Hague: Eleven International Publishing, 2017), 37.

[41] C. Seraglini, n 3 above, 190, talks about a tacit agreement between states and arbitrators.

Having demonstrated that arbitrators have a duty to apply the imperative norms of the states connected with the proceedings, for the sake of completeness it is here worth briefly describing how arbitrators have applied domestic imperative norms. In this regard, it is worth highlighting that, while

> (a)n arbitration tribunal may have to consider the effects of international public policy at different stages of the proceedings. This includes when deciding whether to give full or limited effect to the law chosen by the parties or which is otherwise applicable, if jurisdiction, ie arbitrability, is contested, or where the factual or substantive issues are alleged to be contrary to fundamental international standards[42]

another very relevant ambit of application of public policy rules concerns arbitrators' respect for fundamental canons of procedure. In this regard, it is necessary to make a distinction between substantive and procedural imperative norms.

With regard to substantive imperative norms, as already outlined in the previous Section, there is case law confirming that, for example, arbitrators have taken into account – apart from the peremptory legislation of the law of the seat,[43] of the law applicable to the substance of the claim,[44] or of the law of the place where the enforcement was likely to take place[45] – imperative norms of other countries which, *in concreto*, were relevant for the case, such as the country where a patent was registered[46] or of the place of performance of the obligation in question.[47] The reasoning used by arbitrators, again, consists of a balancing process in light of the circumstances of the individual case: the application of mandatory rules is based on the existence of an effective link and connection between the case and the imperative norms and, for some scholars, on the interest that a state somehow involved in the case has in seeing its

[42] J.D.M. Lew et al, n 15 above, 423.

[43] Which is applicable as a matter of law. See, eg, Section 4 of the 1996 English Arbitration Act.

[44] See the in-depth analysis carried out by J.B. Racine, n 6 above, 237. As an example of this kind it is possible to refer to all the cases where arbitrators have applied EU competition law (on which see J.B. Racine, n 6 above, 255-264).

[45] See eg ICC Award no 953 mentioned by J.D.M. Lew, Applicable Law in International Commercial Arbitration (New York: Oceana, 1978), 543.

[46] See ICC Award no 1230 mentioned by J.D.M. Lew, ibid, 541-542.

[47] See ICC Award no 761 mentioned by J.D.M. Lew, ibid, 542-543. In England, this is a consequence of the well-known English High Court decision of 17 December 1919, Ralli Bros v Compania Naviera Sota y Aznar (1920) 1 K.B. 614, where the English Court stated that it will not enforce a contract which is unlawful at the place of performance. A different approach was assumed by the arbitrators in the well-known Hilmarton case where the arbitrators refused to take into account Algerian law on corruption in an arbitration governed by Swiss law (where the performance of the contract had to take place in Algeria). However, the Court de Justice du Canton de Geneve, Decision of 7 November 1989, Omnium de Traitement et de Valorisation (OTV) v Hilmarton, annulled the award by saying that the arbitrators should have applied Algerian law. This decision was confirmed by the Swiss Federal Tribunal on 17 April 1990.

imperative norms applied.[48]

In this regard, it is particularly relevant to note that, according to the principle *iura novit curia* (that is today considered applicable in international arbitration too, and is named *iura novit arbiter*),[49] arbitrators are entitled to take into account the relevant imperative norms even if the parties did not raise this issue, as long as the parties are then provided with the possibility of expressing their views on the principles and rules that the arbitrators are going to apply. By contrast, arbitrators have refused to take into account mandatory rules of countries which were not related in any way to the case at hand, notwithstanding the fact that one of the parties claimed the application of such mandatory rules. In ICC Case 6320 of 1991,[50] for example, the dispute concerned a choice of law clause for Brazilian law and the proceedings were seated in France. The claimant was Brazilian and the respondent was from the USA. One of the parties asked for treble damages under the US Racketeer Influenced and Corrupt Organizations (RICO) Act arguing that it was an imperative set of norms of the law of the respondent. The Tribunal, however, noted that the goal of the RICO Act was to prevent corruption within the territory of the USA and therefore application to a case located in Brazil was outside of the scope of the norm. According to the Tribunal, there was not even an abstract interest of the USA in seeing the RICO Act applied to this case.[51]

Turning to procedural imperative norms, it should be recalled that the New York Convention sets forth, as a grounds for the non-enforcement of an award, that:

> (t)he party against whom the award is invoked was not given proper

[48] On this topic see D. Hochstrasser, 'Choice of Law and 'Foreign' Mandatory Rules in International Arbitration' 11 *Journal of International Arbitration*, 57 (1994); A.S. Rau, n 29 above, 82.

[49] G. Kaufmann-Kohler, 'The Arbitrator and the Law: Does He/She Know It? Apply It? How? And a Few More Questions' 21 *Arbitration International*, 631 (2005); A. Carlevaris, 'L'accertamento del diritto nell'arbitrato internazionale tra principio jura novit curia e onere della prova' *Rivista dell'arbitrato*, 505 (2007); A. Sheppard, n 23 above, 204. See also International Law Association, Final Report Ascertaining the Contents of the Applicable Law in International Commercial Arbitration, Rio de Janeiro Conference, 2008. See Swiss Federal Tribunal, decision of 15 April 2015, case 4/A_554/2014 where it was affirmed that 'in Switzerland, the right to be heard concerns particularly factual findings. The parties' right to be invited to express their position on legal issues is recognized only to a limited extent. Generally, according to the principle jura novit curia, state or arbitral tribunals are free to assess the legal relevance of factual findings and they may adjudicate based on different legal grounds from those submitted by the parties'. See also, inter alia, UK Commercial Court, Queen's Bench Division, Pacol ltd. v Joint Sotck Co Rossakhar [2000] 1 Lloyd's Rep 109; Paris Court of Appeal, decision of 15 March 2016, De Sutter P. – K., DS2 S.A. et al v Republic of Madagascar. In the arbitral practice, see Caratube International Oil Company LLP v Republic of Kazakhstan, Decision on the annulment application of Caratube International Oil Company LLP of 21 February 2014, ICSID Case no ARB/08/12, para 90.

[50] The case is mentioned and described by A.S. Rau, n 29 above, 83.

[51] S. Lazareff, 'Mandatory Extraterritorial Application of National Law' 11 *Arbitration International*, 146 (1995).

> notice of the appointment of the arbitrator or of the arbitration proceedings or was otherwise unable to present his case. (Art V(1)(b)).

Hence, the Convention seems to set forth a universal standard for the respect of the *audi alteram partem* principle, according to which the parties must be able to express their views on all the matters which have to be decided by arbitrators. More generally, as noted in scholarship, it is necessary that tribunals, notwithstanding the discretion they enjoy in conducting proceedings, ensure respect for 'fundamental norms of procedural fairness'.[52] This is particularly true since Art V(1)(b) is to be read in conjunction with Art V(2)(b), which provides that awards which are contrary to the public policy of the forum may not be enforced.[53] In this regard, it is to be noted that the concept of public policy shall be determined by the relevant national laws and, as we have seen, this concept usually involves procedural standards. Hence, we can assume that respect for the principle of due process is also dictated by Art V(2)(b) of the New York Convention (and, as we will see, judges shall be able to raise this defence *ex officio*).

From the above, it is understandable that the conventional concept of procedural public policy is to be read in conjunction with the relevant national law provisions (mainly the law of the seat and the law of the likely place of enforcement). In order to understand the general trends concerning the content of this concept in domestic laws on international arbitration, one may first of all refer to the principle of due process of law as recognized in fundamental treaties on human rights (primarily Art 6 of the European Convention on Human Rights and Art 47 of the EU Charter on Fundamental Rights).[54] Such a principle has been also recalled by domestic laws on international arbitration. In this regard, it is worth mentioning – by way of example – S. 33 of the 1996 English Arbitration Act, according to which

[52] S. Schwebel and G. Lahne, 'Public Policy and Arbitral Procedure', in P. Sanders ed, n 14 above, 205.

[53] On the strict relationship between public policy and due process in arbitration see, ex multis, A. Atteritano, n 15 above, 219.

[54] See A. Atteritano, n 15 above, 214; G. Carella, 'Arbitrato commerciale internazionale e Convenzione europea dei diritti dell'uomo', in G. Carella ed, *La Convenzione europea dei diritti dell'uomo e il diritto internazionale privato* (Torino: Giappichelli, 2009), 53; M. Benedettelli, 'Human Rights as a Litigation Tool in International Arbitration: Reflecting on the ECHR Experience' 31 *Arbitration International*, 631 (2005); G. Zarra, 'Rinuncia preventiva all'impugnazione dei lodi arbitrali internazionali e compatibilità con l'art. 6 della Convenzione europea dei diritti dell'uomo' *Rivista dell'arbitrato*, 302 (2016); J. van Compernolle, 'La Convenzione Europea dei Diritti dell'Uomo e l'arbitrato' *Rivista dell'arbitrato*, 663 (2017); A. Leandro, 'Arbitration, Multi-tier Waiver of the Access to Courts and the European Convention on Human Rights: Some Remarks on the Tabbane Decision', in E. Triggiani et al eds, *Dialoghi con Ugo Villani* (Bari: Cacucci editore, 2018), 321; A. Sardu, 'Arbitrato volontario e giusto processo nella giurisprudenza CEDU' *Rivista di diritto internazionale privato e processuale*, 691 (2018); M. Nino, 'Il rapporto tra arbitrato e diritto al giusto processo nella Convenzione europea dei diritti dell'uomo: quali risultati e quali prospettive?' *Ordine internazionale e diritti umani*, 756 (2019).

> The tribunal shall -
>
> (a) act fairly and impartially as between the parties, giving each party a reasonable opportunity of putting his case and dealing with that of his opponent (…).

This provision is expressly recognized as mandatory, considering that S. 4 of the same Arbitration Act states that the provisions listed in Schedule 1 to the Act – which includes S. 33 – are mandatory and have effect notwithstanding any possible agreement of the parties to the contrary.[55] It is not by chance that S. 68 of the Arbitration Act provides for a ground for challenge of arbitral awards issued in England when procedural guarantees set forth in S. 33 are not respected.

Likewise, Art 34(2)(a)(ii) of the UNCITRAL Model Law[56] states that arbitral awards may be challenged whenever

> (t)he party making the application was not given proper motive of the appointment of an arbitrator or of the arbitral proceedings or was otherwise unable to present his case.

Similar provisions making non-respect of procedural guarantees a ground for challenge of arbitral awards may also be found in Italian[57] or French law.[58]

Violations of due process have been found in cases where the respondent was not duly informed of the pending arbitration proceedings,[59] where one of the parties did not have the possibility to put forward their position on a factual[60] or legal[61] element on which the tribunal based its decision, or where the tribunal based its decision on elements which were known by none of the parties.[62] Interestingly, many legal systems (and some authors)[63] do not consider that the duty to explain the motivation for arbitral awards is an imperative

[55] See A. Sheppard, n 23 above, 191.

[56] UNCITRAL Model Law on International Commercial Arbitration (1985), with amendments as adopted in 2006. The Model Law is designed to assist states in reforming and modernizing their laws on arbitral procedure so as to take into account the particular features and needs of international commercial arbitration.

[57] See Art 816-bis of the Code of Civil Procedure.

[58] See Art 1504 of the Code of Civil Procedure, on which see J.B. Racine, n 6 above, 18.

[59] US District Court for the Southern District of Florida, *Corporaciòn Salvadorena de Calzado S.A.* v *Injection Footwear Corp.*, decision of 18 February 1982, 533 F. Supp. 290.

[60] House of Lords, Aiden Shipping Co Ltd v Interbulk Ltd (The Vimeira) (No2), decision of 22 May 1986, [1986] A.C. 965.

[61] English Court of Appeal, Kanoria v Guinness, decision of 21 February 2006, [2006] EWCA Civ 222.

[62] UK Technology and Construction Court, Kye Gbangbola and Lisa Lewis v Smith Sherriff Limited, decision of 20 March 1998, [1999] 1 TCLR 136.

[63] L.G. Radicati di Brozolo, 'Controllo del lodo' n 23 above, 634.

norm to be applied to all arbitration proceedings.[64] In this regard, the present author respectfully dissents from. Motivation is an essential part of arbitral awards, which offers the parties the possibility to understand that all their procedural rights have been respected and to comprehend the reasons for their victory or defeat in the case, as well as the reasonableness of the decision.[65]

III. The Alleged Lack of a Legal Order to Which Arbitral Tribunals Pertain. Theories Asserting that Arbitrators Are Exclusively Bound by Transnational Imperative Norms

The analysis of imperative norms in international arbitration should have ended with the previous Section. However, the approach proposed above – according to which arbitrators are bound to apply the relevant *domestic* imperative norms – does not find unanimous approval in scholarship. There are, indeed, a few – although distinguished – mainly francophone authors who argue that arbitration is a transnational form of justice, grounded only on manifestations of party autonomy and therefore not bound in any way to apply domestic imperative norms.[66] There would be, therefore, no place for any

[64] This is the case of, inter alia, UK and Italy. See, eg, Corte di Appello di Firenze, decision of 22 October 1976, Tradax Export v Carapelli S.p.A. and the other cases mentioned in G. Zarra, 'Arbitrato internazionale e ordine pubblico' n 10 above, 558-559. The issue of motivation is analysed in depth by T. Carbonneau, 'Rendering Arbitral Awards with Reasons: The Elaboration of Common Law of International Transactions' 23 Columbia Journal of Transnational Law, 579 (1985); P. Lalive, 'On the Reasoning of International Arbitral Awards' *Journal of International Dispute Settlement*, 1, 55 (2010); G. Cordero Moss, 'Reasoning in Arbitration: What Do Users Want or Need?', in ICC Institute of World Business ed, *Dossier XVIII: Explaining why you lost: Reasoning in Arbitration* (Paris: ICC Publishing, 2020), 93.

[65] In a recent decision issued in Singapore, it was argued that the summary of the relevant facts, the crystallization of the parties' cases and the analysis of the relevant documents and of the merits of certain arguments were sufficient reasons for an arbitral award. See *TMM Division Maritima SA de CV* v *Pacific Richfield Marine Pte Ltd* (2013) 4 SLR 972. This leads us to imply that lack of motivation is considered as part of public policy in Singapore. It is certainly so in Australia, where the Supreme Court of Victoria held that an arbitrator must five reasons commensurate to those provided by judges in their determination. See *BHP Petroleum Pty Ltd* v *Oil Basins Ltd* [2006] VSC 402. The same Court, however, then clarified that the duty of motivation may change on the basis of the complexity of the case at hand. See Westport Insurance Corporation v Gordian Runoff Ltd [2011] HCA 37. In general, it has been affirmed that the award shall be able to inform the parties on the bases (legal and factual) that brought the tribunal to reach its decision. This is also the approach by Courts of New Zealand. See Ngāti Hurungaterangi & Ors v Ngāti Wahiao [2016] NZHC 1486.

[66] See, inter alia, P. Lalive, 'Transnational (or Truly International) Public Policy and International Arbitration', in P. Sanders ed, n 14 above, 258; E. Gaillard, 'International Arbitration as a Transnational System of Justice', in A.J. van den Berg ed, *Arbitration – The Next Fifty Years, ICCA Congress Series 16* (London – The Hague: Kluwer Law International, 2012), 66. In this regard, J.B. Racine, n 6 above, 6, affirms that the only limitation to arbitration should be the public policy of the place of enforcement: 'aussi libérale que soit l'attitude des Etats envers l'arbitrage, une reserve s'impose: celle de l'ordre public. Dans tous les cas en effet 'le juge garde le glaive de l'ordre public sous sa robe. (…) L'ordre public est, en quelque sorte, le seul îlot de resistance à l'autonomie

balancing in arbitration proceedings, considering that party autonomy would be the only value to preserve. According to this approach, strictly anchored in the idea that arbitration is completely autonomous and detached from state justice (so-called 'a-national' or 'floating' arbitration'),[67] in all matters brought before arbitration tribunals, domestic imperatives should be replaced by so-called 'transnational public policy', ie an alleged set of imperative norms which are grounded in the *lex mercatoria*.[68] These norms are claimed to be universally respected in transnational commerce[69] and to represent not the interests of a state but the shared interests of the transnational community of commercial actors.[70] Similarly, other scholars, even if not asserting that the concepts of transnational public policy and mandatory rules have completely replaced domestic imperatives, still believe that a form of transnational public policy exists and that arbitrators are, on the one hand, free to cherry pick from it when making their decisions and, on the other hand, bound to apply it whenever it clashes with domestic imperative norms.[71]

This is a fascinating approach to international arbitration, which might be defined as 'romantic' because it locates international arbitration in an ideal world where justice is detached from state activity. But what is, then, the real content of the concept of transnational imperative norms? Authors and decisions are far from being uniform in this regard.[72] Many scholars, when discussing this form of public policy, make reference to *jus cogens* norms, ie the fundamental principles and rules of public international law.[73] This seems to be the case of the 1989 Resolution of the International Law Institute on 'Arbitration Between States, State Enterprises, or State Entities, and Foreign Enterprises' issued in

de l'arbitrage international'. This opinion recalls the one by R. David, 'Le droit du commerce international: une nouvelle tâche pour les legislateurs nationaux ou une nouvelle lex mercatoria?', in VVAA eds, News Directions in International Trade Law (Dobbs Ferry (NY): Oceana, 1977), 19.

[67] J. Paulsson, 'Arbitration Unbound: Award Detached from the Law of Its Country of Origin' 30 *The International and Comparative Law Quarterly*, 358 (1981); J.B. Racine, n 6 above, 4; J.D.M. Lew, 'Achieving the Dream: Autonomous Arbitration' 22 *Arbitration International*, 179 (2006); E. Gaillard, n 66 above, passim. Contra, see, emblematically, F. Mann, n 11 above, according to whom any arbitration 'is a national arbitration' because it receives authority from a state. The territorial approach is also followed by A.J. van den Berg, n 21 above, 349.

[68] C. Seraglini, n 3 above, 280.

[69] P. Lalive, n 66 above, passim.

[70] C. Kessedjian, 'Transnational Public Policy', in A.J. van den Berg ed, International Arbitration 2006: Back to Basics? (London – The Hague: Kluwer Law International, ICCA Congress Series, Volume 13, 2007), 857.

[71] E. Gaillard, n 1 above, passim; C. Seraglini, n 3 above, 154.

[72] This is not surprising, considering the murky nature of the concept of lex mercatoria. See S.M. Carbone, Autonomia privata e commercio internazionale (Milano: Giuffrè, 2014), 59.

[73] This is also the opinion recently supported by E. De Brabandere, 'The (Ir)relevance of Transnational Public Policy in Investment Treaty Arbitration – A Reply to Jean-Michel Marcoux' 21 *Journal of World Investment & Trade*, 853 (2020). Contra, see J.M. Marcoux, 'Transnational Public Policy as a Vehicle to Impose Human Rights Obligations in International Investment Arbitration' 21 *Journal of World Investment & Trade*, 22 (2020), affirming that the fundamental rights protected by jus cogens norms may fall within the conception of transnational public policy.

Santiago de Compostela (Rapporteurs: Eduardo Jiménez de Aréchaga and Arthur von Mehren), whose Art 2 affirms that

> (i)n no case shall an arbitrator violate principles of international public policy as to which a broad consensus has emerged in the international community.

In this regard, it is worth noting that, while this Resolution addresses arbitration involving states (so-called investment arbitration), it expressly states in the Preamble that it also contains general principles regarding arbitration more broadly.[74] Hence, while respect for international public policy as defined in Article 2 is certainly imposed in international investment cases, there is no reason – from the perspective of the drafters of the Resolution – not to extend it also to other forms of arbitration. This is also (at least indirectly) confirmed by the consideration that, very often, investment arbitration cases are regulated by a certain domestic law and are seated in a domestic legal system; in these cases, investment arbitration cases are very similar, at least from the perspective of the application of imperative norms, to commercial cases.[75]

In the same vein, Gaillard and Savage argue that arbitrators 'should base their judgment on values widely recognized in the international community'.[76] In this regard, indeed, as subsequently recognized in scholarship, it is arguable that the hard core of principles to which these definitions make reference is *jus cogens*.[77] This is also the approach of Lalive, who expressly equates (at least part of) his conception of transnational public policy and domestic imperative norms grounded on public international law.[78] The same Lalive, acting as sole

[74] The fourth Recital of the Preamble expressly states that: 'while there are many principles that apply to international arbitrations in general, this Resolution also draws attention to other principles which are of special importance to arbitrations between States, state enterprises, or state entities, on the one hand, and foreign enterprises, on the other'. (emphasis added)

[75] National laws have a role even in cases celebrated within the framework of ICSID (see Art 42 of the ICSID Convention). As a confirmation of the above see G. Cordero Moss, 'Court Control on Arbitral Awards: Public Policy, Uniform Application of EU Law and Arbitrability', in A. Calissendorff and P. Schöldström eds, *Stockholm Arbitration Yearbook* (London – The Hague: Wolters Kluwer, 2020), 209, affirming that '(t)he specific nature of investment disputes may require considerations that are not generally made in commercial disputes, in particular, due to the public interests involved in investment arbitration. However, it should be remembered that a considerable part of investment disputes is carried out under the rules applicable to commercial arbitration. From a procedural point of view and from the point of view of court control, investment disputes resolved under the rules for commercial arbitration do not differ from regular commercial disputes'.

[76] E. Gaillard and J. Savage, Fouchard Gaillard Goldman on International Commercial Arbitration (London – The Hague: Kluwer Law International, 1999), 863. In the same vein see Fourteau, 'L'ordre public 'transnational' ou 'réellement international'' 138 *Journal du droit international*, 3 (2011); C. Seraglini, n 3 above, 1, also argues that the 'system of reference' for international arbitrators is the law generated by the international community of states.

[77] E. De Brabandere, n 73 above, 853.

[78] P. Lalive, n 66 above, 283. Similarly see J.B. Racine, n 6 above, 368.

arbitrator in ICC Award no 1664,[79] refused to apply the relevant domestic imperative norms – in that case the law of Pakistan, which governed a bank guarantee which was issued by a Pakistani Bank in favor of a Pakistani corporation and the enforcement of which was demanded of arbitrators – and expressly asserted that the only relevant peremptory laws for arbitrators were the rules of truly international public policy.[80]

A critique of this conception of transnational public policy comes directly from the observation that it is a mere repetition of the principles of international law (originating both in customary law and in international treaties, mainly those concerning human rights such as the ECHR) that cannot be derogated by private parties and are considered to be a form of truly international public policy. It seems, therefore, that there is no reason to refer to a further form of public policy which is allegedly proper of international arbitration but is in fact solidly anchored in public international law as applied in domestic legal systems.

Other scholars tend to derive the principles of transnational public policy from a comparative analysis of a large number of domestic legal systems.[81] This is, for example, the case of a well-known arbitration decided under English law, which ruled that the prohibition of bribery and corruption and the related principle according to which a party cannot ask for the enforcement of a contract obtained through corruption (*ex iniuria jus non oritur*) are norms of truly international public policy because they are recognized by the law of most, if not all, countries[82] (as well as by several international conventions). Should we accept that the emergence of a principle from its repetition in different domestic

[79] The content of which is meticulously reported by J.D.M. Lew, Applicable Law n 45 above, 545.

[80] In that case, the arbitrator was asked to impose the execution of a bank guarantee issued by the respondent (Pakistan Bank) in respect of the contractual obligations of a Pakistani corporation towards an Indian creditor. The performance of the guarantee was then allegedly rendered illegal by imperative Pakistani norms issued after the raise of the hostilities between Pakistan and India. The arbitrator, however, refused to take these norms into account saying that as a neutral or international arbitrator he was not obliged to take cognizance of local and politically inspired public policies. Lalive applied, in this regard, international law as described in Lord Mc Nair's 'Legal Effects of War' and deduced that in general international law there is no prohibition of trading with the enemy in presence of armed conflicts.

[81] On this basis, A. Stone Sweet, 'The New Lex Mercatoria and Transnational Governance' 13 *Journal of European Public Policy*, 641 (2006), argues that 'arbitrators are becoming – if with some hand-wringing and reluctance – default law makers for international traders'.

[82] World Duty Free Company Limited v Republic of Kenya, ICSID Case no ARB/00/7, Award, 4 October 2006, para 142 (where several other arbitral and domestic cases are mentioned). Similarly, in the European Gas Turbines v Westman case, the French Court of Appeal of Paris ruled, on 30 September 1993, that 'a contract having influence-peddling or bribery as its motives or object is, therefore, contrary to French international public policy as well as to the ethics of international business as conceived by the largest part of the members of the international community'. Another relevant case in which an arbitrator declared not to have jurisdiction on a contract based on corruption due to transnational public policy is the well-known ICC Case 1110 of 1963 decided by Judge Lagergren acting as sole arbitrator.

legal systems generates an autonomous source of the law applicable in international commercial arbitration only? This solution does not seem possible. Indeed, the only real example of this kind which is put forth by the supporters of this conception of transnational public policy is the abovementioned principle of *ex iniuria jus non oritur*.[83] This principle, however, is so widely recognized by domestic legal systems that Professors Hersch Lauterpacht and Giorgio Sacerdoti noted that this is one of the unquestionable 'general principles of law common to domestic legal systems' to which Art 38, para 1, letter c), of the Statute of the International Court of Justice make reference.[84] Such principles are a source of general international law. Hence, they may be considered, again, part of a truly international public policy (grounded in public international law) and they are applied by arbitrators as part of the relevant domestic imperative norms (ie, it is worth repeating, mainly the ones of the law of the seat and, then, the other legal systems involved in the arbitration). Accordingly, they do not deserve an autonomous place in the legal analysis of the different forms of imperatives.

For the sake of completeness, it is necessary also to point out that in the practice of international investment arbitration some tribunals have held that a rule of transnational public policy (again, the prohibition of bribery and corruption) may be obtained by referring to principles and rules which find application in a significant number of international treaties.[85] Again, what is here

[83] A. Crivellaro, 'Arbitrato internazionale e corruzione' 26 *Rivista dell'arbitrato*, 701 (2019), speaking – on the basis of a comparative analysis – of a truly international public policy principle concerning the prohibition of corruption. M. Hwang and K. Lim, 'Corruption in Arbitration – Law and Reality' 8 *Asian International Arbitration Journal*, 59-60 (2012). See, in this regard, J.M. Marcoux, n 73 above, 4, affirming that '(a)lthough the inclusion of corruption and other forms of illegality appears as fairly uncontroversial, the contours of transnational public policy are inherently precise'.

[84] H. Lauterpacht, Recognition in International Law (Cambridge: Cambridge University Press, 1947), 420-421; G. Sacerdoti, 'Corruption in Investment Transactions: Policy Initiatives, Legal Principles and Arbitral Practice' 24 *ICSID Review – Foreign Investment Law Journal*, 565 (2009); J.B. Racine, n 6 above, 369, makes express reference to principles *ex* Art 38(1)(c) as a source of transnational public policy. The equivalence between transnational public policy and general principles common to domestic legal systems can also by inferred by J.M. Marcoux, n 73 above, 7.

[85] See eg Niko Resources (Bangladesh) Ltd v Bangladesh Petroleum Exploration & Production Company Limited ('Bapex') and Bangladesh Oil Gas and Mineral Corporation ('Petrobangla'), ICSID Case no ARB/10/18, Decision on the Corruption Claim of 25 February 2019, para 434; Vladislav Kim and Others v Republic of Uzbekistan, ICSID Case no ARB/13/6, Decision on Jurisdiction of 8 March 2017, paras 593-597. This is, according to Michael Reisman, the only possible meaning of transnational public policy. See W.M. Reisman, 'Law, International Public Policy (So-called) and Arbitral Choice in International Commercial Arbitration', in A.J. van den Berg ed, n 70 above, 856. Similarly see R. Kreindler, 'Standards of Procedural International Public Policy', in D. Bray and H.L. Bray eds, *International Arbitration and Public Policy* (Huntington (NY): Juris, 2015), 245-249. See, with regard to the applicability of international economic sanctions by arbitrators (considered as a rule of truly international public policy by this author) C. de Stefano, 'L'arbitrabilità dell'embargo internazionale alla prova delle Sezioni Unite' 53 *Rivista di diritto internazionale privato e processuale*, 1998 (2017).

relevant is that the principle or rule at stake is grounded in a source of international law which is applicable in the case. The repetition of a principle in several international treaties may be only used as evidence of the existence (or emergence) of a norm of customary international law.

Some authors, finally, make reference to non-positive standards as sources of transnational public policy. The reference applies both to moral conceptions of imperativeness and to imperative norms grounded in soft- law instruments.

As to the former, according to Julian Lew

> (t)hese truly international or 'pluri-national' criteria are drawn from the fundamental rules of natural law, the principles of 'universal justice', jus cogens in public international law and the general principles of morality and public policy accepted by civilized countries.[86]

This statement is problematic because – apart from referring to forms of public policy generated in public international law – it makes reference to concepts such as 'natural law', 'universal justice' and 'general principles of morality',[87] without providing readers with any guidance as to how to understand these concepts.[88] Moreover, apart from the reference to *jus cogens* principles, due to the very murky borders of the concepts referenced, the idea of transnational public policy would be deprived of any precise meaning.[89] Indeed, Pierre Mayer correctly observed that the reference to transnational principles of justice

> merely serve(s) as justification and do(es) not constitute the source of legal rules which the arbitrator would simply apply.[90]

In addition, readers should be warned of the risk of subjectivism intrinsic to the application of concepts such as 'general principles of morality' and 'universal justice' in domestic litigation and the same considerations apply in international arbitration.[91] Indeed, arbitrators are perfectly conscious of the risk that an

[86] J.D.M. Lew, Applicable Law n 45 above, 534.

[87] In this regard see also J.C. Pommier, Principes d'autonomie et loi du contrat en droit international privé conventionnel (Paris: Economica, 1992), 247, talking about a 'moralité internationale contractuelle'.

[88] Similarly, E. Loquin, n 4 above, 277, affirms that '(l)'arbitre puise à toutes les sources du droit les règles de comportement aptes à satisfaire un sentiment de justice, qui, dans les relations internationals doit tendre à l'universalité'.

[89] S.L. Brekoulakis, 'Transnational Public Policy' n 2 above, 127.

[90] P. Mayer, 'La règle morale dans l'arbitrage international', in P. Bellet et al eds, *Etudes offertes à Pierre Belief* (Paris: Litec, 1991), 384. A.S. Rau, n 29 above, 81, ironically noted that '(r)eliance on 'transnational principles' of 'morality' (like 'transnational principles' of what constitutes effective 'consent' to contract) is attractive precisely because – given that they constitute the core of any developed legal system – they appear to avoid any need for recourse to choice of law. Such are the benefits of a fruitful methodological sloppiness'.

[91] S.L. Brekoulakis, n 2 above, 130, affirms that '(i)t is worrying that some arbitration tribunals feel empowered under a misplaced concept of transnational public policy to render decisions on the

award based on their conception of morality and justice (and not on a system of positive law) may be annulled or not enforced and have therefore regularly avoided this practice.[92] In this regard, it is worth adding that, whenever the parties intend to ask arbitrators to depart from the application of the law and to decide on the basis of their perception of morality and justice, they can provide that the decision shall be taken *ex aequo et bono*.[93] It is not by chance, indeed, that in ICC Case 3540 of 3 October 1980 arbitrators – which were entrusted to decide *ex aequo et bono* – decided to apply *lex mercatoria*, justifying their choice by asserting that, by applying

> the most recent and authoritative doctrine as well as the jurisprudence (...) in determining the substantive law, arbitrators may avoid the rules of conflict of the form, the more so if they have the power of *amiables compositeurs*.

This award perfectly confirms that the concepts of *lex mercatoria* and transnational public policy do not have a foundation in positive law[94] and, on the contrary, they have an equitable, rather than legal, nature.[95] They seem to be a creation of scholarly thinking which can hardly be provided with any autonomous content.[96]

This opinion applies, *mutatis mutandis*, also to alleged public policy principles derived from private sources of contractual regulation, such as UNIDROIT principles, INCOTERMS or codes of conduct, which are mere soft law sources that are applicable to private agreements insofar as the parties

basis of what they consider to be basic standards of morality, no matter how lofty such standards may be'.

[92] C. Seraglini, n 3 above, 273.

[93] G. Zarra, 'Arbitrato commerciale internazionale' n 3 above, 423.

[94] L. Crema, 'Il caso WDF: corruzione e ordine pubblico transnazionale innanzi alla giurisdizione ICSID' 44 *Rivista di diritto internazionale privato e processuale*, 120 (2008).

[95] S.L. Brekoulakis, n 2 above, 125 and 129. As a confirmation of this idea, the author mentions ICC Award no 15300 of 2011, where the sole arbitrators realized that a set of five contracts was a scheme for reverse payments ('kick-backs') to be equated to a form of bribery. He therefore argued that this practice was against 'standards of basic morality' and stated that 'it would not be compatible with fundamental values of international commerce, necessary to allow business being conducted in a loyal surrounding, to lend a helping hand to such agreements'.

[96] As noted by H. Arfazadeh, Ordre public et arbitrage international à l'épreuve de la mondialisation (Bruxelles – Paris – Zurich : Bruylant – LGDJ – Schulthess, 2006), 236, 'une juridiction étatique ne peut, sans mandate et en toute modestie, ni élaborer, ni proclamer des principes transnationaux ou (quasi) universels, de surcroit d'ordre public. Une juridiction étatique, au contraire, ne peux rendre justice qu'au nom de son propre droit, ou en application des principes du droit international privé et du droit international public qui font déjà partie de son droit national'. Similarly, see also J. Fry, 'Désordre Public International under the New York Convention: Wither Truly International Public Policy' 8 *Chinese Journal of International Law*, 89 (2009); H. Fazilaftar, n 32 above, 19. See also, as to the impossibility of outlining a precise content of the lex mercatoria, F. Sbordone, n 1 above, 102.

recall them in their contract.[97] The reference also applies to fundamental principles regulating trade usages,[98] which some domestic decisions[99] considered as binding within the context of the 1980 Vienna Convention on the International Sale of Goods (CISG), whose Art 9, para 2, affirms that

> (t)he parties are considered, unless otherwise agreed, to have impliedly made applicable to their contract or its formation a usage of which the parties knew or ought to have known and which in international trade is widely known to, and regularly observed by, parties to contracts of the type involved in the particular trade concerned.

However, as (correctly noted in scholarship) trade usages are mentioned in Art 9 CISG not as a source of law, but as a mere means of interpretation of contractual provisions to be used in cases of doubt.[100] This means that these usages can neither replace nor integrate the contractual regulation.

Concluding on this point, it is worth borrowing Stavros Brekoulakis' words, arguing that

> (u)nder a conception of transnational public policy that includes non-legal standards, judicial function is dangerously conflated with legislative function and international arbitrators assume the role of the 'regulators of society', which runs counter to the way our world is politically organized today.[101]

Given the above, we might conclude that the idea of arbitration as a system of justice detached from any country of origin and of arbitrators as the judges of 'the international community of businessmen'[102] applying a law – the *lex*

[97] Similarly see J.B. Racine, n 6 above, 373. In this regard see J.M. Marcoux, n 73 above, 16-17; J.M. Jacquet, 'L'ordre public transnational', in E. Loquin and S. Manciaux eds, *L'ordre public et l'arbitrage* (Paris: Lexis Nexis, 2014) 101. In this regard see S.M. Carbone, Autonomia privata n 72 above, 11, correctly highlighting that the same UNIDROIT Principle recognizes in any case the prevalence of domestic imperative norms over contractual regulations agreed by the parties.

[98] These principles are considered as a source of transnational public policy by J.B. Racine, n 6 above, 374, even if the author does not identify their specific content. In the same vein see P. Kahn, Les reactions des milieux économiques, in P. Kahn and C. Kessedjian eds, n 4 above, 491. Contra see T. Mustill, 'The New Lex Mercatoria: The First Twenty-five Years' 4 *Arbitration International*, 111-112 (1988).

[99] See Oberster Gerichstshof Austria, 21 March 2000, in cisgw3.law.pace.edu. Similarly see Oberster Gerichtshof Austra, 15 ottobre 1998 in cisgw3.law.pace.edu.

[100] R. Schmidt-Kessel, 'Article 9', in I. Schwenzer ed, *Commentary to the UN Convention on the International Sale of Goods* (Oxford: Oxford University Press, 2010), 182; R. Goode, 'Usage and its reception in transnational commercial law' *The International and Comparative Law Quarterly*, 46, 35 (1997). Contra, see M.J. Bonell, 'Art. 9', in C.M. Bianca ed, *Convenzione di Vienna sui contratti di vendita internazionale di beni mobili* (Padova: CEDAM, 1992), 39, arguing that these usages may also integrate the contractual regulation in cases of lacunae.

[101] S.L. Brekoulakis, n 2 above, 121.

[102] P. Lalive, n 66 above, 270-271.

mercatoria – originating in the practice of transnational commerce and expressing certain principles of transnational public policy iscertainly very attractive.[103] However, such an idea only complicates an already complicated subject. Arbitration exists because states allow arbitration to exist. The parties' freedom to derogate from state justice exists because the same states recognize the validity of the willingness to refer to arbitration and confer binding force to the award. States want to ensure that the hard core of imperative norms representing the identity of their legal systems are respected in all arbitration cases related to their legal orders;[104] and the parties cannot escape from that.[105] Borrowing the words of Homayoon Arfazadeh, it is possible to assert that

> *(c)ette image idyllique d'un espace transnational homogène et autorégulé est une représentation simpliste et une réalité autrement plus contrastée des relations transnationales. En effet, les dispositions d'ordre public et les lois de police étatiques font partie intégrante de l'environnement juridique des transactions internationales.*[106]

Furthermore, Arfazadeh acutely observes that the relevance of domestic imperatives is even raised by the increase in arbitrable subjects. Indeed, the willingness of states to see their imperative norms applied in arbitration is the counterweight of their agreement to the wide expansion of the number of subjects that can be submitted to arbitration.

In conclusion, provided that the main duty of arbitrators is to provide the parties with an award which is enforceable, arbitrators are bound by the application of the relevant domestic imperative norms.[107] There is no different form of public policy and/or mandatory rules applicable in international

[103] According to this view, arbitrators 'are the guardians of the international commercial order: they must protect the rights of participants in international trade; give effect to the parties' respective obligations under the contract; imply the presence of commercial bona fides in every transaction; respect the customs followed in international trade practice and the rules developed in relevant international treaties; uphold the commonly accepted views of the international community and the policies expressed and adopted by appropriate international organizations; and enforce the fundamental moral and ethical values which underlie every level of commercial activity'. See J.D.M. Lew, Applicable Law n 45 above, 540; J. Kleinheisterkamp, n 9 above, 912, defines (correctly, in our view) this approach as 'utopian'.

[104] M.R. Beniassadi, 'Do Mandatory Rules of Public Law Limit Choice of Law in International Commercial Arbitration?' 10 *International Tax & Business Lawyer*, 11-12 (1992).

[105] Various concerns against transnational public policy have been expressed also by M.W. Reisman, n 85 above, 854; A. Redfern, 'Comments on Commercial Arbitration and Transnational Public Policy', in A.J. van den Berg ed, n 70 above, 871; and M. Pryles, 'Reflections on Transnational Public Policy' 24 *Journal of International Arbitration*, 1 (2007).

[106] H. Arfazadeh, n 96 above, 224.

[107] J.D.M. Lew, Applicable Law n 45 above, 537, observes that: 'The award is the 'raison d'être' of every arbitration; if the award is unenforceable, the whole arbitration proceeding will have been a waste of time and energy. If an arbitrator's award is not enforceable because it violates the public policy of the place of performance, the arbitrator will have failed the responsibility vested in him'.

arbitration. Whenever arbitrators expressly base their decision on transnational public policy, therefore, they have to accept the risk that the award may be successfully challenged or not enforced. This consideration applies in particular where arbitral tribunals do not adequately ground their application and the content of transnational public policy in a source of positive law. This is even recognized by one of the main supporters of the concept of transnational public policy, Prof Pierre Lalive, when he affirms that

> (i)t is understandable that, in general, the international arbitrator would show great caution when using his 'creative powers' and would refrain from resorting unduly to concepts or standards as relative and difficult to define with certainty as that of (transnational) public policy. It is of course much easier for the international arbitrator to refer in a given case to the international public policy *of a State*, since judicial precedents make it comparatively easy to establish its existence and its limits.[108] (emphasis in original).

IV. Public Policy at the Post-Award Stage

Before concluding the present analysis, it is important to note that some authors[109] (followed by very few judicial decisions) argue that domestic legal orders should recognize the transnational nature of arbitration[110] and refuse to regulate the compliance of arbitral awards with domestic imperative norms, having limiting their analysis to the transnational imperative norms emerging from commercial practice. An example of this approach is provided by the well-known Swiss decision in the *Westland* case[111] where, the Federal Tribunal held that it is worth adhering to a

> *notion universelle de l'ordre public, en vertu de laquelle est incompatible avec l'ordre public la sentence qui est contraire aux principes juridiques ou moraux fondamentaux reconnus dans tous les Etats civilisés.*

[108] P. Lalive, n 66 above, 286.

[109] P. Lerebours Pigeonniere, 'A propos du contrat international' 78 *Journal du droit international*, 14 (1951), affirmed that in transnational cases the French Cour de Cassation has adopted an approach 'which does not underlie the particularism of French domestic life and, quite to the contrary, is based on the desire that private transfrontier relations be governed by an international legal order (...) the exception of public policy leads here to the creation within French domestic law of a kind of ius gentium parallel to the domestic common law'.

[110] See, eg, French Cour de Cassation, PT Putrabali Adyamulia c. Rena Holding, 29 June 2007, affirming that 'la sentence internationale, qui n'est rattachée à aucun ordre juridique étatique, est une décision de justice internationale' which is not even required to respect the law of the seat (being it relevant that it complies with the law of enforcement only).

[111] Swiss Federal Supreme Court, F. et U. v W. Inc., 30 December 1994, case 4 P.115/1994.

However, this idea does not find other confirmation in case law and it has been rejected even by authors, such as Gaillard, who generally support the concept of transnational public policy. Indeed, in Gaillard's opinion, recourse to transnational public policy is perfectly legitimate when applied by arbitrators, who do not belong to any particular legal system, but it does not find application in the practice of domestic courts,[112] which are bound to ensure respect for the fundamental principles of the forum.[113] With regard to his affirmation of the irrelevance of transnational public policy before domestic judges, Gaillard's opinion seems perfectly justified. In the EU context, indeed, it is confirmed by Article 3 of the Rome I Regulation, which allows judges of the Member States to apply only national laws (including international law as implemented in domestic legal systems) and precludes recourse to non-State law (ie forms of private codification which are not grounded in any positive law).[114]

However, and for the sake of completeness, it is important to highlight a significant and unique conceptual feature concerning the application of imperative norms in cases regarding challenge or enforcement of arbitral awards.[115] In particular, in these cases the application of imperative norms is to be balanced with the so-called *favor arbitrati* (ie the above-mentioned strong policy favoring arbitration which exists in several domestic legal systems). This has led to a self-restraint of domestic judges who have applied imperative norms as a limitation to the efficacy of arbitral awards quite sparingly.[116] Indeed, a wide recourse to the concept of public policy would undermine the credibility of a certain legal system as an arbitration friendly seat and might have the consequence of leading parties to choose other arbitral seats. Hence, it is understandable that the limited recourse to public policy at the post-award stage is also due to significant economic considerations. A restrictive approach

[112] See E. Gaillard and J. Savage, n 76 above, 955. Similarly see C. Seraglini, n 3 above, 154 and 287.

[113] F. Sbordone, n 1 above, 55-56. In this regard, it has been noted that domestic case law shows that judges difficultly take into account foreign imperative norms when evaluating the compliance of an arbitral award with their legal system. See C. Seraglini, n 3 above, 169.

[114] F. Sbordone, n 1 above, 47; V. Behr, 'Rome I Regulation. A – Mostly – Unified Private International Law of Contractual Relationships within – Most – of the European Union' *Journal of Law and Commerce*, 29, 241 (2011); Z.S. Tang, 'Non-state law in party autonomy – a European perspective' *International Journal of Private Law*, 5, 22 (2012). In this regard, in the law of international arbitration a distinction is usually drawn between 'law' (ie domestic law) and 'rules of law' (ie any substantive rule which is applicable upon express reference by the parties in a contract). See P. Bernardini, L'arbitrato commerciale internazionale (Milano: Giuffrè, 2000), 198.

[115] On this topic see J. Beatson, 'International arbitration, public policy considerations, and conflicts of law: the perspectives of reviewing and enforcing courts' 33 Arbitration International, 175 (2017).

[116] Such a self-restraint takes place, however, also in domestic litigation. In this context, however, the application of imperative norms is limited by the necessity to safeguard the international harmony of decisions and/or by international comity (this latter concept mainly in common law systems).

to public policy at the enforcement stage can indeed be found, inter alia,[117] in France,[118] Korea,[119] Italy,[120] Canada,[121] Hong Kong[122] and Singapore.[123]

Arbitral awards have been successfully challenged or not enforced only in very rare circumstances, which can be summed up in two categories: (i) when they did not respect procedural due process; and (ii) when they have been based on substantive laws which were in striking contrast with the *lex fori*.

As to violations of public policy falling under the former category, we can here recall the analysis under Section 2 above. It is here only necessary to clarify that – at least in the countries where the European Convention of Human Rights is applicable – domestic courts are entitled to ensure respect by arbitrators for the guarantees enshrined in article 6 of the ECHR (as interpreted by the Strasbourg Court).[124] Should arbitrators not have respected the due process principle, national courts are entitled to annul or not enforce the award.

As to domestic court decisions not enforcing arbitral awards for substantive violations of public policy, after restating that these cases are extremely rare due to the self-restraint that domestic judges usually practice in the application of the public policy exception,[125] it is necessary to analyze *the* two significant cases,

[117] See the in-depth comparative analysis carried out by A.G. Maurer, The Public Policy Exception Under the New York Convention (Huntington (NY): Juris Publishing, 2012). For an analysis of Chinese practice see P. Rossi, 'Public Policy and Enforcement of Foreign Awards: An Appraisal of China's Judicial Practice' 31 Diritto del commercio internazionale, 299 (2017).

[118] See, eg, Cour de Cassation 2 December 2015 no 1367.

[119] Korean Supreme Court, Adviso N.V. v Korean Overseas Construction Corporation, 14 February 1995.

[120] Corte d'Appello di Firenze, Nuovo Pignone v Schlumberger, 21 February 2006; Corte d'Appello di Milano, Tensacciai v Terra Armata, 15 July 2006.

[121] Court of Queen's Bench of Alberta, 26 September 2007.

[122] Hong Kong Court of Final Appeal, 9 February 1999, Hebei Import & Export Corporation v Polytek Engineering Company Limited; and 9 October 2007, Karaha Bodas Company LLC v Perushahaan Pertambangan Minyak Dan Gas Bumi Negara (Pertamina), CACV 121/2003.

[123] Singapore Court of Appeal, 1 December 2006, Asuransi Jasa Indonesia (Persero) v Dexia Bank S.A. [2007] 1 S.L.R. 597.

[124] This is not the place for extensive analysis of such cases. See, for some recent decisions, Eur. Court H.R., Tabbane v Switzerland, March 2016, and Mutu & Pechstein v Switzerland, 2 October 2018. In this regard, also for an extensive analysis of scholarship and case law, it is possible to refer to M. Benedettelli, n 54 above, passim; G. Zarra, 'Rinuncia preventiva' n 54 above, passim; J. van Compernolle, n 54 above; A. Leandro, n 54 above, passim; A. Sardu, n 54 above, passim; M. Nino, n 54 above, passim.

[125] See US Court of Appeal for the Second Circuit, Parsons & Whittemore Overseas Co. v Societe Generale de l'Industrie du Papier (RAKTA), 23 December 1974, 508 F.2d 969, where it was affirmed that public policy may be applied only in case of violation of the 'most basic notions of morality and justice'. Similarly see US District Court for the Southern District of New York, MGM Production Group Inc. v Aeroflot Russian Airlines, 14 May 2003, 573 F. Supp. 2d 772 (S.D.N.Y. 2003), in which, interestingly, the Court noted that the payment of commissions allegedly running against US sanctions against Iran was not sufficient to activate the public policy defense. Contra, see the recent English Court of Appeal decision of 12 February 2020, MODSAF v IMS [2020] EWCA Civ 145, where – without discussing of public policy – English judges precluded the enforcement of interests in favor of the Iranian Ministry of Defense because this would have violated EU sanctions

ie the *only* cases where arbitral awards have been (i) successfully challenged in Switzerland due to the award's contrariness to public policy; and (ii) not enforced in England due to the award's contrariness to public policy.

The successful challenge in Switzerland for public policy reasons,[126] occurred in *Matuzalem* v *FIFA*.[127] An award of the Court of Arbitration for Sport (CAS) condemned the Brazilian football player Matuzalem to pay approximately 7 million euros to the Football Club Shakhtar Donetsk due to his arbitrary desertion of the Club's activities. The award also stated that Matuzalem was precluded from playing professional football until he paid the entire amount to Shakhtar Donetsk. The Swiss Federal Tribunal, however, considered that the award constituted an implicit and *sine die* prohibition of work for the football player, something that was in striking contrast with his fundamental rights under the Swiss Constitution.

The *Soleimany v Soleimany* case[128] is the only case where the enforcement of an international award was refused in England on public policy grounds. Sion Soleimany and his son Abner organized the smuggling of carpets from Iran to England and corrupted some diplomats who illicitly transferred the carpets in their luggage. When the business ended, the father refused to pay to the son his part of the earnings. Due to their Jewish faith, the two agreed that the dispute between them had to be resolved by the Beth Din (a Jewish authority), who decided that Abner was entitled to receive an amount of money from Sion arguing that the illegal nature of the contract was irrelevant for Jewish law. When Abner sought enforcement of the award in London, however, the Court of Appeal stated that, notwithstanding the strong *favor* for arbitration existing in England

> (this) is the very type of judgment which the English courts would not recognize on the ground of public policy. We stress that we are dealing with a judgment which finds as a fact that it was the common intention to commit an illegal act, but enforces the contract.[129]

It is important to note that the Court insisted on the circumstance that the existence of an illicit act was recognized by the Beth Din and not by the English

against Iran. See A. Atteritano, n 15 above, 336, explaining that US scholarship talks about a 'pro-enforcement bias'.

[126] Generally speaking, Swiss courts tend to exclude recourse to public policy arguing that it is possible only where the award is in 'violation of the fundamental principles of the Swiss legal system, which hurt the innate feeling of justice'. See Geneva Court of Justice, Import and Export Co. v G. S.A., decision of 11 December 1997.

[127] Swiss Federal Tribunal, Francelino da Silva Matuzalem v Federation Internationale de Football Association, 27 March 2012, case 4/A_558/2011.

[128] English Court of Appeal, 4 March 1998, Abner Soleimany v Sion Soleimany [1998] EWCA Civ 285.

[129] Paras 32-33.

judges, who considered that they were precluded from re-examining the merits of the dispute. This element distinguishes this case from another corruption case, *Westacre* v *Jugoimport*,[130] where the Court refused to apply the public policy exception because – notwithstanding the strong evidence filed in the English proceedings in favor of corruption – the arbitral tribunal considered that corruption was not proved. In this regard, it is interesting to note that – in the first instance decision confirmed by the Court of Appeal's judgment – Colman J noted that:

> (o)n the one hand there is the public policy of sustaining the finality of awards in international arbitration and on the other hand the public policy of discouraging corrupt trading (...) In my judgment, it is relevant to this balancing exercise to take into account the fact that there is mounting international concern about the prevalence of corrupt trading practices (...) However, although commercial corruption is deserving of strong judicial and governmental disapproval, few would consider that it stood in the scale of opprobrium quite at the level of drug-trafficking. *On balance, I have come to the conclusion that the public policy of sustaining international arbitration awards on the facts of this case outweighs the public policy in discouraging international commercial corruption.*[131] (emphasis added)

In conclusion, it seems worth pointing out that recourse to public policy at the post-award stage is a very rare phenomenon. It takes place only in cases where the contrast between the arbitral award and the fundamental principles of the forum is so striking as to necessarily mandate the annulment or non-enforcement of the award, regardless of the *favor arbitrati*. In this kind of reasoning, as in the rest of the arbitral proceedings, there is no space for transnational imperatives.

[130] English Court of Appeal, Westacre Investments Inc v Jugoimport-SDRP Holding Co. Ltd, 12 May 1999, [1999] EWCA Civ 401.

[131] [1998] 3 WLR 770, 798-800.

Corporate and Financial Markets Law

(In)efficient Cost Allocation in Italian Proxy Contests

Giovanni Niccolò Antichi*

Abstract

The research aims to examine the regulatory model adopted in Italy relating to Proxy solicitation. It will be verified whether Proxy solicitation, as actually regulated, adapts to the high level of ownership concentration that characterizes the Italian stock exchange and therefore provides adequate solution for agency's problems to which these ownership structures give rise. Thus, it will be demonstrated that, in its current formulation, Proxy solicitation in Italy appears poorly suited to guaranteeing an appreciable standard of shareholder democracy, especially for reasons related to financial aspects.

I. Overview

When flipping through manuals, monographs, essays, and articles relating to proxy solicitation from all over Europe and the United States,[1] I came across few doctrinal texts that clearly explained the different ways in which proxy solicitation can be carried out and its related corporate governance risks.[2]

* PhD in Business Law, LUISS Guido Carli.

[1] Proxy solicitation is a process of obtaining shareholders' proxies for votes in favour of, or against proposals, and is often used by those who are dissatisfied with a public company's performance, want to change. There are various types of proxy solicitation depending on who the proxy solicitor is, how it is carried out, and whose finances are used to fund it. These are issues addressed later in this paper.

[2] Only a few authors concern themselves with proxy solicitation's expenses (most of them are from the US): see E.R. Aranow and H.H. Einhorn, *Proxy Contests for Corporate Control: A Treatise on the Legal and Practical Problems of Management and Insurgents in a Corporate Proxy Contest* (New York: Columbia University Press, 2nd ed, 1963); L.A. Bebchuk and M. Kahan, 'A Framework for Analyzing Legal Policy Towards Proxy Contests' 78 *California Law Review*, 1071, 1073 (1990); D.M. Friedman, 'Expenses of Corporate Proxy Contests' 51 *Columbia Law Review*, 951, 951-964 (1951); S.W. Mintz, 'Use of Corporate Funds for Proxies and Other Expenses in Fight Over Corporate Management' 8 *New York University Intramural Law Review*, 90, 92 (1953); L.S. Machtinger, 'Proxy Fight Expenditures of Insurgent Shareholders' 19 *Case Western Reserve Law Review*, 212, 212-229 (1968); F.C. Latcham and F.D. Emerson, 'Proxy Contest Expenses and Shareholder Democracy' 4 *Western Reserve Law Review*, 5, 5-18 (1952); L.D. Stifel, 'Shareholder Proxy Fight Expenses' 8 *Cleveland-Marshall Law Review*, 339, 339-350 (1959); Note, 'Corporations. Powers of Stockholders. Stockholders Can Reimburse Victorious Insurgents from Corporate Treasury for Proxy Solicitation Expenses' 69 *Harvard Law Review*, 1132, 1132-1135 (1956); Note, 'Corporations – Payment of Proxy Solicitation Expenses – an Aspect of Corporate Democracy' 31 *New York University Law Review*, 825, 825-831(1956); F.D. Emerson and F.C. Latcham, 'Proxy Contests: A Study in Shareholder Sovereignty' 41 *California Law Review*, 393, 393-438 (1953); B.G. Buchanan et al, 'Shareholder Proposal Rules and Practice: Evidence from a

I consider it appropriate to take the opportunity, at the beginning of the eleventh year after sweeping reforms to Italian proxy solicitation rules were enacted, to comment on how these rules are enforced in respect of companies listed on the Italian stock exchange (Borsa Italiana) and in light of the above doctrinal texts.

Firstly, I will deal with the practice of proxy solicitation in Italy. I will provide a brief introduction outlining the concept to it according to Arts 136–144 TUF (decreto legislativo 24 February 1998 no 58, known as the 'Testo Unico della Finanza') and 135–139 Reg Emitt of the *Commissione Nazionale per le Società e la Borsa* (CONSOB)[3] – ie the public authority responsible for regulating Italian financial markets including the Italian stock exchange. Then, I will introduce different types of proxy solicitation and their underlying rationale.

Proxy solicitation can be divided into two different categories according to who engages in it: ie either board members as legal representatives of a company (incumbents); or shareholders or third parties (insurgents). Sections III and IV are dedicated to discussion on this.

Finally, I will deal in depth with how proper cost allocation with respect to who bears the costs proxy solicitation can contribute to the revitalisation and increased use of it, especially when it is carried out by minority shareholders. I will also outline possible solutions to cost allocation issues to ensure substantially equal treatment of proxy solicitors. I suggest that CONSOB should integrate these changes into regulation of issuers.

1. Italian Proxy Solicitation Regulation

Regulation of the representation of shareholders at company meetings of Italian listed companies, and the solicitation of proxies for voting at such meetings, has been subject to many layers over time to the point where, today, it is regulated on four different levels one level of which is the by-laws' articles applicable to it.

The first basis of the regulation of proxy solicitation is in Art 2372 of the Italian Civil Code. This provision contains general principles of company law such as that those who have the right to vote at company meetings of listed companies can give a written proxy to another who will attend the meeting. The proxy records how the person who gave it would like the proxy holder to vote on their behalf (ie it cannot be a blanket proxy) and it must relate to a particular meeting and not company meetings in general.

The second layer regulating proxy solicitation is contained in special legislation justified by promoting the efficiency of financial markets and is contained in the following legislation: title III 'Issuers', chapter II 'Discipline of listed companies'; sections II-*ter* 'Proxies'; and III 'Solicitation of proxies', of the TUF. These provisions

Comparison of the United States and United Kingdom' 49 *American Business Law Journal*, 739, 739-803 (2012).

[3] Regulation implementing the decreto legislativo 24 February 1998 no 58 (TUF), concerning the regulation of issuers (last CONSOB Resolution no 21508 of 22 September 2020).

provide that those who have the right to vote can nominate only a single representative to hold their proxy vote for each meeting, except for some specific exceptions contained in Art 135-*novies* (2-3).

These provisions, in turn, refer to the third layer of regulation: title IV, chapters I 'Proxies' and II 'Solicitation of proxies' of the Reg Emitt CONSOB relating to specific procedures and the obligations of a proxy solicitor in Italy which are regulated in detail in Arts 135-139.

The last layer of regulation is contained in companies' by-laws and statutes as well as individual issuers' shareholders general meeting regulations. The regulation governing the Italian stock exchange does not contain anything concerning proxies.

Since the entry into force of decreto legislativo 27 January 2010 no 27 and decreto legislativo 18 June 2012 no 91 (the 'corrective decree') adopting the shareholders' rights directive (integrated in 2017 by the European Parliament and Council Directive 2017/828/UE of 17 May 2017 – so called SRD II) in which the legislator radically changed the provisions in section III (Arts 136–144) of the TUF, any request to confer proxies for specific voting proposals or accompanied by recommendations, statements or other suggestions influencing a vote at a company meeting, addressed to more than two hundred shareholders, is considered a proxy solicitation. The requirement of the person making the request to also own shares in the relevant company was removed from the legislation. Now, proxy solicitation can be made directly by the promoter of shares,[4] by them disseminating a proxy statement and a proxy form without using a broker.[5] A vote on measures is then exercised directly by the promoter or its substitutes.[6] Previously, a proxy solicitation could be promoted only by an intermediary who assumed management of the entire procedure, including the collection of the voting proxies. Today, however, the proxy solicitor can do it themself, also delegating others.

While there may be structural factors that make the USA environment different to that in Italy, proxy solicitation was, in any case, rarely used in Italy until few years ago.

The provisions of the TUF aims to encourage the exercise of voting rights by

[4] Today, this important innovation also allows directors, through an issuer, to conduct in proxy solicitation. Before 2012, it was uncertain, even whether Banco Popolare Soc. Coop. and Sopaf S.p.A. in 2011, as issuers, had promoted proxy solicitation concerning general bondholders' meeting.

[5] The fact that the Italian legislator has decided to remove the requirement for a broker represents a strong push towards a freer financial market, not burdened by the previous bureaucratic constraint.

[6] The notion of proxy solicitation in Italy is very restrictive. For example, in the USA a proxy solicitation also exists when a communication is directed to more than ten shareholders, while in Spain a proxy solicitation is considered to have occurred when one person represents more than three shareholders in respect of a vote. See L. Enriques, 'Quanto è armonizzato il diritto societario europeo?', in G. Carcano et al eds, *Regole del mercato e mercato delle regole. Il diritto societario e il ruolo del legislatore* (Milano: Giuffrè, 2016), 169; K.J. Hopt, 'Corporate Governance in Europe: A Critical Review of the European Commission Initiatives on Corporate Law and Corporate Governance' 12 *New York University Journal of Law & Business*, 139, 155 (2015).

individual shareholders. This is unlike American law, where the proxy solicitation regulation established in 1934 limited management who abused proxy solicitation in the absence of rules to self-elect and remain in office.[7]

2. How Italian Proxy Solicitation Regulation Is Deficient

The Italian legislation should provide for substantive equal participation in corporate decision making for the 'insurgent' (shareholders) and the 'incumbent' (directors).[8] Actually, there is a huge distance between formal and substantive positioning of the proxy solicitor in order to ensure effective equality of treatment of directors and shareholders. While from a formal point of view it may appear that directors and shareholders are treated in the same way with respect to proxy solicitation, in reality – as shareholders are not allowed to finance their proxy solicitation from the issuer's coffers – they are at a substantial disadvantage compared to directors. The Italian legislation should give the same access to corporate funds to both for the purposes of proxy solicitation.

The expenses incurred to carry out proxy solicitation should always be traceable and reasonable, and explicitly indicated in a proxy statement because the purpose of proxy solicitation goes towards greater transparency in corporate choices.

It should also be expected that a reimbursement mechanism be present for insurgent shareholders to recoup proxy solicitation expenses and an *ex post* authorisation for the use of corporate funds by management to carry out a proxy solicitation.

These topics will be explored further below.

3. Overview of Proxy Solicitation in Italy

The number of Italian proxy solicitations carried out has been growing since 2012. Until 2016, most proxy solicitation was carried out by jammers without a specific aim, with there being only a few instances performed by issuers.[9] Since 2017, there have been an increasing number of solicitations carried out but never comparable to the numbers carried out in non-Italian jurisdictions.

In 2020 at the time of writing – largely owing to the Covid-19 pandemic – the only known proxy solicitation carried out in Italy is one in which Banca

[7] On this see CONSOB, 'Nota tecnica in materia di sollecitazione e raccolta di deleghe di voto e voto per corrispondenza', available at https://tinyurl.com/y54g2hm2 (last visited 30 June 2021).

[8] These terms are elaborated on in sections III and IV below.

[9] These are the few cases known to date: Retelit S.p.A. (2015), Italmobiliare S.p.A. (2014), Tiscali S.p.A. (2014), Seat Pagine Gialle S.p.A. (2014), Italcementi S.p.A. (2014), Telecom S.p.A. (2014), Intesa Sanpaolo S.p.A. (2014), Juventus S.p.A. (2014), Milano Assicurazioni S.p.A. (2013), Cape Listed Investment Vehicle in Equity S.p.A. in concordato preventivo (2013), Impregilo S.p.A. (2012), Enel S.p.A. (2012), Parmalat S.p.A. (2012), Banca Carige S.p.A. (2012), Screen Service Broadcasting Technologies S.p.A. (2012), Sopaf S.p.A. (2011), A2A S.p.A. (2011), Meridiana S.p.A. (2011), Eades Group (2011), A.S. Roma S.p.A. (2011), Banco Popolare Soc. Coop. (2011), Mediobanca Banca di Credito Finanziario S.p.A. (2011).

Carige S.p.A. favoured a proxy solicitation for a resolution for a paid and inseparable share capital increase for € 700 million. In 2019, Mediaset S.p.A. promoted the support by proxy voting of a resolution at its own extraordinary general meeting, for the approval of a common cross-border merger plan relating to a merger by incorporation of Mediaset S.p.A. and Mediaset España Comunicación S.A. in Mediaset Investment N.V.

There have been many other cases of proxy solicitation by shareholders in Italy, such as by the issuer Italiaonline S.p.A. during a savings shareholders' meeting in 2019 by Sunrise Investments S.p.A.

That year, 2019, was also lively in this respect for Trevi-Finanziaria Industriale S.p.A. which, during a spring bondholders' meeting, instigated a proxy solicitation as an incumbent. Then it was involved in another proxy solicitation in a September shareholders' meeting which was performed jointly by FSI Investimenti S.p.A. and Polaris Capital Management LLC.

Another shareholders' meeting in which a proxy solicitation was instigated by a shareholder, is the case of Fiber 4.0 S.p.A. (promoter) in the meeting of Retelit S.p.A. in April 2018. During that same spring in 2018, A2A S.p.A. and Lario Reti Holding S.p.A. also carried out another one in respect of an ordinary shareholders' meeting of Acsm-Agam S.p.A.

Between 2017 and 2018, significant Italian banks such as Intesa Sanpaolo S.p.A. and UniCredit S.p.A. also promoted solicitation as issuers in connection with their own shareholders' meetings.

Most often, solicitations are used in connection with resolutions for the conversion of savings shares into ordinary shares, approval of cross-border merger proposals, or election of directors and the fees to be paid to them. The most frequent proponents of solicitations are issuers because they can tap into corporate finances to fund solicitation. However, the number of proxy fights involving insurgent shareholders using solicitation is also growing.

The number of management proxy solicitations promoted in Italy is currently less than ten per year (known cases) and very few proxy fights have been carried out. The most famous fight in Italy was in 2012 involving Impregilo S.p.A. in which a shareholder, Salini S.p.A., carried out a proxy solicitation in connection with a resolution at a shareholders' meeting for the modification of the company's constitution to guarantee minorities representation through election of a director sitting on the board of directors. Salini S.p.A. managed to harness votes to rival Gavio S.p.A. and take on the majority of the board.

From the cases just listed, there seems to be a correlation between the existence of a cost allocation regime and the proxy solicitor's size. In fact, the more it has the opportunity to finance itself, the greater is the range of action by the intermediary on the number of shareholders reachable by proxy.

II. The 'Law & Economics' of Proxy Solicitation

It is interesting to investigate the dynamics behind the reasons why some companies engage in proxy solicitation and to understand, from a purely economic point of view, what advantages they gain in doing so. On the other hand, it is also useful to investigate what limits decisions to engage in proxy solicitation.

A law & economics analysis will lead us on a logical-deductive path aimed at understanding the details of proxy solicitation and its financial implications for its promoter.

1. Proxy Solicitation by Management and Its Problems

The figure of the promoter of proxy solicitation has now been liberalised even in the Italian legal system.[10] In a few years, the promoter has undergone a Copernican-like paradigm revolution within shareholders' meetings.[11] This however, generates strong doubts relating to an obvious conflict of interest issue; indeed, if an issuer is mobilised by its own finances and uses the managers of a company to promote a proxy solicitation, this necessarily places majority shareholders' interests in conflict with those of minority shareholders. The perverse effect of this position leads to the belief that, regardless of the many precautions to protect a minority shareholder,[12] the systematic approach is against them, because they are not guaranteed financial support on which to base their own alternative proxy solicitation. This is support which, conversely, majority shareholders have through an issuer.[13]

On the premise that there is a clear gap between the position of an issuer and that of shareholders in relation to the financing capacity of a proxy solicitation,[14] the goal will be to analyse how, and to what extent, the costs of proxy solicitation initiated by an issuer, should be borne by the company.

I will not address the situation in which directors (as shareholders) are promoters of a solicitation, but when they, as the directors, undertake to promote a solicitation by an issuer. While in the US access to corporate funds is allowed to directors in both cases, in Italy the legislation has ensured greater equality of treatment among challengers to directors' positions because it prevents directors,

[10] In Italy, the neutral expression of 'promoter' is used to both refer to an incumbent or insurgent that engages in proxy solicitation. Previously the promoter was prevented from carrying out solicitations directly but had to rely on a broker to do so.

[11] To emphasize the delay with which Italy has regulated proxy solicitation, see M. Bianca, *La rappresentanza dell'azionista nelle società a capitale diffuso* (Padova: CEDAM, 2003), 208. In the USA, it has been functioning well since the 1930s: see R.C. Pozen, 'Democracy by Proxy' *Wall Street Journal*, available at https://tinyurl.com/y2cmxfhw (last visited 30 June 2021).

[12] Let us think about limits within which incumbents may state proxy voting: see below section II.2.

[13] S. Bruno, *Ruolo dell'assemblea di s.p.a. nella corporate governance* (Padova: CEDAM, 2012), 211-212; G. Sandrelli, 'Note sulla disciplina "anti-scorrerie" ' *Rivista delle società*, 186, 186-223 (2018).

[14] This will be demonstrated below section III.

as promoters on their own account, to draw on corporate funds. However, I consider that the problem persists even in the case of proxy solicitation by an issuer, as long as the latent interest of directors is personal or partisan and not that of the company.[15]

If this is the case, how to run both provisions of proxy solicitation directly from management (incumbent) and from insurgent shareholders should be understood, even when performed through intermediaries, within the balance of shareholders' guarantees.[16] In other words, if anyone can carry out proxy solicitation, including directors, it is necessary to understand its limits in order to guarantee the rights of all shareholders.

a) Suggested Changes to the Enforcement of Provisions to Provide for Substantial Equal Treatment Between Incumbent and Insurgent

Italian legislation provides that a promoter of a proxy solicitation, including where it is a company, must prepare a proxy statement and a proxy form to be sent to shareholders in order to obtain their proxy votes. The rules governing the form of the statement must be adhered to and cannot be modified.[17]

A proxy statement aims not only to provide information to shareholders in accordance with Arts 136–144 Reg Emitt,[18] but also to serve as an appeal to shareholders to complete voting proxies and take a position at a company meeting. Therefore, an incumbent can also include a cover letter with the proxy form when sending the proxy statement.[19] The incumbent may also send informal communications to shareholders designed to gauge support for an initiative. This can take place more easily in cases of proxy solicitation carried out by insurgent shareholders, who will outline criticisms and observations to the board of directors in order to induce a change in strategic policy or in views held by the directors.[20] This is different in cases where the promoter is the incumbent: if shareholders are already satisfied with the policies of a board, the distribution of extra information could be a double-edged sword and bring about situations in which shareholders do not support a board.

At this point, we are already faced with an important debate. If a prior informal communication is considered part of proxy solicitation, as defined in

[15] For this reason, I will use the term 'incumbent' to refer to both circumstances.

[16] An equal position between directors and shareholders with proxy solicitation can be achieved only by ensuring that both have the same access to corporate funds. This issue is analysed further in section V.

[17] M. Bianca, n 11 above, 208.

[18] The *Regolamento Emittenti Consob* is the implementing regulation of *Testo Unico della Finanza*, concerning the issuers' regulation in Italy.

[19] Arts 136-144 state the minimum materials that must be provided when carrying out a solicitation, but do not limit nor prevent other related materials also being provided that support the initiative.

[20] This point will be discussed more thoroughly in section V.

Art 136 TUF, it should then be subjected to strict legislation to protect the interests of the company. If, however, mere informal communication is considered not to come within solicitation practices, its form should of course not be subject to the specific requirements regarding proxy solicitation.

According to CONSOB's technical notes, informal communications are excluded from the scope of the issuers' regulation regarding proxy solicitation.[21] Therefore, all communications not directly aimed at proxy voting requests, such as calls by a broker, are exempted from the regulations. Regardless, the authenticity of information provided to shareholders must be guaranteed. Although determining the content of communications is at the discretion of the promoter, CONSOB will carefully double-check its compliance with the requirements of the issuers' regulation because the risks associated with providing false and misleading information, especially in relation to the election of directors, are very high and irreparable.

When two or more opposing management policies are caught in the crossfire of contested corporate control, then there is likely to be a proxy fight to resolve the matter, in which case the battle will not end by parties sending cover letters, but by them following through with further communications in the standard form prescribed by the issuers' regulation.

b) Limits to Proxy Solicitation Provided by Art 138 Reg Emitt

An incumbent's proxy solicitation involves consideration of many aspects of company meeting principles, and corporate governance. However, the only two limitations that an incumbent may incur within proxy solicitation are:

(a) It cannot vote differently from the instructions received from shareholders (Art 138(4) Reg Emitt); and

(b) It must also accept voting instructions that do not comply with its own proposals (Art 138(2) Reg Emitt).

As others have already observed,[22] these limitations have been put in place to protect shareholders. However instead they may also marginalize them.[23] Although the existence of a proxy solicitation by an issuer is inevitable and

[21] See n 18 above.

[22] See, for example R. Sacchi, 'Voto in base alla data di registrazione e voto per delega dopo l'attuazione della Direttiva azionisti' *Giurisprudenza commerciale*, I, 31, 55 (2012).

[23] Even P.G. Jaeger, 'Le deleghe di voto', in F. Bonelli et al eds, *La riforma delle società quotate* (Milano: Giuffrè, 1998), 101; G. Minervini, 'Art. 136', in G. Alpa and F. Capriglione eds, *Commentario al testo Unico delle disposizioni in materia di intermediazione finanziaria* (Padova: CEDAM, 1998), II, 1243; R. Sacchi, 'Sollecitazione e raccolta delle deleghe di voto', in F. Bonelli et al eds, *La riforma delle società quotate* (Milano: Giuffrè, 1998), 387; F. Ghezzi, 'Art. 136-140 e 144', in P. Marchetti and L.A. Bianchi eds, *La disciplina delle società quotate nel Testo Unico della Finanza d.lgs. 24.02.1998 n. 58* (Milano: Giuffrè, 1999), 1267, agree that the issuer must be inhibited from promoting a proxy solicitation. In spite of this, A. Gambino, 'Art. 139', in G. Alpa and F. Capriglione eds, *Commentario al testo Unico delle disposizione in materia di intermediazione finanziaria* (Padova: CEDAM, 1998), II, 1266, support the idea that an issuer can promote solicitations on its own.

sustainable, at the same time the protections provided by CONSOB in order to avoid risks to shareholders in connection with them do not seem adequate. It would be far more desirable to introduce an organic discipline of cost allocation, just to overcome the enormous inequality among incumbents (directors through an issuer) and insurgents (the shareholders or third parties wishing to be promoters).

While agreeing that the legitimacy of proxy solicitation by an issuer without adequate safeguards represents a setback in the prevention of abuse in the exercise of proxy voting compared to the earlier rules (previously the issuer was not allowed to promote solicitation),[24] the solution should not be to prevent an issuer from promoting a proxy solicitation,[25] nor to always impose a *two-way proxy* rule.[26]

Adequate regulation of cost allocation would achieve a balance between healthy competition in governance and shareholders' corporate guarantees.[27] Nor is the allocation of the costs of proxy solicitation to an issuer contrary to the principles that has inspired European and Italian legislators and regulators.[28] Indeed, differently from the views of others,[29] the solicitation of proxies regulated by Italian legislation, is apparently done in the interests of the promoter since the eventual success of a solicitation would benefit the entire corporate structure and not only those who have supported the promoter, especially when the solicitation comes from shareholders or third parties wanting to gain control of a company.[30] Therefore, if the ultimate goal of allowing proxy solicitation is

[24] R. Sacchi, n 22 above, 55.

[25] That obstacle is based on the extension of the ban on a company to exercise a vote in its name and to vote by proxy. According to R. Sacchi, n 22 above, 57, this view is also supported by the European Parliament and Council Directive 2007/36/EC of 11 July 2007 as called SHRD (Art 10.1, second period) as it would prevent the national legislature from prohibiting the company receiving unsolicited proxies, not even to prohibit the company to assume the role of promoter. The author would consider this rule of Reg. Emitt. illegal because it is contrary to Art 2357-*ter* Civil Code.

[26] This device, even if supported by R. Sacchi, n 22 above, 59; P.G. Jaeger, n 23 above, 108; G. Napoletano, 'Deleghe di voto (Artt. 136-144)', in L. Lacaita and V. Napoleoni eds, *Il testo unico dei mercati finanziari: società quotate, intermediari, mercati, opa, insider trading: commento al d.lgs. 24 febbraio 1998, n. 58* (Milano: Giuffrè, 1998), 130, represents a circumvention of the problem rather than a solution. Also, it would probably induce a significant reduction in the probability that proxy solicitation will be promoted in favor of all shareholders, rather than to a certain category of them. Also, P.G. Jaeger, 'Convenzioni di voto – rappresentanza azionaria', in M. Rotondi ed, *Inchieste di diritto comparato, I grandi problemi della società per azioni nelle legislazioni vigenti* (Padova: CEDAM, 1976), V, 699, states that 'it is clear that, by itself (the *two-way proxy* rule), it would not offer substantial chances to members who wish to organize opposition to the control shareholder'.

[27] The Italian regulation merely requires that all costs of proxy solicitation must be borne by the promoter (Art 136(9) Reg Emitt).

[28] See A. Blandini, 'L'intervento e la rappresentanza in assemblea e l'art. 10 della Direttiva 2007/36/CE: prime considerazioni e proposte' *Le Società*, 511, 515 (2009).

[29] Contrarily, R. Sacchi, n 22 above, 58, considers that the interest protected during proxy solicitation is always that of the shareholders. However, CONSOB regulated in the same way as in the US, in which the guaranteed interest is that of the promoter, who also bears all the costs of proxy solicitation.

[30] See G. Presti, 'La nuova disciplina delle deleghe di voto' *Banca Impresa Società*, 35, 56 (1999). The author, based on the assumption that the costs are borne by the promoter, argues that

good corporate governance, I do not understand why the cost of the solicitation should not fall on the issuer.[31]

2. Proxy Solicitation by Insurgent Shareholders

Proxy solicitation by insurgent shareholders in Italy takes place in the same way as that done by an incumbent and therefore, the applicable regulatory regime in both cases is the same.[32] Other jurisdictions prefer divergent approaches to regulating these two situations, especially given the weak position of the shareholder as a promoter. For example, the US Securities Exchange Commission has dedicated Rule X-14A-8 of the Regulations specifically to proposed proxy solicitation by shareholders.[33]

The ontological choice of the Italian legislator has been underpinned by a principle of substantial equality of treatment between directors and shareholders, to avoid inequalities for the benefit of management.[34] I consider that therefore proxy solicitation expenditures are largely not discussed in Italy.

To say that on a formal level positions of incumbent and insurgent are substantially equal, albeit within the limits of Art 138 Reg Emitt in the case of incumbent proxy solicitation, is untrue. Access to corporate funds, therefore, becomes a discriminating factor in promoting proxy solicitation.

III. Expenses of Proxy Solicitation – Efficient Allocation of Costs

A proxy solicitation's promoter incurs various costs in connection with the solicitation, depending on the size of the company, its number of shareholders,

the Italian legislator, through the solicitation of proxies, uses a private and selfish interest of the promoter to obtain a general improvement in corporate governance practices.

[31] See P.G. Jaeger, n 23 above, 701, where, while acknowledging that at the time (as now) a solution has not been reached, he proposed two ways to fund proxy solicitation: the company should bear the costs of proxy solicitation in both cases; or when directors, also through an issuer, are promoters of a solicitation, the costs should be borne by the directors.

[32] For a complete examination of the current Italian legislation, see E. Ricciardiello, 'La nuova disciplina in materia di sollecitazione delle deleghe di voto: inizia la stagione italiana dei *proxy fights*?' *Giurisprudenza commerciale*, I, 151, 151-176 (2012); M. Campobasso, 'La tutela delle minoranze nelle società quotate: dall'eterotutela alla società per azioni "orizzontale" ' *Banca borsa titoli di credito*, 139, 142 (2015); R. Sacchi, n 23 above; M. Bianca, n 11 above.

[33] The 'shareholder proposal', however, has very different aims to how proxy solicitation promoted by shareholders is regulated in Italy. In fact, when one or more shareholders are willing to submit a proxy solicitation in the US, they are required to request management to include their voting proposal in the proxy statement of the issuer. However, if the issuer, by means of the directors, opposes the shareholders' proposal, it must also include in its proxy statement that of the shareholders' supporting their proposal. The issuer must also include in the proxy form the specific choice through which shareholders can express their approval or disapproval of the proposal. Alternatively, in cases in which an issuer does not promote a solicitation, or with a directors' election, the issuer must provide to insurgents a shareholder list, so they can disseminate their proxy statement independently.

[34] This is particularly emphasised by P.G. Jaeger, n 26 above, 697.

and the nature and purpose of the solicitation.

Previous experience, especially that in the US,[35] has shown that the whole process of solicitation requires assistance and the costs are unpredictable. It should be possible, for example, for a group of shareholders to introduce proxy solicitation to counter proxy solicitation organised by incumbents, especially in the case of election of board members. In these circumstances, however, costs can soar dramatically and cannot be reduced, except by not progressing with such solicitation.

Before promoting a proxy solicitation, both insurgents and incumbents should be required to estimate their potential expenditures, as well as to prepare a strategic plan. Such costs include the fees of lawyers, accountants, and other professionals, such as public relations experts, as well as proxy advisors. Without these skills, rarely will a solicitation hit the mark. These are essential sums, and not everyone with an interest in the outcomes for corporate governance can afford them.

This is where one of the problems of the interpretation and enforcement of Italian proxy solicitation regulation arises. While other jurisdictions regulate who bears the costs of solicitation,[36] the Italian legislator completely fails to address such cost allocation, almost as if it is not concerned with the proper functioning of the proxy solicitation itself.

As appears to be clear from the preparatory work undertaken to integrate into issuers' regulation, CONSOB has addressed the question of the costs of proxy solicitation. However, its conclusions have so far been short-sighted. If the aim is for promoters to be able to reach all shareholders without adding excessive cost to the proxy solicitation's promoter who may also be a minority shareholder, to permit that a notice can be published in the press, instead of requiring the use of mechanisms with significant costs borne directly by the issuer, as CONSOB does, is not enough. I consider that giving a proper discipline to proxy solicitation costs will essentially contribute vitality to the existing provisions in Italy, which have not caught on as in other European equity markets.[37]

Regarding the dissemination of proxy material, CONSOB requires that a proxy form, accompanied by a proxy statement, must be delivered to the parties concerned, including through intermediaries or depositaries. Awareness of the high costs of distribution, albeit at the expense of proper corporate information (the final scope of the provision), has, however, provided shareholders with the option of turning to intermediaries or depositaries of the shares to indirectly acquire the documentation for the solicitation. However, such a procedure does not prevent an intermediary from conveying information to shareholders, either directly or through brokers. The preference is to waive this last procedure owing

[35] For details about leading cases in the State of New York and Delaware, see section IV.1.

[36] The US Federal legislation provides for complete disclosure of a proxy contest's expenditures in Item 4 of Schedule 14A to Rule X-14.

[37] Surely, there are also structural factors that make the US situation different to the Italian one.

to the high costs the developer would need to support it.

This solution is incomplete. If, indeed, the problem of proxy solicitation relates to its costs, then CONSOB should provide cost-sharing mechanisms. In this sense, it is important to analyse: firstly, the limits within which management can use corporate funds (as an incumbent), a matter completely ignored even by CONSOB;[38] and, second, if reimbursement of insurgents' expenditures can be made directly from corporate funds.

In Italy today, insurgents rarely instigate a proxy solicitation against management policies. Often, insurgents are firmly anchored to the granite majorities of shareholders that characterize the Italian financial market (which, however, are slowly crumbling).

Moreover, a similar mechanism, and not even a foreign teleological purpose of a forecast of such content,[39] can be seen in the Italian Civil Code. Art 2393-*bis*(5) is substantially about a similar case to corporate liability, although in the field of a minority's liability action against directors. In the case of acceptance of such a claim by a court, a company is required to refund the costs of the minority, if they cannot be recovered against directors. Leaving the discussion on this point to section V below, I wonder whether to extend this provision, even in the case of proxy solicitation, complies with constitutional principles (Art 3 of the Italian Constitution relating to freedom and equality) and the rationale for proxy solicitation rules.

1. The Transparency Requirement and Costs of Soliciting Proxies

Proxy solicitations are operations that take place within a regulated stock market, and so CONSOB always has the power to check whether operators are behaving in a way that does not harm the public interest.[40] As a result, the issuers' regulation requires that anyone who wishes to promote a proxy solicitation to send a notice to CONSOB, who – after having noted the information contained in it – might want the promoter to provide further information.[41]

Nothing is yet enshrined regarding the costs of a solicitation. In other words, CONSOB, unlike its international 'colleagues',[42] has decided not to deal with the

[38] Nothing is said in the Italian issuers' regulation about this topic.

[39] I support this idea so that it can be introduced into our legal system, either at regulator level or soft law.

[40] See M. Bianca, n 11 above, 274, where he argues that the fact that the regulator prepared special schedules on which promoters are required to draw up a proxy statement and a proxy form, is symptomatic of how CONSOB addresses accurately identifying the disclosure requirements aimed at comprehensive and transparent information addressed to the recipients of proxy solicitation.

[41] It happened, for example, in Banca Carige S.p.A.'s proxy solicitation concerning a general bondholder' meeting in 2012.

[42] The Securities Exchange Commission in Item 4 of Schedule 14A (in accordance with Section 14 (a) of the Securities Exchange Act of 1934) requires a promoter to make disclosures about the source, amount, and extent of the expenses related to the proxy solicitation. In the European legal systems I reviewed (ie those in Italy, France, Spain and Germany), mostly the promoter declares in a proxy statement that all expenses will be borne by them.

subject of solicitation accounting, even if it is a fundamental part of the reimbursement charged to corporate funds. If CONSOB does not demand transparency in spending, it is difficult to start discussing the limits directors face in the use of corporate funds for promoting proxy solicitations. Neither will we be able to provide reimbursement of expenses in cases of insurgents' solicitation.

For these reasons, we therefore cannot remain indifferent to this serious gap, which the Italian regulator does not seem to want to fill. I hope that CONSOB will take notice of this issue and add it in the attached no 5B and 5C of Reg Emitt.[43]

2. Use of the Internet for Proxy Solicitation

CONSOB has adopted rules permitting promoters to distribute proxy materials through Internet websites and to inform shareholders using just an availability notice.[44] Before a shareholder meeting, a promoter can send a written notice of Internet availability to shareholders, indicating that proxy materials are available and explaining how to access them.[45] Otherwise CONSOB should provide a database in which all promoters could make their proxy materials available.[46]

This could also be an opportunity for CONSOB to regulate the relationship between promoters and intermediaries/brokers[47] and the proxy solicitation's public communications (all kinds of information given to the market) that today are subject only to the general rules of Art 114 TUF. Distributing proxy materials via the Internet could reduce the cost of soliciting proxies to the benefit of insurgent shareholder.[48]

IV. Solicitation by an Issuer – Use of Corporate Funds

The most significant benefit that board members have in a proxy solicitation as an issuer, is certainly direct and unconditional access to corporate funds with

[43] These attached no 5B and 5C provide example models of proxy statements and proxy forms. The first model requires that only the promoter must indicate any funding received for the promotion of the solicitation.

[44] The SEC has done this too, see Rule 14a-16. It became mandatory from 1 January 2009 for all issuers with respect to proxy solicitation commencing on or after that date.

[45] See A.L. Goodman et al, *A Practical Guide to SEC Proxy and Compensation Rules* (New York: Wolters Kluwer, 6th ed, 2019), 9-23.

[46] See D.O. Garris and C.A. Duke, 'SEC Proposes the Use of the Internet for Proxy Solicitations' 20 *Corporate & Securities Law Advisor*, 11-12 (2006). The SEC, for example, has established an Edgar archive, but it does not allow promoters to use the Edgar system in lieu of a publicly accessible Internet website, because of technological limitations. Otherwise, giving this public interest function to CONSOB or another authorized private company will guarantee shareholders' adequate information.

[47] I am referring here to how, and under which limitations, intermediaries and brokers are asked to distribute proxy materials to beneficial shareholders. See in more detail E.R. Aranow and H.H. Einhorn, n 2 above, 212.

[48] See J.H. Trotter and J.G. David, 'Proxies Prepare to Go Digital: The SEC's E-Proxy Solicitation Proposal' 14(4) *Corporate Governance Advisor*, 5, (2006).

which to defray costs related to the solicitation. As this privilege allows management to preserve themselves in the top positions with only the support of the majority shareholders, it is necessary to limit the use of corporate funds in this way. To that end, I will summarise the relevant principles that American jurisprudence has been developing over recent decades. Then, I will examine the main relevant judgments of the courts of New York, Delaware and the UK.[49]

1. General Principles in Non-Italian Jurisdictions – The Concept of Reasonable Expenditure

Today, both in the Italian legal system and in international jurisdictions, how management of a company can use corporate funds is not generally regulated.[50] However, principles are contained in judicial decisions regarding this, and I refer to these principles to extrapolate on their ideas and to support their integration into the Italian legal framework, of course with due consideration being given to the structural differences in corporate governance between different legal systems.[51]

The case law that I will now elaborate on relates to the following assumption:[52] the principles according to which management must be monitored emanate from

[49] These principles have been developed in the US mainly with reference to cases in which proxy solicitation is promoted by directors. However, in the Italian national financial market in which there is identity between the controlling shareholder and directors, the interest that motivates directors to take charge of solicitation is the same personal or partisan one. Therefore, even in the case of proxy solicitation by an issuer in Italy, we can observe the same spending limits.

[50] The Securities Exchange Commission regulates the disclosure of expenditure by management but does not regulate management's access to corporate funds, see n 42 above.

[51] See G. Bachmann, 'Der "Europäische Corporate Governance-Rahmen" - Zum Grünbuch 2011 der Europäischen Kommission' *Wertpapier-Mitteilungen*, 1301, 1305 (2011); T. Baums, *Bericht der Regierungskomission Corporate Governance. Unternehmensführung, Unternehmenskontrolle, Modernisierung des Aktienrechts* (Köln: Verlag Dr. Otto Schmidt, 2001); C. Bellavite Pellegrini, '*Corporate governance* e assemblea delle società quotate in Italia: un'indagine empirica' *Rivista delle società*, 401, 434 (2006); L. Enriques, 'La *corporate governance* delle società quotate: sfide e opportunità' *Giurisprudenza commerciale*, I, 493, 493-509 (2012); A.M. Fleckner and K.J. Hopt, *Comparative Corporate Governance* (Cambridge: Cambridge University Press, 2013); H. Fleischer, 'Zukunftsfragen der Corporate Governance in Deutschland und Europa: Aufsichtsräte, institutionelle Investoren, Proxy Advisors und Whistleblowers' 40(2) *Zeitschrift für Unternehmens- und Gesellschaftsrecht*, 155, 165 (2011); P. Hommelhoff and K.J. Hopt, *Handbuch Corporate Governance* (Köln: Schäffer-Poeschel Verlag Stuttgart, 2nd ed, 2009); P. Mäntysaari, *Comparative Corporate Governance: Shareholders as a Rule-maker* (Heidelberg: Springer, 2005); P. Marchetti, 'Il ruolo dell'assemblea nel T.U. e nella corporate governance', in F. Alcaro et al, *Assemblea degli azionisti e nuove regole del governo societario* (Padova: CEDAM, 1999); R. Kraakman et al, *The Anatomy of Corporate Law. A Comparative and Functional Approch* (Oxford: Oxford University Press, 3rd ed, 2017); P. Davies, *Introduction to Issuer Law* (Oxford: Oxford University Press, 2nd ed, 2010); G. Estaban Velasco, *El gobierno de las sociedades cotizadas* (Madrid-Barcelona: Marcial Pons-Garrigues & Andersen, 1999), 678; A. Pomelli, 'Offerta pubblica d'acquisto o scambio prevalente ed altre questioni aperte in tema di offerte concorrenti' *Giurisprudenza commerciale*, II, 682, 687 (2017).

[52] Corporate property theory can be relevant in Italy in the light of directors' duties of care and loyalty.

'corporate property theory',[53] and state that the resources of an issuer (being owned by all shareholders) must be used solely for corporate purposes and therefore for the company.[54] The use of corporate funds by board members for proxy solicitation in their own interest, or even majority shareholders aiming just to keep control of the company, is improper and should not be tolerated.

Further, the application of these corporate governance principles to proxy solicitation poses serious questions that ought to be analysed.[55] Without a doubt, management can draw on corporate funds to prepare and organize general meetings, but an important and separate question is the extent to which board members can reasonably spend corporate resources to influence voting decisions.

One of the first judgments on this issue was handed down in 1906 in the UK, namely *Peel* v *London & Northwestern Railway Co.*[56] In this case, the Court of Appeal ruled that it was legitimate for management to use corporate funds as long as the aim was supporting their policies and not the interests of directors as individuals.[57]

On this issue, several rulings have been made in the State of Delaware,[58] all censuring management's conduct designed to defend their policies at company meetings. In other words, it was considered reasonable for management to have expended corporate funds when done with the purpose of disseminating information to shareholders on corporate policies aimed at increasing awareness and judgment in the exercise of voting rights, whether or not that information was persuasive. Any other costs, expended only with the aim of directors maintaining their roles within management, are to be considered improper.[59]

This stance, however, was not as well accepted then as it is now. Again, under the law of the State of Delaware, a decision was made a few years later in *Steinberg* v *Adams*[60] that criticised the approach to the concept of 'reasonableness', as it was understood at that time, in connection with on what management should

[53] As everyone knows, the debate began with A. Berle and G. Means, *The Modern Corporation and Private Property* (New York: Transaction Publishers, 1932).

[54] Any other interest pursued must be considered *ultra vires*. See E.R. Aranow and H.H. Einhorn, n 2 above, 491, fn 2.

[55] See ibid 492.

[56] [1907] 1 Ch. 5, a dispute had arisen between the management of a company and a group of its shareholders who challenged the unconditional use of corporate funds by directors to circulate information to shareholders in favour of strategic policies that had already been undertaken. The judge ruled in favour of the directors who had made use of corporate funds because of their duty to inform shareholders about the policies undertaken and the reasons in support of them, although the purpose of the information was to also persuade and influence shareholder voting.

[57] E.R. Aranow and H.H. Einhorn, n 2 above, 493.

[58] See Hall v Trans-Lux Daylight Picture Screen Corp., 20 Del. Ch. 78 (1934); Empire Southern Gas Co. v Gray, 29 Del. Ch. 95 (1946); Hand v Missouri-Kansas Pipe Line Co., 54 F. Supp. 649 (D.C. Del. 1944).

[59] E.R. Aranow and H.H. Einhorn, n 2 above, passim.

[60] *Steinberg* v *Adams*, 90 F. Supp. 604 (S.D.N.Y. 1950) (applying Delaware law). See also *Rosenfeld* v *Fairchild Engine & Airplane Corp.*, 309 N.Y. 168, 176 (1955).

be permitted to expend company money. The judges in this case decided that the matter of reasonable expenses in a proxy solicitation was covered by the 'business judgment rule'.

This issue was not so clear, even in case law in the State of New York. In a leading case in 1907, *Lawyers' Advertising Co.* v *Consolidated Ry. Lighting & Refrigerating Co.*,[61] the Court considered that unusual, superfluous, or unnecessary expenses, although incurred in good faith, should not be borne by the company.

Some rulings over the years have addressed reasonable expenditure limits incurred by directors during a proxy solicitation. The issue was discussed in at least two other cases in the US. In *Cullom* v *Simmonds*,[62] the Supreme Court of the State of New York declared, based on the minority shareholder's claim, that costs expended of just over a hundred thousand dollars on proxy solicitation were unreasonable. However, in *Levin* v *Metro Goldwyn Mayer Inc* 1967 (New York),[63] five shareholders complained that directors had improperly committed the issuer to pay for lawyers, public relations experts and consultants, as well as committing internal human resources for the purpose of proxy solicitation. The action challenging the costs was dismissed because they were held to have been expended for the greater interest of access to corporate information.

2. Identification of Costs that Management May Charge to the Company for Proxy Solicitation

From the above discussion of the principles that directors must follow with a proxy solicitation, two different legislative approaches appear possible.[64]

The first is that directors have the right to use corporate funds for expenses reasonably incurred to convince shareholders of the correctness of their policies, provided that these policies are pursued in good faith and in the interest of the company, and for proxy conferment.

The second, more restrictive approach, is that it is possible for a board of directors to use the resources of a company only to inform shareholders of management decisions taken and the reasons in support of them, and also with the help of proxy advisors, for the purpose of achieving a quorum at a company meeting, and nothing more. The difference lies in the contemplation of the proxy conferment.[65]

Regardless of the difficulty of discerning which expenses are incurred for a constitutive quorum and not to influence voters,[66] the real rock that is found in the

[61] 187 N.Y. 395 (1907).
[62] 285 App. Div. 1051 (N.Y. App. Div. 1955).
[63] 264 F. Supp. 797 (S.D.N.Y. 1967).
[64] See E.R. Aranow and H.H. Einhorn, n 2 above, 498-499.
[65] M. Bianca, n 11 above, 292.
[66] The difference it is whether the rules will allow the directors to keep operating despite the adversity of the shareholders, or simply allow them to give correct information on the work and then let the shareholders be free to decide them.

Italian stock market is determined by a massive presence of majority shareholders who make management and control hardly contestable. This, however, is slowly becoming less the case. Just think now there is a massive presence of institutional investors and hedge funds in Italy, especially in the banking sector.[67]

The first legislative approach outlined above – though more delicate and difficult to achieve – should be chosen, provided that, in the face of greater ease of use of corporate funds by directors in a proxy solicitation, there are strict limits provided ex post to evaluate directors' behaviour and the reasonableness of expenses, maybe even by a court.[68]

It remains now to review some practical circumstances so that the underlying reasoning for the preferred approach is clear. Directors must, of course, submit a notice of a company meeting to shareholders at the company's expense. Nevertheless, if directors intend also to attach a proxy statement and form to the notice, it is relevant to assess the legitimacy of their behaviour.[69] The legality of forwarding a proxy solicitation statement jointly with a notice of a meeting would not be legally questioned, as a proxy statement must specify the date of the meeting and must not also say whether the notice has already been sent or not.[70] Indeed, while lowering the cost of any distribution of the proxy statement, and the proxy form supports efficiency in the expenditure of corporate funds, as at the same time they can be used to persuade shareholders, rather than simply inform them. It becomes difficult to understand what kind of expenditure was incurred by directors in the interest of the company and what was incurred with

[67] The latest detailed data on this can be found in: Glass Lewis, 'Guidelines, 2017 Proxy Season Italy', available at www.glasslewis.com; ISS, '2016 Europe Summary Proxy Voting Guidelines', available at www.issgovernance.com; Frontis Governance, 'Principi di Corporate Governance e Politiche di Voto per il Mercato Italiano. Stagione assembleare 2018', available at www.frontisgovernance.com; F. Bianconi and S. Bruno, *Evoluzione degli assetti proprietari ed attivismo assembleare delle minoranze (proxy season 2013)* (Roma: Luiss Ceradi-Georgeson, 2013), 7, available at https://tinyurl.com/y6d6jfre (last visited 30 June 2021); and, with reference to banking industry, conference proceedings, *Assemblea e Corporate Governance: Proxy Season 2014*, Milan, 2 July 2014.

[68] Behind this scheme of proxy solicitation costs' reimbursement lies the same reasoning governing a minority's liability action against directors. As shareholders are acting on the behalf of a company, pursuing its institutional interest, shareholders should not incur costs when acting without any personal motive.

[69] In France, notice of a company meeting and proxy statement are systematically sent together. Regarding attempts by directors to abuse this process, see Y. Guyon, *Droit des affaires. Droit commercial general et Sociétés* (Paris: Economica, 12th ed, 2003), I, 301.

[70] From reading Art 136(2) Reg. Emitt. Discloses no textual data against this interpretation. An extract of the article states: 'The notice shall indicate: a) the identity of the promoter and the company issuing the shares for which the proxies are sought; b) the date of the shareholders' meeting and the list of items at the agenda; c) how the proxy statement and the proxy form are published as well as the Internet site that these documents are available on; d) the date beginning from which the party with the voting right may request the prospectus and the delegation form from the promoter or view it at the market operator; e) the proposals for which the solicitation is to be carried out'; all data that has to be contained in a proxy statement as provided by the model (attached no 5b of CONSOB issuers' regulation).

the intent of persuading shareholders as to their exclusive interest. This is because the actual costs of in notifying shareholders of the meeting, and those for the distribution of the proxy statement and the proxy form, may be the same.

The same can be said of expenditures related to the organisation of a general meeting. There are no reasons why proxy advisors' fees may not also be covered, as such advisors, for example, assist management in developing industrial and strategic policies, in counting proxies at the pre-meeting and in tabulating votes during the meeting.

The proxy advisor's role is broad, owing to the need for them to take into account a variety of corporate interests and matters. Board members may also use these advisors to ensure the presence of shareholders at a meeting for the proper functioning of the company.[71] However, board members could also instruct proxy advisors to plan a proxy solicitation, thus creating a link between the interests of the company and those of the directors.

To what extent, then, is it reasonable to limit the use of corporate funds to pay the fees of proxy advisors?[72] This question should be discussed on a case-by-case basis, returning first to the directors' duty of care, and possibly to later judicial review of such payments.

V. Solicitation by Insurgent Shareholders – Reimbursement of Expenses by Corporate Funds

In this section, I will discuss how the costs of solicitation should be allocated when shareholders, in any form or composition, are promoting a proxy solicitation.

In identifying the exact expenses related to a proxy solicitation that should be reimbursed by the company to shareholders who initiated it, it is inevitable to refer to the concept of reasonableness, already referred to in section IV.1, about the use of corporate funds by management.

All expenses to compensate professionals, such as financial advisers and lawyers, for printing the proxy materials, notifying shareholders and the cost for proxy advisors' services, should be chargeable to the company, even if judicial review regarding the reasonableness of total expenditure should also exist.

A particular issue can be envisaged relating to the amounts necessary, according to Art 136(7) Reg Emitt, for the transfer of identification data to intermediaries involving the number of registered shares of the issuing company and the relative amount of shares of different classes, those who have the right to vote, who have not expressly prohibited communication of their data, and those

[71] E.R. Aranow and H.H. Einhorn, n 2 above, 500; G. Balp, *I consulenti di voto* (Milano: Egea, 2017), 46.

[72] With the same yardstick, the expenses paid to public relations advisors should also be considered, as well as those for legal and financial advisors. Relevant is the extent to which directors can claim the use of expert advisors of the issuer for the organisation of a proxy solicitation.

who have opened accounts as intermediaries and the amount of the issuer's shares respectively recorded on these accounts, as well as those of members. The promoter should bear these costs, which are essential to begin a proxy solicitation, but they might not all need to be repaid by the company to the promoter.[73]

1. Reimbursement of Costs to a Proxy Solicitor Where There Is Successful Proxy Solicitation by Shareholders

Reimbursement of expenses is at the heart of this entire discussion: to give vitality to proxy solicitation in the Italian financial system, the regulator has to provide that the issuer should be accountable for part of the expenses incurred and anticipated by insurgent shareholders and downsize access to corporate funds by management. We should not forget that the concept of proxy solicitation was created in a different legal context and that adjustments to it will be needed because of the different distributions of power within Italian corporate structure.[74]

Considering the urgency of drafting legislation or regulation that provides suitable finance tools in the hands of proxy solicitor as different from directors, I suggest two different solutions based on the current legislation.[75]

Judgments handed down in the US in the 1950s are leading the way for opening a debate on these topics in Italy.[76] In the US, the question was whether reimbursement by a company to a promoter of proxy solicitation expenditure could be subject to approval at a general meeting. Making reimbursement subject to proper corporate governance oversight, ensures both the use of corporate funds to a reasonable extent by directors, and that the reimbursement of insurgents' expenses are instrumental in achieving consistency of information at the base of expressing an informed vote. Yet, if the reimbursement of expenses incurred by insurgents is subject to the balance of corporate powers, it becomes legitimate in itself. Therefore, referring any decision on the matter regarding the reimbursement of expenses to a general meeting is unnecessary,[77] since authorisation of any illicit payments, though unanimous, has no effect and the related expenses could not even be reimbursed. For the same reasons, if a payment is legitimate, the shareholders' approval is not required.[78]

The foregoing leads me to say that, beyond all the issues arising relating to authorisation of reimbursement of expenses incurred by insurgents,[79] it might

[73] Related costs could be high, but when using e-proxy solicitation they may be less.

[74] For more detail, see A.L. Goodman et al, n 45 above, 3-9, concerning practical aspects of proxy solicitation.

[75] The different legislative choice is in the fact of making (or not) by the general meeting authorization of the expenses' reimbursement incurred by the insurgent.

[76] As noted above, see *Steinberg* v *Adams*, 90 F. Supp. 604 (S.D.N.Y. 1950); *Rosenfeld* v *Fairchild Engine & Airplane Corp.*, 309 N.Y. 168, 176 (1955).

[77] I am talking about ordinary general shareholders meeting, not extraordinary.

[78] E.R. Aranow and H.H. Einhorn, n 2 above, 512.

[79] If a shareholder has made use of proxy solicitation, this means that they do not have the

be logically and legally argued that, where the costs of solicitation incurred by insurgents are reasonable, documented, and subject to a corporate purpose, (not merely the purpose of gaining personal control of the company), then management could, also without authorisation at a general meeting, order the reimbursement of such expenses via an automatic mechanism.[80]

Another circumstance, not so remote, can be seen when insurgents remain partially successful. In these cases, having a hypothetical automatic mechanism for reimbursement would certainly be more controversial, and requiring a decision of a general meeting would be preferred.

2. Reimbursement of Costs to Proxy Solicitor Where Proxy Solicitation Is Unsuccessful

There are authors who have argued with some difficulty that even in cases of the failure of proxy solicitation, a company should be required to reimburse the costs to the promoter.[81]

If the corporate purpose requirement is replaced by greater disclosure of information about corporate policies, then it would also be justified for expenses incurred by insurgents who are unsuccessful in a proxy solicitation, to be reimbursed, as their actions are teleologically related to the increase of the degree information provided to the shareholder.[82] In other words, even if the scope of the failure ensures a debate between two or more opposing views about management decisions, shareholders can simply benefit from the debate in terms of increased awareness.

Others have also supported the view that[83] if directors use corporate funds to disseminate corporate information that is intended to influence shareholders' judgments when voting on a resolution (which most of the time are proposed by management), the same should be allowed in cases where proposals concern issues other than those of the directors.

However, there are two issues related to such an approach. The first is the

majority of votes. In this sense, then, if the insurgent shareholder managed to get a substantial number of proxies in order to push through its proposal to vote once, it does not mean that they would fail again, especially when the general meeting will be called to vote authorising the reimbursement of expenses incurred by them.

[80] Art 2393-*bis*(5) of the Civil Code provides that shareholders can initiate action against directors for damages arising from wilful misconduct or negligent behaviour. And, if the claim is accepted, the company reimburses the plaintiff's judicial expenditures and those incurred for the ascertainment of the facts which the judge does not charge to the losing party or which may not be possible to recover upon enforcement against them. This is the only case in which the legislator foreseen an automatic reimbursement in case of winning suit without general meeting's approval. This mechanism could be also used in the context of proxy solicitation.

[81] See F.C. Latcham and F.D. Emerson, n 2 above, 13; S.W. Mintz, n 2 above, 96.

[82] See D.M. Friedman, n 2 above, 958.

[83] I am referring here to the US literature, the most attentive to these issues being: Note, 'Proxy Solicitation Costs and Corporate Control' 61 *Yale Law Journal*, 229, 229-237 (1952).

risk of encouraging proxy solicitations that are completely irresponsible and not adequately justified. The second links to formulas for calculating the percentages of reimbursable expenses:[84] for example, the issuer could supply reimbursement based on the percentage of votes obtained by proxy solicitor, or by affording full reimbursement of the expenses necessary for information's disclosure and partial for fees of lawyers, accountants, and other professionals.

VI. Conclusion – Actualisation of Italian Proxy Solicitation Rules

Hoping that this work will help to encourage debate on the issue of cost allocation for different types of proxy solicitation for companies listed on the Italian stock exchange a debate that has not yet begun in Italy, I conclude by offering my own view, which is similar to the laws now in force in the States of New York and Delaware, on how expenditure should be borne for proxy solicitations in Italy.

Returning to the distinction between formal and substantive positioning of the proxy solicitor, in order to ensure effective equality of treatment between directors and shareholders, the same access to corporate funds should be given to both of them. Also, access to such funds should be limited as follows funds:

(a) Expenses must always be traceable and reasonable (and indicated in the proxy statement);

(b) The purpose of proxy solicitation should further greater transparency in corporate choices; and

(c) The company meeting should approve reimbursement to insurgent shareholders or authorises ex post the use of corporate funds by management, following a judgment of fairness by an independent internal committee.

[84] See F.D. Emerson and F.C. Latcham, n 2 above, 436. These two American authors are of the opinion that the rate of reimbursement should be conditioned to the ability of the promoter to obtain proxies, the so-called 'proportional reimbursement formula'.

Corporate and Financial Markets Law

Shifting the SME Corporate Model Towards Sustainability: Suggestions from Italian Company Law

Laura Marchegiani*

Abstract

While Corporate Social Responsibility (CSR) is currently at the center of debates regarding company law all over the world, the discourse on this topic remains predominantly focused on large enterprises operating at a multinational level. The purpose of this paper is to introduce some reflections on the relationship between CSR and smaller companies. It examines what organisational solutions can be found in the regulatory framework of Italian company law to encourage small and medium-sized enterprises (SMEs) to transition towards sustainability. Firms of this dimension represent ninety nine per cent of European businesses and account for more than ninety per cent of the world's business enterprises that need to make this transition. Despite the fact that SMEs are defined as the 'backbone of Europe's economy', organisational models of sustainability in SMEs have not yet been studied in depth, and the usefulness of company models that combine altruistic and lucrative corporate purposes, and above all impose a sustainable manner of action on company activities, are still to be analysed comprehensively as they relate to enterprises of a smaller dimension.

The main contribution of this article is to identify the effects of the introduction of the 'Società Benefit' model into Italian company law and examine the first empirical evidence from its application. Useful operational tools are drawn from it, especially for smaller companies, which, inspired by this business model, can develop their own sustainability strategies by relying on an organisational model that highlights comprehensive communication and analysis of non-financial performance.

I. Introduction

It is widely known that Corporate Social Responsibility (CSR) has reached a prominent position in the current debate on corporate law and corporate governance. In a recent book, two distinguished company law scholars declare their ambition to 'establish sustainability-related study of corporate law and corporate governance as a field'[1] and they outline numerous initiatives aimed at

* Associate Professor of Commercial Law, University of Macerata.

[1] B. Sjåfjell and M. Bruner, 'Corporations and sustainability', in B. Sjåfjell and M. Bruner eds, *Cambridge Handbook of Corporate Law, Corporate Governance and Sustainability* (Cambridge: Cambridge University Press, 2019). The application of a strong conception of sustainability places companies' action, as regulated by company law and corporate governance rules, above the minimum target of 'social foundation' and within the limits of 'planetary boundaries'. It is an original application to the corporate law and corporate governance field of well-known concepts

reconsidering the future of the corporations that are developing on both sides of the Atlantic.

More recently, and with particular emphasis, the very purpose of the corporation has been questioned. On the one hand, the 'Future of the Corporation Program' promoted by the British Academy suggests reformulating the concept of corporate purpose in a sense that is 'not solely about profit, but about public purposes that relate to the firm's wider contribution to public interests and societal goals'. On the other hand, from the heart of the American economic system, comes the exhortation to redefine the purpose of the corporation, 'to promote an economy that serves all Americans'.[2]

The CSR mentality is growing fast in managerial theory. Its focus has always included institutional arguments in its toolbox that analyse business organisations, and this has increasingly affected the legal discourse around CSR approaches.

Within this framework, European and European Union (EU) Member States' legislation is increasingly characterised by aims of ensuring a sustainable footprint, especially by promoting the mandatory disclosure of non-financial information by companies and encouraging voluntary models for purposeful businesses. At a national level, the French loi 'Plan d'Action pour la Croissance et la Transformation des Entreprises' (loi PACTE, 22 May 2019 no 486) provided the opportunity to companies to introduce a *raison d'être* into company bylaws. This is characterised as a purpose which encompasses principles at the very basis of a company's mission and values for which the company intends to allocate resources to carry out its activity. This reform aspires to comprehensively redesign the mission of enterprises. It has its origin in the drive to place the undertaking of a business within a framework of a responsible economy and to construct a third way between public action and the market economy, aiming at conciliating financial objectives of companies

developed by scientists belonging to other fields: see J. Rockström et al, 'Planetary Boundaries: Exploring the Safe Operating Space for Humanity' 14(2,32) *Ecology and Society (2009)*; W. Steffen et al, 'Planetary Boundaries: Guiding Human Development on a Changing Planet', 347(6223) *Science*, 736-747 (2015) and K. Raworth, *Doughnut Economics, Seven Ways to Think Like a 21st Century Economist* (White River Junction-Vermont: Chelsea Green Publishing, 2017). This approach makes it possible to review the famous triple-bottom-line scheme, discussed by the author himself still from a three-dimensional perspective based on the pillars of environmental, social and economic sustainability: J. Elkington, '25 Years Ago I Coined the Phrase "Triple Bottom Line". Here's Why It's Time to Rethink It' *Harvard Business Review*, 25 June 2018.

[2] Business Roundtable, 'Statement on the Purpose of a Corporation' of 19 August 2019, available at https://tinyurl.com/364h7j8t (last visited 30 June 2021). The aim of the signatories would not be so transparent: see M.J. Roe, 'Why Are America's CEOs Talking About Stakeholder Capitalism' *Project Syndicate*, 4 November 2019, available at https://tinyurl.com/ubuk2b47 (last visited 30 June 2021). For an update of American CEOs' commitment to the benefit of all of their stakeholders, especially during pandemic and racial crises which dramatically affected the United States in 2020, see Business Roundtable, 'One Year Later: Purpose of a Corporation', available at https://tinyurl.com/2jza3drh (last visited 30 June 2021).

with social and environmental goals.[3] The French loi PACTE has been heavily criticised in France including the questioning of its real innovativeness.[4] However, ethics and compliance are the current passwords in the implementation of corporate governance practices in the French ecosystem of industrial companies.[5]

The European legal framework is accelerating its transformation towards a sustainable approach underpinning the operations of businesses and is also intervening at the board level to enhance corporate sustainability performance. It is not audacious to suppose that the exit of the United Kingdom (UK) from the EU and consequently from the EU decision-making table has been playing a role in the promotion of a new legislative initiative for sustainable corporate governance. This new model is based on a controversial Ernst & Young (EY) report produced on behalf of the European Commission.[6]

The report starts from the assumption – not rigorously proven[7] – that many listed companies pursue shareholder value creation in a manner that is incompatible with long-term strategies of the company and the pursuit of environmental and societal goals.[8]

The shareholder primacy myth[9] seems destined to give way to a new approach, of stakeholderism moving from the traditional role of presupposition of legitimacy in business theory to a central concept in corporate governance regulation.

[3] For further information on the origin of this reform, see N. Notat and J. Senard, 'L'entreprise, objet d'intérêt collectif', Rapport du 9 mars 2018, available at https://tinyurl.com/ruyrbb4t (last visited 30 June 2021).

[4] P. Conac, 'Le nouvel article 1833 du Code civil français et l'intégration de l'intérêt social et de la responsabilité sociale d'entreprise: constat ou revolution?' *Rivista ODC*, 497, 500 (2019).

[5] J.C. Magendie, 'Ethique et conformité dans les entreprises' *Revue des sociétés*, 730 (2019); M.A. Frison-Roche, *Pour une Europe de la compliance* (Paris: Dalloz, 2019); J. Ballet et al, *L'entreprise et l'éthique* (Paris: Seuil, 2011).

[6] European Commission-EY, 'Study on Directors' Duties and Sustainable Corporate Governance', Final report (2020), available at https://tinyurl.com/3x43uppp (last visited 30 June 2021).

[7] The question of the methodology adopted in the study was discussed in depth during the consultation and has been indicated as a weakness of the initiative. Scholars also are aware that the Commission's objective to focus on long-term value creation and improvement of resiliency of European undertakings in current market and social terrible development deserves maximum support: see for example A. Bassen et al, 'University of Hamburg Feedback Statement' (2020), available at https://tinyurl.com/92vue6vp (last visited 30 June 2021).

[8] The report has been criticised for being based on scant significant empirical data, both in terms of the number and variety of categories considered: see, amongst other commentary, the feedback of the Confederation of Finnish Industries, of 6 October 2020, available at https://tinyurl.com/4wnu2s22 (last visited 30 June 2021) where it is pointed out that the web survey on which the EY report was based was limited to sixty-two stakeholders with only twelve people being interviewed in twelve countries. This weakness in the report's quantitative analysis could make the empirical basis of the study unreliable for EU-wide application.

[9] J.R. Macey, 'The Central Role of Myth in Corporate Law', ECGI Law Working Paper no 519/2020, available at https://tinyurl.com/jejdbnsw (last visited 30 June 2021).

In recent months, however, the COVID-19 crisis, starting in China at the end of 2019 and spreading all over the world during 2020 and 2021, offers a new significant input that resonates powerfully with climate-emergency concerns, as declared existed by governments and scientists in December 2016. The declaration aimed at reconstructing a business model that was more aware of environmental responsibilities and more resilient to global and systemic crises.

In this complex and fast evolving framework, minor companies, belonging to the category of so-called small and medium sized enterprises (SMEs), are mostly outside the center of the debate because CSR issues are seen to mainly concern large enterprises, operating at a multinational level and particularly those listed on regulated markets.[10]

Nevertheless, SMEs represent ninety nine per cent of European businesses[11] and account for more than ninety per cent of the world's business enterprises.[12] It is already clear that the engagement of SMEs in responsible business conduct is crucial to world economies, considering that these enterprises help create employment opportunities, drive economic growth and a more equitable distribution of income in society.[13] Business organisation theory, in fact, has already developed copious literature about CSR and SMEs, which is why we can also expect future development of the legal debate on this issue.

The first formal definition of the concept of CSR was made in the seminal work of Bowen,[14] who defined it as 'the set of moral and personal obligations that the employer must follow, considering the exercise of policies, decisions or courses of action in terms of objectives and values desired by society'. Subsequently, it has assumed ever more importance in a debate with multiple topics: from the ideas of Drucker, who underlines the need to take public opinion into account in organisational decision making processes, regardless of the size of a corporation or an industry,[15] to the contributions by Davis, where he discusses the role played by the trust of stakeholders for business success and strength, giving way to a theoretical trend known as corporate constitutionalism.[16] The company, in fact, bases its success on the responsible exercise of power, taking into account the interests of its stakeholders. If it does not live up to this

[10] European Commission-EY, n 6 above, 1.

[11] Compare data provided by European Commission on the website https://tinyurl.com/6ex5x57e (last visited 30 June 2021).

[12] See, for example, United Nations, 'Supporting Small Business through Covid-19 Crisis' (2020), available at https://tinyurl.com/4px3tdr7 (last visited 30 June 2021), when the United Nations refers to the data provided by the International Council for Small Business.

[13] W. Luetkenhorst, 'Corporate Social Responsibility and the Development Agenda' 39 *Intereconomics*, 157, 158 (2004).

[14] H. Bowen, *Social Responsibilities of the Businessman* (New York: Harper & Brothers, 1953), 6.

[15] P. Drucker, *The Practice of Management* (New York: Harper & Row, 1954).

[16] K. Davis, 'Can Business Afford to Ignore Social Responsibilities?', 2(3) *California Management Review*, 70-76 (1960); see also S. Bottomley, *The Constitutional Corporation: Rethinking Corporate Governance* (Aldershot: Ashgate Publishing, 2007).

goal, it is doomed to failure and expulsion from the market.

Although the debate is now quite old and has had growing relevance in the late twentieth and early twenty-first centuries,[17] – so much so that it has become a real paradigm of economic development for major international organisations – the literature available on CSR of SMEs is rather scarce.[18] Therefore, once the notions of SMEs and CSR have been defined, taking into account the voluntary pattern that constantly remains at the basis of the social and environmental commitment of the for-profit enterprise, it is certainly possible to highlight the main regulatory data in Italian company law in order to build on this hybrid form of enterprise to arrive at a prospective statute to promote socially responsible SMEs.[19]

II. Defining Small and Medium-Sized Enterprises

The definition of SMEs does not have only one meaning. While constantly based on quantitative parameters, what is an SME varies in different geographical areas, in different types of industry and depending on the organisational forms of businesses.[20]

Even in the academic field, the concept of an SME is quite varied. However, it is customary to distinguish SMEs from larger companies by the presence of qualitative characteristics, such as businesses controlling a small market share, being subject to the direct management of their owner and lacking bureaucratised organisational structures.[21]

In Europe, the notion of a SME is contained in the EU Recommendation 2003/36, and is fundamentally based on quantitative data, such as numbers of employees, turnover and balance sheet total, although other factors must also be considered for a correct application of the rules intended for them, especially

[17] OECD, 'Better Policies for 2030: An OECD Action Plan on the Sustainable Development Goals' (Paris, Meeting of the OECD Council at Ministerial Level Paris, 1-2 June 2016), available at https://tinyurl.com/m7pykc9y (last visited 30 June 2021); United Nations, 'Transforming Our World: The 2030 Agenda for Sustainable Development' (New York, United Nations General Assembly September 2015) available at https://tinyurl.com/yhjeth7h (last visited 30 June 2021).

[18] M. Libertini, 'Economia sociale di mercato e responsabilità sociale dell'impresa' *Rivista ODC*, 1, 8 (2013); A. Kechiche and R. Soparnot, 'CSR Within SMEs: Literature Review' 5(7) *International Business Research*, 97-104 (2012); R. Vázquez-Carrasco and M.E. López-Pérez, 'Small&Medium-Sized Enterprises and Corporate Social Responsibility: a Systematic Review of the Literature' 47 *Quality & Quantity*, 3205-3218 (2013).

[19] See, in relation to the Italian system, F. Massa ed, *Sostenibilità. Profili giuridici, economici e manageriali delle PMI italiane* (Torino: Giappichelli, 2019).

[20] G. Berisha and J.S. Pula, 'Defining Small and Medium Enterprises: a critical review' 1(1) *Academic Journal of Business, Administration, Law and Social Sciences*, 17-28 (2015).

[21] L. Spence and J.F. Lozano, 'Communicating About Ethics with Small Firms: Experiences from the UK and Spain' 27 *Journal of Business Ethics*, 43 (2000); G. Enderle, 'Global Competition and Corporate Responsibilities of Small and Medium-Sized Enterprises' 13(1) *Business Ethics: A European Review*, 50, 51 (2004).

incentives. A company, in fact, while remaining below the quantitative threshold indicated by the recommendation, could have access to significant additional resources because it is owned by, linked to, or partnered with, a larger enterprise.[22] Therefore, together with quantitative requirements, other aspects relating to ownership, partnership and linkages must be considered when categorizing a SME, so as to ensure that it is a genuine SME.

If we then look at the qualitative data, such as is done in Italy under Art 2083 of the Civil Code, the characteristics of a smaller company mainly revolve around direct and personal management which operate in an informal way and are based on interpersonal relationships, focusing on direct communication with stakeholders and the dedication of particular consideration towards employees, community and consumers.[23]

Although it has always been investigated from the perspective of large companies, especially multinationals, CSR naturally belongs to the sphere of SMEs. SMEs' agile organisational form, indeed, facilitates the transmission of ethical values from the owner of the SME to stakeholders and the community in which the business is located, looking for their endorsement and support. The owner's perspective is quite particular: they tend to impart a cooperative spirit to the management of the business and impresses an ethical corporate culture on the entrepreneurial organisation, where profit is not the only goal, as achieving results of creating shared value within the community is an equally important purpose.

III. CSR: Too Vague a Notion?

In addition to the points made above, it is worth noting that the extremely popular notion of CSR has been developed as a rather vague concept, to indicate the impact that running company activities has on society and on the environment. From time to time, a wide range of conduct, combined with the purpose of mere profit, has been linked to the concept of social responsibility: including engaging in a philanthropic action, applying stewardship principles and pursuing social goals.

In this extreme interpretation, the voluntary nature of social responsibility represents a fixed point, possessing a double meaning. Firstly, the non-mandatory essence of CSR is inferred from the fact that the pursuit of social goals entails the undertaking of significant costs or risks for the company, which

[22] 'For enterprises with a more complex structure, a case-by case analysis may therefore be required to ensure that only those enterprises that fall within the 'spirit' of the SME Recommendation are considered SMEs', European Commission, 'User Guide to the SME Definition', available at https://tinyurl.com/3kdjj9y7 (last visited 30 June 2021).

[23] J. Lepontre and A. Heene, 'Investigating the Impact of Firm Size on Small Business Social Responsibility: A Critical Review' 67(3) *Journal of Business Ethics*, 257-273 (2006).

could influence its market success or failure.[24] Second, as explicitly clarified in the Green Paper adopted by the European Commission on CSR,[25] being socially responsible does not mean only ensuring compliance with legal and statutory obligations, but involves the assumption of altruistic values and commitment to the community in which a company operates.

The evolution of the CSR concept has passed through, and has been characterised by, various ways of conceiving the relationship between business and society. The well-known opinion of Milton Friedman, which is considered as starting the myth of shareholder maximisation, acknowledged that

> there is one and only one social responsibility of business – to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game.[26]

Since the nineteen-eighties, in economic theory social responsibility has increasingly assumed the existence of a rationality criterion in business management,[27] taking up a vision of a commitment to efficiency according to social, environmental and ethical concerns, thus enhancing behaviors based on social and environmental sustainability as a source of business opportunities, innovation and competitive advantages.[28]

Moreover, the notion of CSR remains a broad concept even in the most recent definition adopted by the EU, which encompasses multiple values at the very basis of responsible conduct: it remains, under several aspects, a precise synonym of business ethics.

The European Commission defines CSR as 'the responsibility of enterprises for their impacts on society' and it also clarifies that CSR is something different from mere compliance with laws and that regulations underlining 'respect for applicable legislation, and for collective agreements between social partners, is a *prerequisite* for meeting that responsibility' (emphasis added).[29]

In terms of language, we can observe that we need a stipulative definition of

[24] C.C. Walton, *Corporate Social Responsibilities* (Belmont: Wadsworth Publishing, 1967).

[25] European Commission, Green Paper 'Promoting a European Framework for Corporate Social Responsibility' [COM(2001) 366 def] of 18 July 2001, available at https://tinyurl.com/f9s5fhdw (last visited 30 June 2021).

[26] M. Friedman, 'The Social Responsibility of Business is to Increase its Profits' *The New York Times Magazine*, 13 September 1970, available at https://tinyurl.com/yw6akp4v (last visited 30 June 2021); Id, *Capitalism and Freedom* (Chicago: University of Chicago Press, 1962).

[27] A. McWilliams and D.S. Siegel, 'Corporate Social Responsibility: Strategic Implications' 43 *Journal of Management Studies*, 1-18 (2006).

[28] P.F. Drucker, 'Social Impacts and Social Problems', in Id, *The Essential Drucker 2001* (Oxford: Routledge, 2001); M.E. Porter and M.R. Kramer, 'Strategy & Society. The Link Between Competitive Advantage and Corporate Social Responsibility' 84(12) *Harvard Business Review*, 78-92 (2006).

[29] European Commission, Communication 'A Renewed EU Strategy 2011-2014 for Corporate Social Responsibility' [COM(2011) 681 final] of 25 October 2011, available at https://tinyurl.com/2s9rktzt (last visited 30 June 2021).

CSR that is able to facilitate the modeling of organisational structures (and culture) and societal values aiming to promote the wider adoption of responsible behavior in running businesses.

In this sense, the way in which we can imbue a general clause mandating CSR with meaningful significance is through a strategic approach. Strategic CSR is a well-known orientation which starts from an obvious affirmation: only if CSR investments are also good for the business itself, can they work as a driver of innovation, economic growth and societal prosperity. This vision is shared by a large number of authors who underline how strategic CSR tends to align to the well-known objective of the creation of value in the long term, which is also dear to European and national legislators.[30] When business leaders are aware that a proactive attitude towards shareholders is able to generate gains for the business itself and at the same time achieve social benefits, they are inclined to set a strong affirmative CSR agenda in doing business.[31]

This approach seems quite compatible with Friedman's theory, referred to above, according to which the responsibility of a corporate executive is to conduct the business in accordance with the desires of the owners, 'which *generally* will be to make as much money as possible while conforming to their basic rules of society, both those embodied in law and those embodied in ethical custom' (emphasis added). It is worth noting that the definition of CSR inferable from Friedman's position goes beyond the respect for legal rules and ethical custom. It consists of the behavior of corporate executives who run the company for purposes differing from those of its shareholders' interests, voluntarily adhering to socially desirable conduct which is neglected by the law and the ethical norms.[32]

It is evident that large-scale success of a CSR approach cannot be detached from acceptance and trust in the competitive value of ethical business practices

[30] Compare, for example, European Commission, Communication 'Action Plan on Financing Sustainable Growth' [COM(2018) 97 final] of 8 March 2018, available at https://tinyurl.com/nd8hw3fy (last visited 30 June 2021): 'this is necessary, if the EU is to develop more sustainable economic growth, ensure the stability of the financial system, and foster more transparency and long-termism in the economy'. On the uncertainty that arises from the use of the long-term concept, as a panacea for sustainability concerns, see M. Stella Richter jr, 'Long-Termism' *Rivista delle società*, 16-52 (2021).

[31] Arguments form the strategic approach are widely shared in the economic literature: see, among other references, R.E. Freeman, *Strategic Management: A Stakeholder Approach* (Cambridge: Cambridge University Press, 2010); R.E. Freeman et al, *Stakeholder theory: The State of the Art* (Cambridge: Cambridge University Press, 2010); E. Garriga and D. Melé, 'Corporate Social Responsibility Theories: Mapping the Territory' 53 *Journal of Business Ethics*, 51-71 (2004); S. Wheeler, *Corporations and the Third Way* (Oxford and Portland-Oregon: Hart Publishing, 2002).

[32] This would be the real field of Corporate Social Responsibility: for this definition of voluntariness, see C. Angelici, 'Divagazioni sulla "responsabilità sociale" d'impresa' *Rivista delle società*, 3, 7 (2018).

and consideration of stakeholders' needs,[33] especially for the purpose of legitimacy and increasing the reputation of the firm in the market and in the community.

If organisational integration of CSR acts or activities in the business has a proven capacity to increase company value and its profits in the long term, then expenditures on strategic CSR activities become long-term investments that are likely to yield financial returns.[34]

This point of conceptual equilibrium, which enhances social responsibility as a productivity factor and a tool for creating value, fits in coherently with other actions of European authorities, related to the encouragement of long-term shareholder engagement[35] and the strong promotion of sustainable investing.[36]

IV. CSR in Italian Company Law, from Large Enterprises…

Despite its cultural and theoretical appeal, the implementation of CSR goals has had a lukewarm welcome in the Italian discourse on corporate law.

In Italy, the contrast between 'institutionalism' and 'contractualism' in the theory of the firm has always been strong. A contractualist approach has been prevalent in the discourse on company law since the postwar period. This largely unhindered preference, which implies the rejection of any legal construction that references institutionalist theories of the firm, is not only based on the strong contractualist culture mentioned above, but also on a profound distrust of approaches, like institutional ones, which seem to have an ideological link with fascist corporatism.[37] This has meant that there has been no room in Italian company law for a vision of CSR that was not merely a voluntary vision. Not surprisingly, the latest organic reform of company law, approved in 2003, did not consider the issue of CSR and there is still no trace of

[33] P. Ruggiero and S. Cupertino, 'CSR Strategic Approach, Financial Resources and Corporate Social Performance: The Mediating Effect of Innovation' 10(10) *Sustainability*, 3611 (2018).

[34] G.P. Lantos, 'The Boundaries of Strategic Corporate Social Responsibility' 18(7) *Journal of Consumer Marketing*, 595-632 (2001).

[35] European Parliament and Council Directive (EU) 2017/828 of 17 May 2017 amending Directive 2007/36/EC on the encouragement of long-term shareholder engagement [2017] OJ L132/1.

[36] See European Commission, n 30 above, and, just before the pandemic crisis hit Europe, Sustainable Europe Investment Plan (SEIP), the investment pillar of the Green Deal: European Commission, 'Proposal for a Regulation of the European Parliament and of the Council establishing the Just Transition Fund', of 14 January 2020, [COM (2020) 22], and European Parliament and Council Regulation (EU) 2020/852 of 18 June 2020 on the establishment of a framework to facilitate sustainable investment and amending Regulation (EU) 2019/2088 available at https://tinyurl.com/nzysta2s (last visited 30 June 2021). The framework of European authorities' interventions has been shaped by the European Commission Communication 'The European Green Deal' [COM (2019) 640 final] of 11 December 2019, available at https://tinyurl.com/yyah6jpk (last visited 30 June 2021).

[37] Full explanation in M. Libertini, n 18 above, 11.

general provisions in Italian law such as Art 1833 of the French Civil Code, according to which '*la société est gerée dans son intérêt social, en prenant en consideration les enjeux sociaux et environnementaux de son activité*' or section 172 UK Companies Act which imposes on directors the duty of an enlightened decision making process, having regard to a series of factors listed in the section, which refer to the promotion of environmental, social and governance objectives.[38]

The view can be taken that the recognition of CSR principles is implicit in the provision of Art 41, para 2, of the Italian Constitutional Charter which establishes that private economic initiative 'cannot take place contrary to social utility or in such a way as to damage security, freedom or human dignity'. However, this provision, according to a widely shared interpretation of it, does not mean that the company must necessarily pursue social ends or assume sustainability as a central objective in its strategy and operation. It only makes it clear that the freedom of economic initiative must not be exercised in conflict with fundamental human rights, defining this not as a fundamental right itself, but only as a regulated freedom, which is limited by a series of rules and principles established by the legislator.[39] Therefore, it is clear that Italian company law has not shown a particular concern for CSR issues, whose regulation is substantially left to the mandatory rules provided for by laws dedicated to environmental and social protection. Nevertheless, in recent years some disciplines have appeared, drawing inspiration from the international models in this area, that aim to boost more sustainable action by Italian enterprises. This comes also in the wake of European measures.

First of all, we have to consider the law defining the 'Business Act' (legge 11 November 2011 no 180), which was adopted to ensure the full application of the European Commission's Communication 'Think Small First – A Small Business Act for Europe'.[40] The purpose of the law is to promote national and regional legislation consistent with the scope of the Small Business Act; and in so doing, it identifies some fundamental principles that should ensure the further strengthening of the sustainable growth and competitiveness in SMEs.

The Italian Business Act aims, incidentally, to 'promote the inclusion of

[38] This is the wording of section 172, of the Companies Act 2006 (UK), subsection 1: 'A director of a company must act in the way he considers, in good faith, would be most likely to promote the success of the company for the benefit of its members as a whole, and in doing so have regard (amongst other matters) to: a) the likely consequences of any decision in the, b) the interests of the company's employees, c) the need to foster the company's business relationships with suppliers, customers and others, d) the impact of the company's operations on the community and the environment, e) the desirability of the company maintaining a reputation of high standards of business conduct, and f) the need to act fairly as between members of the company'.

[39] M. Libertini, n 18 above, 20; Id, 'A "Highly Competition Social Market Economy" as a Founding Element of the European Economic Constitution' *Concorrenza e mercato*, 491 (2011).

[40] [COM(2008) 394 final] of 25 June 2008, available at https://tinyurl.com/4na3j6re (last visited 30 June 2021).

social issues and environmental matters in the conduct of business activities and in their relations with stakeholders'. Even if the law has the characteristics of a declamatory discourse, rather than a strict prescriptive formulation, it represents proactive support for sustainability as it introduces social aims among the general principles of the business legal system, which may be relevant in the interpretation of more specific legal rules and should improve collaboration between business and public authorities.

A second, very strong boost coming from EU law, is connected with the implementation of Directive 2014/95/EU, on non-financial and diversity information.[41] As a sign of increasing CSR juridification, the Directive introduces mandatory communication, including a description of the policies pursued in relation to environmental social governance (ESG) matters and due diligence processes implemented by the company and its supply chain. Although it does not impose legal obligations of conduct on undertakings, it offers, by a comply or explain mechanism, a strong reputational incentive to adopt ESG strategies and practices.[42]

The Directive concerns large undertakings which are public-interest entities (as defined by Directive 2013/34/EU) and public-interest entities which are parent undertakings of a large group.[43] However, it is possible that even SMEs can provide a non-financial statement containing information about environmental, social and employee matters, because Italy has applied the optional provision underlined by recital 14 of the Directive, opening the regulation to the discretionary adoption by small undertakings.

The recital shows, on the one hand, the legislator's conscientiousness in not imposing a disproportionate burden on minor enterprises in terms of the cost of reporting sustainability information;[44] and on the other hand it allows

[41] European Parliament and Council Directive 2014/95/EU of 22 October 2014, amending Directive 2013/34/EU on the disclosure of non-financial and diversity information by certain large undertakings and groups [2014] OJ L330/1. The regulation is actually under review: see The European Commission Proposal for a Directive of the European Parliament and of the Council amending Directive 2013/34/EU, Directive 2004/109/EC, Directive 2006/43/EC and Regulation (EU) no 537/2014, as regards corporate sustainability reporting of 21 April 2021 [COM(2021) 189 final] available at https://tinyurl.com/2uaxtusw (last visited 30 June 2021).

[42] Note that the Sustainable Finance Action Plan expanded the non-financial reporting requirement to include disclosure of initiatives to reduce the impact of climate change: Communication of the Commission, Guidelines on non-financial reporting. Supplement on reporting climate-related information [2019] OJ C209/1.

The recent European Commission Proposal for a Directive on corporate sustainability reporting (n 41 above) confirms the choice not to put new reporting requirements on small enterprises, except for SMEs with securities listed on regulated markets. The burden to listed SMEs is also limited, as they will be allowed to report their sustainability information using simpler standards than the standards that will apply to larger undertakings.

[43] 'In each case having an average number of employees in excess of five hundred, in the case of a group on a consolidated basis', Directive 2014/95/EU, Whereas 14.

[44] Sustainable paths of SMEs would in any case be assured to the extent that larger companies are obliged to disclose information on the due diligence processes also regarding its supply and

enterprises which are outside the perimeter of the mandatory disclosure requirements to publish a non-financial report on a voluntary basis, with the declaration of compliance with the decreto legislativo 30 December 2016 no 254, thus being able to demonstrate to stakeholders an active engagement in ESG strategies and objectives.

V. … To Small-Sized Companies: The Italian Benefit Corporation

It was noted above that there are rules, such as those on the communication of non-financial statements, that SMEs can follow on a purely voluntary basis in order to communicate their social commitment to stakeholders affected by the business activities and to the public as a whole. Recently, however, organisational models have been enriched in Italian company law with a qualification to a company's purpose, namely that of 'Società Benefit'. This concept aims to encourage companies to assume the obligation of creating or pursuing a general and one or more specific public benefits, assessed against a third party standard, in addition to the purpose of profit.[45]

The Italian 'Società Benefit' was inspired by the model of the North-American Benefit Corporation, first introduced into the Maryland legislation in April 2010, on the basis of a 'Model Business Corporation Act' proposed (and promoted) by B-Lab, a not-for-profit organisation that certifies as B-Corp for-profit companies which meet rigorous standards of social and environmental performance, accountability and transparency.[46] The intent, in the archetypical legal experience, was to build a safe harbor for directors, should they wish to take into consideration, as interests along which they run the company, concerns of other important constituencies, such as employees, customers, local or regional economy, local or global environment and other factors.[47]

subcontracting chains (Whereas 6, Directive 2014/95/EU). Recently, the European Parliament Resolution of 10 March 2021 with recommendations to the Commission on corporate due diligence and corporate accountability, available at https://tinyurl.com/ykju68se (last visited 30 June 2021) following a large study of European Parliamentary Research Service released in October 2020, available at https://tinyurl.com/2xrcknw7 (last visited 30 June 2021) proposes the adoption of a hard law instrument aiming at strengthening corporate accountability for human rights and environment abuses while pursuing the objective of boosting good governance at the European and international level.

[45] Italy implemented the 'Società Benefit' legislation with legge 28 December 2015 no 208.

[46] For more information about the activity of B-Lab, which leads a community including three thousand nine hundred BCorp in seventy countries, providing for BCorp Certification and a B Impact Assessment which is widely used to measure and manage social and environmental performance of businesses, see https://tinyurl.com/38tv43wu (last visited 30 June 2021). Actually, 37 States have passed benefit corporation legislation and four are working on it.

[47] Compare W.H. Clark et al, 'The Need and Rationale for the Benefit Corporation', version of 18 January 2013, available at https://tinyurl.com/y84an3km (last visited 30 June 2021); H.K.

The model introduced by the Italian legislator fits into the European legal context with an economic model that is historically more institutional and intermediate than the Anglo-American one. Here, the provision of Benefit Corporation law responds to the need to enable shareholders to optimally pursue a balance between a purpose of profit and the achievement of social ends,[48] particularly by integrating stakeholders' needs into organisational activity and performance plans.[49] It is therefore an organisational model that pertains to the issue of CSR to the extent that it entrusts private autonomy, through the inclusion in the bylaws of a general and one or more specific public benefits, to strike a balance among societal interests and the profit purpose, which remains the typical objective of the company.

It is not easy to say whether this approach contradicts the voluntary nature of CSR, because there is the possibility that, according to the bylaws of a company, the public benefit pursued by a company could be delineated in a very generic form and therefore, in many cases, the pursuit of the general public benefit (acting responsibly, sustainably and transparently) and of the specific one (serving one or more specific purposes of common benefit) will be left to the exercise of administrative discretion by managers of the company.

But what is the relationship between SMEs and benefit corporations?

In Italy, most of the benefit companies are SMEs, as defined in the Commission Recommendation of 6 May 2003 no 361.[50] Therefore, it is to this size range of companies, rather than to larger public companies, that the benefit model seems to apply.[51] As with other nations that have introduced a similar

Lidstone et al, 'The Long and Winding Road to Public Benefit Corporations in Colorado'(2019), available at https://tinyurl.com/asjn99ek (last visited 30 June 2021).

[48] The meaning of 'social enterprise' concept is quite different between Europe and United States, R. Esposito, 'The Social Enterprise Revolution in Corporate Law: A Primer on Emerging Corporate Entities in Europe and the United States and the Case of Benefit Corporation' 4(2) *William & Mary Business Law Review*, 639-714 (2013); R. Katz and A. Page, 'The Role of Social Enterprise' 35 *Vermont Law Review*, 59-103 (2010); B. Means and J.W. Yockey eds, *The Cambridge Handbook of Social Enterprise Law* (Cambridge: Cambridge University Press, 2019). It is worth mentioning that Italian Company Law provides for another figure, the 'Impresa Sociale' regulated by decreto legislativo 3 July 2017 no 112, in which the purpose of profits must not to be the main objective of members, which should rather pursue the social mission underlined by Art 2, decreto legislativo no 112/2017. The substantial difference between 'Impresa sociale' and 'Società Benefit' under Italian Company Law, lies precisely in the fact that the functional impact of the public benefit purpose must not be higher than that of the profit purpose, G. Marasà, *Imprese sociali, altri enti del terzo settore, società benefit* (Torino: Giappichelli, 2019), 23-24.

[49] G. Riolfo, 'The New Italian Benefit Corporation' 21(2) *European Business Organization Law Review*, 279-317 (2020).

[50] [2003] OJ L124/36. See, on this topic, M. Bianchini and C. Sertoli, 'Una ricerca Assonime sulle società benefit. Dati empirici, prassi statutaria e prospettive' *Analisi giuridica dell'economia*, 201, 206 (2018) and, explaining a case study in the context of Italian SMEs, M. Del Baldo, 'Acting As a Benefit Corporation and a B Corp to Responsibly Pursue Private and Public Benefits. The Case of Paradisi Srl (Italy)' 4(1) *International Journal of Corporate Social Responsibility*, 1-18 (2019).

[51] M. Stella Richter jr, 'Corporate social responsibility, social enterprise, benefit corporation: magia delle parole?' *Vita notarile*, 953-968 (2017).

change in corporate law, it has been welcomed in Italy as a powerful tool that should stimulate SMEs to develop sustainable strategies and to enhance those already in place by aiming at boosting the transition of the entrepreneurial system towards sustainable development.[52] Even though Italian SMEs have a low intensity of CSR conduct, they have a great willingness to move toward sustainability practices, and to better formalise unstructured CSR processes already in place, in the presence of adequate incentives.[53]

By adopting the 'Società Benefit' model, companies that are mobilizable towards CSR aims could progress to more advanced cohesive organisational forms, where the areas of engagement and approach to CSR communication are systematic and creative, expressing themselves in a large variety of shapes. This can include the involvement of employees in decisions, a high level of transparency in decision-making processes, taking actions relating to sustainable manufacturing and extensive collaboration with local community and not-for-profit organisations.[54]

Dual mission management and common benefit communication are at the centre of the law relating to Italian Benefit Corporations. First of all, the provision of one or more specific benefits aims at fulfilling the objective of responsible, sustainable and transparent management, making the pursuit of the blended mission binding for shareholders and managers, who have the delicate task of balancing potentially antithetical interests.

Directors, therefore, have far greater discretionary powers, conforming in a similar way to that contemplated in the management of (profit) corporations according to team production theory,[55] where choices that reduce profits in favor of aims of stakeholders other than shareholders can only be prevented by the majority of shareholders threatening to revoke or not to reconfirm the appointment of directors.[56]

The introduction of the Benefit Corporation model into Italian company law has awakened an age-old debate on the purpose of the corporation, raising the question of whether the new legislation was intended to influence the main interpretation of Art 2247 Civil Code. This provision, in fact, provides that, in the exercise of an economic activity, companies must have an egoistic,[57] as well

[52] E. Giovannini, 'Prefazione', in F. Massa ed, *Sostenibilità* n 19 above, XIII-XIV.

[53] M.M. Molteni and M. Lucchini, *I modelli di responsabilità sociale nelle imprese italiane* (Roma: Franco Angeli, 2004), 121.

[54] ibid. For an explanation of the cohesive model of conduct, A.Y. Mermod and S.O. Idowu, *Corporate Social Responsibility in the Global Business World* (Berlin-Heidelberg: Springer, 2014), 177.

[55] M.M. Blair and L.A. Stout, 'A Team Production Theory of Corporate Law' 85(2) *Virginia Law Review*, 247, 248 (1999).

[56] F. Denozza and A. Stabilini, 'La società benefit nell'era dell'investor capitalism' *Rivista ODC*, 1, 14 (2017).

[57] Using the broader concept of egoistic purpose aiming to interpret the provision in Art 1, para. 376, legge no 208/2015, where it states 'in addition to the aim of split profits', we can easily

as a lucrative purpose. However, in the past it was never doubted, at least since the decline of the *ultra vires* acts theory, that companies were able to pursue public benefit purposes,[58] or move their activity towards environmental or social sustainability.[59]

Following the introduction of legge no 208/2015, there has been debate as to whether this interpretation is still plausible, or if the introduction of the 'Società Benefit' model has strengthened the lucrative (or egoistic) purpose of companies that do not have a public benefit mandate in their corporate charter. The question arises from the obscure formulation of Art 1, para 379, legge no 208/2015, where it is provided that 'companies others than BC [Benefit Corporations], if they also intend to pursue public benefit purposes, are required to amend articles of association or bylaws'. Hence the counterintuitive conclusion is that only with a change to their bylaws can a company orient its strategies and objectives towards sustainable performance. However, this conclusion seems to be contrary to the regulation on non-financial information, which allows firms, both at a European (recital 14, Directive 2014/95/UE) and national (Art 7, decreto legislativo no 254/2016) level to adopt a non-financial statement on a voluntary basis. It would therefore be illogical to think that the adoption by management of sustainability strategies was subject to the condition of amendment of bylaws at a shareholders' meeting.

It is important to underline that Italian law, unlike the American Model Business Corporation Act, requires the appointment of a benefit officer, which is typically a formal role useful for improving endo-managerial processes aimed at enhancing the hybrid and long-term orientation of the company, as is characteristic of cohesive enterprises. The main differences between the American Model Business Corporation legislation and the Italian model are set out in the following table.

	Model Business Corporation legislation	Italian 'Società Benefit' legislation
General public benefit (in the articles of incorporation)	Mandatory	Mandatory
One or more specific public benefit(s) (in the articles of incorporation)	Optional	Mandatory

resolve the question posed by the contradiction between § 376 – which seems to refer to companies with a lucrative purpose solely – and § 377, according to which the BC quality can be assumed by each of the types of companies mentioned in Book V, Title V and VI of the Italian Civil Code, including cooperatives. On the subject, G. Riolfo, n 49 above, 287-288, considers that the second provision is absorbent with respect to the first one.

[58] M. Stella Richter jr, n 51 above, 957, 961-962.

[59] G. Marasà, n 48 above, 18-20.

Third-party standard	Mandatory	Mandatory
Directors' standard of conduct	Board of Directors (BoD) must *consider* conflicting factors	BoD must *balance* pecuniary interests of shareholders and GPB-SPBs
Benefit director	Mandatory for listed companies; optional for the others. They 'shall be an individual who is independent' (§302 (b))	Not provided
Benefit officer	Optional	Mandatory
Benefit report	Mandatory	Mandatory
Benefit name	Not provided	Optional
Stakeholders' forum (or similar tools)	Not provided	Not provided

Until recently, the Italian 'Società Benefit' did not enjoy any fiscal benefit, tax relief or other financial advantages. Only with the approval of the recent legge 17 July 2020 no 77, which has introduced a new rule entitled 'Promotion of the Società Benefit system', can an organisation which adopts the Benefit Corporation model obtain a tax credit equal to fifty per cent of the costs of incorporation or transformation (*recte*: amendment of bylaws). It is therefore quite clear that reputational advantages serve as the main incentives for the adoption of the 'Società Benefit' form. This is why the communication of the social responsibility involved in the benefit model adoption remains crucial and it is mentioned in a series of rules relating to them. The denomination 'Società Benefit' can be used in the name of the company – although this is not mandatory – only if a public benefit purpose has been inserted into the articles of incorporation or in the bylaws of the company.

As pointed out above, a Benefit Corporation is obliged to draw up an annual report which constitutes its main accountability tool. It also represents the only useful means for stakeholders and the supervisory authority, which in Italy is the AGCM (Autorità Garante della Concorrenza e del Mercato), to evaluate the pursuit of programmed not-for-profit benefits and the non-deceptive nature of related communications, within the framework of the regulation of unfair commercial practices.

The adoption of integrated reporting methods, which the non-financial disclosure practice is also moving towards,[60] represents a further point – albeit

[60] G. Nigri and M. Del Baldo, 'Sustainability Reporting and Performance Measurement Systems: How do Small- and Medium- Sized Benefit Corporation Manage Integration?' 10(12) *Sustainability*, 4499 (2018).

a voluntary one – in the strategic improvement of multi-purpose business models, since it manages to favor the integration of social responsibility into management functions, fostering greater cohesion of objectives and protection from the risk of opportunism.

VI. Conclusions

Even though in Italian Company Law there are no rules similar to those outlined above in Art 1833 of the French Code Civil or by section 172 of the UK Companies Act relating to CSR considerations, the issue of CSR has come to the centre of the corporate law discourse and interesting schemes are offered within the Italian sphere for the development of sustainable management strategies by SMEs.

The Italian legal system, thanks to the guiding force of constitutional values (like the 'social utility' of Art 41 of the Constitutional Charter) and the EU adhesion to the highly competitive social market economy model under Art 3 of European Union Treaty, already hold all the tools for developing the discourse on business purpose in a modern way. Moreover, EU harmonisation has introduced, and is still developing, information obligations and compliance duties that are likely to affect the behaviour of companies, especially in a context in which access to finance may depend on a more attentive awareness about sustainability. However, if the Italian legislator wants to extend issues concerning social responsibility beyond the traditional field of large multinational companies and involve SMEs, whose contributions are essential for achieving the ambitious goals of sustainability,[61] it needs to provide incentives and support for the adoption of cohesive organisational forms, rather than just develop more prescriptive legislation, which could overburden small businesses with excessive costs.

Virtuous entrepreneurial realities are already widespread in the Italian business environment. They are characterised by the strong personal imprint of the owner of a business and by the transmission of family and personal values by the owner in the value chain. These firms are defined as 'spirited businesses'.[62] They arise from personal and family values as well as from the attachment to local communities and find ever greater legitimacy in the sensitivity of communities. This is also due to the direct communication and spontaneous convergence of businesses on issues of environmental and social emergency.

The provision of the Benefit Corporation model is not an enabling measure

[61] See, European Commission, Communication 'An SME Strategy for a Sustainable and Digital Europe' [COM(2020) 103 final], available at https://tinyurl.com/54rt3fuy (last visited 30 June 2021), 1: 'Europe's 25 million small and medium enterprises (SMEs) are the backbone of the EU economy'.

[62] M. Del Baldo, 'Corporate Social Responsibility and Corporate Governance in Italian SMEs: The Experience of Some "Spirited Businesses"' 16(1) *Journal of Management and Governance*, 1-36 (2012).

– as it is clear that even profit companies can carry out single acts and activities with a public benefit purpose. However, companies can perform a promotional role in providing CSR organisational tools such as the appointment of a benefit officer, complying with an obligation to file a yearly report on its non-financial performance using an independent third-party standard, the enlargement of the discretionary management powers to allow for the consideration of non-financial stakeholders and the balancing of conflicting interests. This can all lead to blending social impact with competitive advantage. If we intend to take CSR seriously,[63] the role of the 'Società Benefit' model may also be used to enhance a progressive approach towards improving sustainability strategies in SMEs and to lead to a widespread adoption, also in the corporate field, of real and not just fictitious forms of CSR.

[63] According to the valuable suggestion of M. Libertini, 'La comunicazione pubblicitaria e l'azione delle imprese per il miglioramento ambientale' *Giurisprudenza commerciale*, I, 331, 334 (2012). The risk that the public declaration of a benefit purpose by the corporation becomes the premise of opportunistic behaviors could be just around the corner, if one considers that (general and special) benefit purposes are defined by most statutes in a very vague form and the equilibrate achievement of societal and lucrative aims are not fully guaranteed in practice: G. Mion and C.R. Loza Adaui, 'Understanding the Purpose of Benefit Corporations: An Empirical Study on the Italian Case' 5(4) *International Journal of Corporate Social Responsibility*, 1, 12 (2020).

THE ITALIAN LAW JOURNAL

An International Forum for the Critique of Italian Law

Vol. 07 – No. 01 (2021)

Short Symposium

'The Terrible Duty. Legitimacy and Usefulness of Punishment'

Short Symposium on the Punishment

Introduction

Mario De Caro* and Francesco Toto**

The discussion on the justification and purpose of punishment is as old as philosophy but, like all genuine philosophical questions, no agreement has been reached in this regard – not even in the last decades, in which evolutionary theory, neuroscience, and cognitive psychology have been offering contributions in favour of all the various views at stake in the discussion. In this regard, it is surprising that, while how punishment is inflicted has significantly evolved over time, the categories used to justify it and explain its functions are very similar to those forged by past thinkers. In this light, our aim here is to offer some state-of-the-art reconstructions of the views on punishment held by some of the most prominent philosophers of justice, from Aristotle up to the contemporary times.

The first article is by Flavia Farina, who focuses on the Aristotelian concept of punishment. Although Aristotle did not discuss this topic in a systematic way, he offered some reflections on it that generated many discussions. By analyzing those reflections, Farina highlights the philosophical depth of the problem, which draws on some key concepts in Aristotle's practical philosophy, such as the distinction between deliberate and non-deliberate actions and that between voluntary and involuntary actions. A multifunctional account of punishment thus comes into view: punishment has multiple aspects and cannot be reduced to any one of them. These are the corrective aspect (aiming to restore equality within the community and to attain reparative justice), the repressive or deterrent aspect (aiming to instil fear in order to discourage socially unacceptable practices), and the educational aspect (aiming to form or reform the convict's conduct and dispositions). Aristotle's most interesting reflections concern the last of these aspects, especially in light of his view that acquired character is unchangeable. According to many interpreters, this view implies that if punishment may discourage morally bad actions, it cannot encourage virtuous ones. Against such a

* Full Professor of Moral Philosophy, Roma Tre University; Visiting Professor, Tufts University.

** Assistant Professor of History of Philosophy, Roma Tre University.

reading, Farina argues that, in addition to having an undisputed instrumental value for those who distribute the punishment (ie the whole community), the coercion connected with punishment also has a potential ethical value according to Aristotle, one that aims at the good of those who suffer it.

Luc Foisneau's article is dedicated to Hobbes's conception of punishment as it was developed between the Leviathan and the Dialogue between a Philosopher and a Student of the Common Laws of England. Foisneau begins by discussing the justification of a sovereign authority's right to punish its subjects. In particular, the primacy of self-preservation, on which Hobbes insists, may seem to prevent citizens from authorising the sovereign to inflict punishment on them or on their neighbours. However, Hobbes finds a solution to this problem by arguing that in the covenant the right to punish is left to the sovereign, so that it coincides with the *jus in omnia* characteristic of the state of nature as a state of war. On this basis, Giorgio Agamben has maintained that for Hobbes the concept of sovereignty is tied to the normalisation of a 'state of exception', in the sense that the sovereign right to punish, seen as a right to war, constitutes the heart of sovereignty and strips the subjects bare of all their rights. Against such an interpretation, Foisneau claims that the Hobbes's justification of the right to punish – which is compatible with his denial of free will – lies in its effectiveness as a deterrent, that is, in its ability to promote the subjects' obedience, ie, to promote the conformity of their will and conduct to the laws promulgated by the sovereign in order to safeguard social security. At the same time, Foisneau highlights Hobbes's insistence on the somewhat problematic distinction between punishment and acts of hostility: in addition to aiming at the production of obedience, punishment must be established via law by a sovereign authority and imposed by a judge in accordance with that law. From this perspective, which safeguards legal guarantees and aims at the efficiency of the legal system as a whole, Foisneau discusses a case that seems to confirm the reduction of punishment to war and the close relationship between sovereignty and the state of exception. This is the 'statues of Provisors', a case in which the criminal is deprived of any legal protection not as citizen, but as an enemy and is thus subject to a foreign sovereign power.

Francesco Toto's article focuses on a well-known difficulty in Rousseau's theory of punishment. Rousseau seems to ground the right to punish in the right of war (a view Hobbes also defended, with the difficulties we have just seen). Yet the idea that criminals are punished, possibly with the death penalty, not as citizens but as a 'public enemies' seems to contradict many aspects of Rousseau's political theory. On the one hand, the enemy is external to the state, and is punished without regard to considerations of utility or to the prospect of rehabilitation and reintegration. On the other hand, the application of the law to criminals should be carried out also in their own interest as a way of 'forcing them to be free' – of ensuring conformity their will to the 'General will', which is also their

own – in order to reintegrate them into the community. Toto indicates a way out of this problem by interpreting Rousseau's reference to the 'public enemy' in a restricted way. In accordance with the subject discussed in the Social Contract – that is, political law – the criminal punished as an enemy must be a political criminal, ie, an agent who undermines the foundations of democratic legitimacy (the most important case being that of individuals who, vested with some state authority, try to obtain despotic power by evading the control of the sovereign people).

Dario Ippolito's article investigates some Enlightenment approaches to the deontology of punishment by both comparing the views of Montesquieu, Beccaria, Filangieri, Genovesi, and Kant, and analysing some of the most influential readings of 18th century criminal reformism such as those of Michel Foucault and Giovanni Tarello. From a historical-critical point of view, Ippolito shows how Foucault cannot reduce the heterogeneity of the *philosophes'* positions to a unitary economic rationality and their humanitarianism to the ideological mask of discipline required by this rationality, if not at the price of distorting their specificities and mixing together incompatible theoretical commitments. A discussion of Tarello's proposed distinction between three main 'schools' or 'ideologies' (humanitarianist, utilitarianist, and proportionalist or retributivist) highlights, instead, an inverse distortion. In the 18th century, utilitarianist and retributivist principles could not only coexist in the same authors, but each of them could lead to opposing positions belonging to different ideologies (for example, utilitarian justifications could be offered both for and against the death penalty). From a more strictly conceptual point of view, Ippolito dwells on two meta-legal principles, the 'principle of proportionality' (which prescribes a quantitative correlation between the gravity of the penalty and that of the crime) and the 'principle of homogeneity' (which prescribes a qualitative correlation between the type of crime and the type of penalty). He shows that the first principle, introduced with a view to mitigating penalties, is based on utilitarian considerations: proportionality has a deterrent function, and is premised on a view of the correlation between crime and punishment as a political artifice, an institution, a human responsibility. By contrast, the second principle considers the correlation between crime and punishment in the naturalistic and retributivist terms of a 'natural correspondence'.

Francesca Fantasia discusses Kant's philosophy of punishment, focusing on Kant's views on legitimacy on punishment, on the conditions under which an action may be punishable, on the modalities of punishment, and more generally on the relation between pure practical reason and pragmatic reason. With regard to the first point, Fantasia shows how for Kant the legitimacy of punishment is analytically contained in the idea of law: there is no right without authorisation to coercion, and therefore without punishment of violation. On the second point, Fantasia argues that for Kant an action can only be punished if it is carried out

voluntarily against public law. This retributivist conception of punishment is doubly linked to the theme of humanity, the characteristic of humans as ends in themselves. First, humanity dictates that punishment is not imposed in a merely instrumental way, with a view to something other than retribution for a crime committed. Second, humanity imposes proof in trial as a requirement for punishment, specifically proof that an autonomous subject inflicted harm on another equally autonomous subject in a way that breaks a law. Thus, either the lack of a law or trial or a violation of the principle of equality (both between the parties involved and between them and the judge), is sufficient to invalidate punishment. When it comes to the modalities of punishment, Fantasia discusses the well-known assimilation of the necessity of punishment to a categorical imperative. The proper necessity of punishing a crime committed by a sane agent and proven in trial is unconditional, independent of any utilitarian considerations. The *lex talionis* is the only criterion, independent of empirical considerations and consistent with the principle of equality between parties, by which punishment can be made commensurate with the crime. Yet because it recognises the lexical priority of the principle of humanity with respect to the principle of equality, the *lex talionis* also imposes internal limits on commensurability. Not even the most heinous of crimes can authorise a punishment that violates the dignity of the criminal. After discussing Kant's rebuttal of Beccaria's criticism of the death penalty – and showing that it is not so much a defence of the death penalty as a defence of the coercive and anti-utilitarian character of law – Fantasia ends by clarifying Kant's conception of punishment as a categorical imperative, which issues from pure practical reason. This conception does not aim to rule out all justifications of the instrumental value of the punishment, as considered by pragmatic reason, but rather to subordinate the latter to the former, and therefore to attribute to the 'vindictive' value of punishment a primacy over all other kinds of value, such as deterrent or re-educational value.

Sabina Tortorella examines Hegel's theory of punishment as it is developed in the Elements of the philosophy of right, paying particular attention to the difference in perspective between the section on 'abstract law' and that on 'ethics'. She underscores both the specificities of these different perspectives – ie the way in which one aims to delineate a purely rational foundation of punishment and the other to make explicit its historically determined aims and conditions of applicability – and their convergence within a unitary framework. Within this framework, the first perspective is at once the logical presupposition and historical result of the second, since the internal logic of law emerges in its truth only with the formation and stabilisation of the modern state. Indeed, for Hegel, the state is the only entity entitled to publicly establish laws governing the relations between private individuals and punishments for their violation. Tortorella also shows how, on both perspectives, punishment necessarily fulfils a function of universalisation of particular wills and realisation of the idea of law. In this

sense, punishment cannot be reduced, along retributivist lines, to the simple compensation of wrong, since it cannot mend the tear between reality and justice caused by the wrong without at the same time contributing to the *Bildung* of the criminal as an ethical subject, to repairing the laceration between his particular will and universal will, and thus to his consequent reintegration within the community. In the same way, punishment cannot simply be considered a violence that the state must inflict on the criminal in fulfilment of a duty towards the injured party, because it is always, at the same time, a right of the criminal himself, and in this sense less a limitation of his freedom than its realization.

Marco Piazza's work examines the theories of punishment of La Mettrie and Nietzsche, as illustrated in *Discourse on happiness*, on the one hand, and in *Human, all Too Human* and *On the Genealogy of Morality*, on the other. His aim is twofold. From a historical point of view, Piazza tests the hypothesis that Nietzsche's conception of punishment develops in a critical dialogue with the French philosopher and physician. Piazza concludes that Nietzsche certainly had indirect knowledge of La Mettrie – through Albert Lange's *History of Materialism*, at the very least – and it is plausible (though not provable) that he also had direct access to La Mettrie's texts, or to their translations. From a theoretical point of view, Piazza frames the two philosophers' conceptions of punishment in the broader context of their respective metaphysics, anthropologies and moral philosophies, highlighting both convergences and divergences. Among the views that Nietzsche and La Mettrie share are some general theses that constitute the background of the concept of punishment: determinism, the arbitrariness of the ideas of good and evil, the birth of these ideas within a project of domination, and the unsustainable price that they require the individual to pay in terms of the sacrifice of his or her natural drives. More specifically, and more immediately connected to the theme of punishment, the two share the ideas of a merely socio-political and non-philosophical foundation of justice and criminal law, of the merely social function of punishment, and of the uselessness of remorse. Where Nietzsche and La Mettrie diverge, on the other hand, they reveal differences in their respective conceptions of materialism, with Nietzsche further radicalizing La Mettrie's views. These differences concern the understanding of the fundamental motive of human action (the search for pleasure in La Mettrie, the search for power in Nietzsche); the evaluation of this motive (the selfishness which, for La Mettrie, is not morally reprehensible but rightly punishable inasmuch as socially harmful, and which Nietzsche proposes to free from the repression to which it is subjected by prejudice and morality); and the evaluation of the legal-moral sphere itself (which La Mettrie considers philosophically unfounded but practically or socially necessary, and which Nietzsche on the contrary hopes to leave behind).

Finally, Patrizio Gonnella's article proposes a critique of the reeducation paradigm of punishment that combines a plurality of themes and sources, from

memorialism to philosophical reflection, from jurisprudence to anthropology. Gonnella draws attention to the ambivalence of this paradigm and the easy abuses to which it lends itself. He calls into question the centrality that the Italian penal system continues to accord to prison, noting that the opening to alternative penal measures has not succeeded in slowing the growth of the prison population, which is no longer linked to the growth (or decrease) rates of crime. At the same time, he underscores the incompatibility of prison reality with the rhetoric of re-education: following Massimo Paravini, a multiplicity of functions (punitive, programmatic, expressive, strategic) can be attributed to prison, but in reality prison remains today, as it was during the Fascist period, a place of suffering, humiliation and exclusion, governed by asymmetrical and at best paternalistic relations of power. These relationships are based on unwritten and often arbitrary rules, removed from any democratic control and largely extra-legal if not explicitly illegal. Today prison remains a factory of recidivism. Faced with the risk that the reeducation paradigm is the mask of revenge and arbitrariness – when not also of the particular interests that in the United States have, for example, transformed the prison into a real business – Gonnella proposes to bring the value of human dignity – ethical but constitutionally recognised – back to the forefront. The primacy of human dignity, the author argues, can limit the arbitrariness of the power to punish and help avoid the correctionalist degeneration of rieducational 'treatment'.[1]

[1] The papers collected here derive from a seminar held at Roma Tre University in January 2019 as a part of the project '*Dinamiche pubbliche della paura e cittadinanza inclusiva*' financed by the same university in the framework of the call '*Azione 4: azione sperimentale di finanziamento a progetti di ricerca innovativi e di natura interdisciplinare*'.

Short Symposium on the Punishment

The Therapeutic Function of Punishment in Aristotle

Flavia Farina*

Abstract

In the *Nicomachean Ethics* Aristotle describes punishment as a sort of cure. However, a well-defined and complex theory of punishment is nowhere to be found in Aristotle's works: all mentions of punishment occur in works significantly different in focus and the argumentative contexts also vary. Despite these difficulties, as Aristotle states that punishment is a cure, the possibility to ascribe to Aristotle a reformative theory of punishment will be taken into account. The aim of this paper is thus twofold: on one side, I will argue that while a theory of punishment is indeed to be found, punishment itself is not to be reduced to one simple function. I will further argue that, while Aristotle is skeptical about the possibility of changing one's character, the possibility of a reformative theory of punishment is consistent with his claims about the 'almost' impossibility of moral reform.

I. Introduction

In the *Nicomachean Ethics* (NE) Aristotle describes punishment as a kind of cure,[1] which – like all cures – works through opposites. Just as virtues – Aristotle argues – are concerned with actions and passions, and as passions are accompanied by pleasure and pain, punishment, in the case of character, is effective by means of pains and pleasures. Nonetheless, Aristotle neither develops nor offers a well-defined and coherent account of punishment. A theory of punishment is nowhere to be found in Aristotle's works and an account of the functions punishment can be said to accomplish is missing. Unlike Plato, who explicitly addresses punishment in book IX of the *Laws*, Aristotle is not committing himself to a complex theory of punishment. We can ask, then, whether it is possible at all to ascribe a theory of punishment to Aristotle. Besides these initial points, another difficulty must be considered. All mentions of punishment appear in Aristotle's practical works, namely the *Rhetoric, Politics and Ethics*, but these texts are significantly different in focus; hence, the argumentative contexts in which punishment is mentioned also vary considerably.

The aim of this paper, then, is twofold. Firstly, despite the lack of a complex account of punishment, I will consider whether it is still possible to reconstruct

* Research Fellow, RomaTre University.

[1] *NE* II 3, 1104b16-18.

a theory of punishment: on the basis of several passages, the different functions ascribed to punishment will be taken into account. Therefore, in the first section, an overview of the different functions of punishment in Aristotle's works will be provided. I will argue that while a theory of punishment is indeed to be found, punishment itself is not to be reduced to one simple function. I will further argue that Aristotle approaches the subject from an ethical rather than merely instrumental perspective.

Secondly, as several *NE* passages state that character is difficult to change, the question arises as to whether punishment can rightfully be said to have a therapeutic function, as an effective instrument to bring about a change in the agent's habits. Consequently, the consistency between the idea of punishment as a cure and Aristotle's ethical views on character will be evaluated.

II. Theories of Punishment

Generally speaking, punishment is the deliberate infliction of pain or loss on an individual by the state or a community. At the same time, punishment is the instrument through which laws exercise their coercive power.

When Aristotle speaks of punishment in the *Politics* (Pol. IV 14, 1298a5), he is referring specifically to practices such as exiling, the death sentence, property confiscation, fines and other penalties. All of the aforementioned practices are concrete means of punishment, but they can serve different purposes.

In ancient Greece and Athens, orators and philosophers developed three fundamental theories of punishment: the corrective, the deterrent, and the reformative. They attempted either to justify these theories or to develop new accounts and alternatives to their opponents' arguments.[2] None of the aforementioned theory can be said to lie outside Aristotle's consideration. Nevertheless, the absence of another theory of punishment is remarkable: the cleansing theory.

Tragic examples of the cleansing function of punishment can be found in Oedipus and in the Erinyes and Orestes episode. There, punishment is seen as a form of purification of the soul of the wrongdoer. Still, the presence of the Prytaneion, a tribunal in charge of judging inanimate objects and animals, provides specific evidence of the purifying function of punishment. Objects and animals found guilty of homicide were cast outside the city borders in order to prevent them from spreading diseases. Athenians considered all unnatural death to be a

[2] An overview of the different theories of punishment in Classical Athens can be found in two papers by D. Cohen: D. Cohen, 'Crime, Punishment, and the Rule of Law in Classical Athens', in M. Gagarin and D. Cohen eds, *The Cambridge Companion to Ancient Greek Law* (New York: Cambridge University Press, 2005), 211-235; D. Cohen, 'Theories of Punishment', in M. Gagarin and D. Cohen eds, *The Cambridge Companion to Ancient Greek Law* (New York: Cambridge University Press, 2005), 170-190.

matter of extreme gravity[3] and had a mainly religious view of homicide. Vengeance, on the one hand, and cleansing, on the other, were not optional. As death was a pollution, the polluter, whether it be a person or an object, was to be banished to preventing the spread of the disease. Aristotle mentions this practice only once, in Pol. 1262a32. Referring to Plato's *Republic*, Aristotle foresees the impossibility of applying punishment as cleansing in a state where family connections are disregarded. Punishment as cleansing was primarily (albeit not exclusively) concerned with familicide. Aristotle's objection goes as follows: if, in Plato's *Republic*, parents are unknown to their children and vice-versa, it will be impossible to know the identity of the family member killed and of the killer and, consequently, their relation, and thus to proceed with the cleansing. Since punishment as cleansing is merely used as a dialectical means to criticize Plato's views on the family and the state, Aristotle can be said to be largely unconcerned with the practice itself.[4]

As already stated, it is not easy to find a complete and well-defined theory of punishment in Aristotle. Punishment seems to serve different functions in different argumentative contexts. Aristotle's skepticism about the possibility of changing people's character only seems to complicate matters. If it is true that Aristotle is skeptical about the possibility of moral reform, we might conclude that the reformative theory is nothing more than an opinion shared by some of his contemporaries, with no further implications for Aristotle's ethical philosophy. It would be nothing but *endoxa*, a view which Aristotle himself does not share. We will return to this point in the next section of the paper. The corrective and the deterrent theories (ie restoring the broken balance and deterring people from committing a crime), however, seem to be more coherent with Aristotle's views as expressed in the *Nicomachean Ethics*.

In *NE* V 4, 1132a6-19, Aristotle states that the judge has the duty to restore balance and to re-establish equality in a community. Indeed, whenever an injustice between two parties is committed, one has more than its fair share and the other less:

> Therefore, this kind of injustice being an inequality, the judge tries to equalize it; for in the case also in which one has received and the other has inflicted a wound, or one has slain and the other been slain, the suffering and the action have been unequally distributed; but the judge tries to equalize things by means of the penalty, (10) taking away from the gain of the assailant (NE V 4, 1132a6-10).

The primary function of corrective justice is thus to restore a lost balance.

[3] W.T. Loomis, 'The Nature of Premeditation in Athenian Homicide Law' *Journal of Hellenic Studies*, 92, 95 (1972).

[4] R. Sorabji, *Necessity, Cause and Blame, Perspectives on Aristotle's Theory* (Ithaca-NY: Cornell University Press, 1980), 289.

In the presence of a crime, whether in commercial transactions or in other matters, one of the parties has taken more than its fair share and more than it deserved: in this case, one party has experienced a gain and the other, conversely, a loss. Corrective justice is thus concerned with private transactions. They are to be distinguished into two kinds: voluntary and involuntary.[5] Voluntary transactions (such as selling, buying, etc) are mainly contracts between individuals, while involuntary transactions (theft, assault, murder) are liabilities for the payment of compensation to another citizen. In both cases, justice and punishment readdress the balance between individuals: in voluntary transactions, however, the balance is restored by redressing the breach of an agreement (eg by means of a fine); in involuntary transactions, penalties must be inflicted on the wrongdoer.

Aristotle acknowledges that speaking of gain and loss can be misleading and not always appropriate in relation to the crime committed. If we are dealing with measurable goods (as in voluntary transactions), it is easy to see if there is a disequilibrium that corrective justice has to restore by means of a fine. This is the case, for instance, with the breaching of a contract. But what happens in a murder case? It would seem rather odd to speak of gain and loss in this case or of a balance to be restored. Aristotle nonetheless states that, even in this case, it is still possible to speak of gain and loss in a derivative and analogical way.[6] Here too corrective justice can and should restore the balance and the proportion lost. The judge therefore restores and re-establishes the balance in all cases of injustice.

Aristotle is dealing here with a mathematical proportion. If one of the two parties, let it be called A, had to get four but got two, and the other party, let it be called B, consequently got six, the equilibrium will be re-established by means of a mathematical average. A lacks two and B exceeds by two. As a consequence, A shall have the two it is lacking and the judge will restore the balance, understood as the mean between an excess and a defect.

> By this, then, we shall recognize both what we must subtract from that which has more, and what we must add to that which has less; we must add to the latter that by which the intermediate exceeds it and subtract from the greatest that by which it exceeds the intermediate (NE V 4, 1132b2-5).

Corrective justice, however, is not equivalent to reciprocation. Punishment is not a revised kind of *lex talionis*, whereby someone who has deprived someone else of something has to suffer the same privation he or she has caused.

> Now reciprocity fits neither distributive nor rectificatory justice (…); eg if an official has inflicted a wound, he should not be wounded in return, and if someone has wounded an official, he ought not to be wounded only but punished in addition. Further, there is a great difference between a

[5] *NE* V 2, 1131a2-9.
[6] *NE* V 4, 1132a10-14.

voluntary and an involuntary act (NE V 5, 1132b24-30).

Punishment is not exclusively defined by reciprocity, as one may be led to think. So the proportionality principle should not take only the disproportion between the parties into account but also the agent's attitude. What makes a difference is whether the agent has committed an injustice voluntarily or involuntarily.[7]

Aristotle recognizes that punishments should be more or less severe depending on the agent's attitude at the moment of committing the crime. If an agent strikes someone and this results in bodily injuries, the agent who has struck the victim might be punished more or less harshly, depending on whether he or she could have foreseen the negative effects of his or her own actions. The judges, then, will consider whether the ultimate results of the wrongdoer's actions can be classified as an error, a misfortune or an act of injustice. Punishment will be more or less harshly settled, proportionally to the error committed.

> And it must make us distinguish between wrongdoings on the one hand, and mistakes, or misfortunes, on the other (Reth. I 13, 1374b5-7).

The distinction that Aristotle draws between errors, misfortunes and wrongdoings constitutes a significant contribution to Athenian legal theory. Although precedents of the aforementioned distinction can be found in Greek laws about the classification of wrongdoings (eg the distinction between murder and manslaughter), no distinction between premeditated homicide and voluntary homicide is to be found in Greek laws. Athenian laws distinguished between three cases of homicide: voluntary, involuntary, and justifiable.

The distinction also implies that the various cases will be settled by different courts. Generally, cases of deliberate homicide (*ek pronoias* or *hekousios*) were held at the Aeropagus, cases of involuntary homicide at the Palladion, and cases of justifiable homicide at the Delphinion. No distinction was drawn, therefore, between homicide *ek pronoias,* ie premeditated homicide, and voluntary homicide. In *Against Aristocrates* Demosthenes reports that according to law in case of legitimate defense, the person defending herself from an assaulter can kill the wrongdoer and escape punishment exclusively if the homicide is carried out without premeditation. Otherwise, the homicide will fall under the category of *hekousios* homicide. The absence of deliberation is thus a feature of legitimate defense but in no case is it considered to be an aggravating

[7] Evidently, the description of the agent's disposition as voluntary or involuntary can apply to both voluntary and involuntary transactions. The distinction between voluntary and involuntary transactions is therefore a difference established not on the basis of the agent's disposition but rather on the kind of transaction. Voluntary transactions are transactions where the two parties usually reach an agreement voluntarily while in involuntary transactions, such as theft, one of the two parties is involuntarily, ie unwillingly, deprived of something.

circumstance.[8]

Aristotle, on the contrary, primarily distinguishes between premeditated and non-premeditated actions (NE V 8, 1135b8-11); secondly, in relation to the latter category, he distinguishes between voluntary and involuntary actions.

In order to be just, punishment must take these differences into account, as they reflect different dispositions. An agent who commits a crime with premeditation acts from bad character states or, generally speaking, because she is evil or vile. Therefore, she will be punished more harshly than an agent who committed an error she could not have reasonably foreseen at the time of the action itself. Even more significantly, an agent who could have foreseen the negative effects of her actions (but failed to do so because of negligence) will be punished more harshly than one who could not have reasonably foreseen them.

Proportionality, then, is not merely settled *ex parte objecti*, meaning with regard to the victim or the crime itself and the misbalance the act has caused. Rather, it is also established *ex parte subjecti*, ie by considering the agent's disposition, which led to the occurrence of the crime.

Aristotle further addresses the issue of the deterrent function of punishment.

> Witness seems to be borne to this both by individuals in their private capacity and by legislators themselves; for these punish and take vengeance on those who do wicked acts (unless they have acted under compulsion or as a result of ignorance for which they are not themselves responsible), while they honour those who do noble acts, as though they meant to encourage the latter and deter the former (NE III 5, 1113b22-25).

Hence, punishment has the function of both repressing and deterring the person who has committed a crime, in order to discourage the repetition of the same crime by the same agent, once she has been punished. The fear of new punishments plays a crucial role in the accomplishment of that function.

However, deterrence does not only work in relation to the already convicted person. It also serves as a warning and a reminder to all citizens. The punishment inflicted upon a single agent can speak to all members of the community. The goal, then, is not only to discourage bad and unlawful behaviors in the individual but also to discourage all other members of the community from engaging in the same behaviors. The same holds for public honors, whose function is not just to lavish praise on the person who deserves them but to encourage other citizens to engage in honorable actions or behaviors.

> But it is surely not enough that when they are young they should get the right nurture and attention; since they must, even when they are grown

[8] A useful reconstruction of homicide law in Classical Athens is offered by A. Merker, *Le Principe de l'Action Humaine – selon Démosthène et Aristote – Hairesis - Prohairesis* (Paris: Les Belles Lettres, 2016), 331-342.

> up, practise and be habituated to them, we shall need laws for this as well, and generally speaking to cover the whole of life; for most people obey necessity rather than argument, and punishments rather than what is noble (NE X 9, 1179b35-1180a5).

People are therefore discouraged from committing unjust and unlawful actions by the fear of punishment.

Insofar as it provides rewards and punishments, then, the law also has a protreptic function, in addition to serving as a deterrent. Both functions can help create suitable social and political conditions in which an agent can be educated. It is not sufficient, in Aristotle's view, to have been taught what is noble, honorable and good as children. As virtue is a habit and is acquired through the constant repetition of acts of the same sort, it is necessary that the laws continue to show both adults and children which behaviors are to be avoided and which are to be pursued. This function of law finds its main expression specifically in punishment and in the coercive power embodied by law.

A counterexample might make the point clearer. Say a child has been educated well: she has pursued the noble, abstained from injustice and has grown into a virtuous agent. If this same agent found herself living in a society where laws encouraged exactly the opposite line of conduct, she would probably engage in behaviors and actions contrary to her own education. It might also be the case that the laws are, generally speaking, good – they promote what is noble and the kind of behavior the child has been educated to embrace – but do not have any coercive power, for they are ineffective in punishing the unjust and in offering rewards to the just. Aristotle seems rather skeptical about the possibility that an agent who is learning to be a good, virtuous person could preserve her habits in a state that exercises its coercive power ineffectively.

> However that may be, if (as we have said) the man who is to be good must be well trained and habituated, and go on to spend his time in worthy occupations and (15) neither willingly nor unwillingly do bad actions, and if this can be brought about if men live in accordance with a sort of intellect and right order, provided this has force (NE X 9, 1180a14-18).

Besides, the function of a legal order is also to direct educators and primary parental figures (ie the father in Classical Greek society) toward those things that the child ought to be taught. Generally speaking, Aristotle believes that family education is intrinsically related to the father's knowledge of laws. Family education is considered to be more effective in virtue of the blood tie between father and son. The laws, on their part, must direct citizens towards happiness and the common good, and help them achieve these goals. Fathers and educators, then, must take into consideration what the laws command, in order to offer their children the best possible upbringing.

> Now, since politics uses the rest of the sciences, and since, again, it legislates as to what we are to do and what we are to abstain from, the end of this science must include those of the others, so that this end must be the good for man (NE I 2, 1094b4-8).

Laws therefore play a central role with respect to virtue and they are shown to be a privileged instrument for educating citizens, especially because of their coercive force and the power they have to encourage or discourage certain behavior. As a consequence, they are able to forge the citizen body. Punishment, therefore, also acquires an educational function because of its capacity to inflict pain and give pleasure. Education is based on rewards and punishments, because virtue is about pleasure and pain. Virtue consists – among other things – in finding pleasure in good things and noble acts, and pain in bad things and evil actions.

> For moral excellence is concerned with pleasures and pains; it is on account of pleasure that we do bad things, and on account of pain that we abstain from noble ones. Hence we ought to have been brought up in a particular way from our very youth, as Plato says, so as both to delight in and to be pained by the things that we ought; for this is the right education. Again, if the excellences are concerned with actions and passions, and every passion and every action is accompanied by pleasure and pain, for this reason also excellence will be concerned with pleasures and pains (NE II 3, 1104b9-15).

As I have tried to show, punishment serves an educational function in relation not only to characters yet to be shaped but also to fully formed ones, as it continues to point people toward what is to be pursued and what is to be avoided. However, a remark is in order here. Moral education[9] in Aristotle is different from behavioral conditioning. While the latter may be carried out – and indeed often is carried out – for the benefit of the controller or conditioner, the former is carried out in the interests of the learner and for the sake of his or her happiness, *eudaimonia*, and virtue. As it has been seen in *NE* 1094b4-8, Aristotle thinks of politics and legislation as aiming at the good for man, ie his happiness, defined as the activity of a complete life in accordance with complete virtue.

As to the educational function of punishment, therefore, it must not be confused with a mere conditioning practice, as the link with the deterrent function might suggest. A father punishing his child for a bad action could be seen as deterring him or her from engaging in the same action in the future. In this respect, punishment could be understood in merely instrumental terms. However, this is not exactly what Aristotle has in mind. As has been stated, behavioral conditioning primarily serves the controller's interests: the conditioner may punish the conditioned subject in an effort to maintain the existing order. I am not suggesting

[9] J. Echeñique, *Aristotle's Ethics and Moral Responsibility* (Cambridge: Cambridge University Press, 2012), 37-40.

that Aristotle denies the usefulness of such practices, however this is only half of the story. While acknowledging that it is useful to encourage or discourage certain behaviors by means of punishments and honors, Aristotle also has in mind the idea of moral education as something that is pursued for the learner's sake, for his or her good. The father punishing a child, while discouraging him or her from engaging in a certain action, is also acting for the sake of the child's virtuous upbringing, ie for the sake of the learner's own virtue and happiness.

It is now possible to consider the question whether punishment can also be assigned the third function mentioned above, the reformative. As already noted, Aristotle's skepticism with regard to moral reform and character change is one of the main issues to be dealt with once we start looking for a reformative or therapeutical role of punishment. This will be the focus of the next section.

III. A Reformative Theory of Punishment?

1. Some Difficulties

The general overview just provided ensures a deeper understanding of punishment in Aristotle's works. It appears that Aristotle does not formulate a single theory of punishment, but rather assigns punishment different functions. However, the educational function has been shown to be closely related to virtue and happiness and to be irreducible to an instrumentalist view of punishment. Education has the happiness of the learner and his or her good, ie virtuous, upbringing as its goal. Hence, it is concerned with the moral attitudes of individuals.

Aristotle further argues that punishment plays a role in deterring people from repeatedly engaging in bad actions: a man who has already established and acquired bad habits will at least abstain from engaging in unjust actions not because he wishes to, but out of fear of punishment. This amounts to nothing more than a deterrent function: for the individual in question does not refrain from engaging in bad acts because of his virtuous states of character. While Aristotle does not neglect the deterring function of punishment, we have seen that, in the case of education, this function is embedded within a moral view of punishment.

What we need to ask, then, is whether punishment can also have a therapeutic function: what we should look for in this case is a genuine moral reform, a conversion – so to say – from vice to virtue.

If a commander remains on the battlefield simply because he fears he might otherwise be punished upon returning to his home city, he cannot be considered courageous or virtuous. He will go into battle not because of his courageous habits but out of fear of suffering. So, his actions are not guided by virtuous dispositions, as stated in the *Eudemian Ethics* (EE).

> Further, we praise and blame all men with regard to their choice rather

> than their acts (though activity is more desirable than excellence), because men may do bad acts under compulsion, but no one chooses them under compulsion. Further, it is only because it is not easy to see the nature of a man's choice that we are forced to judge of his character by his acts. The activity then is more desirable, but the choice is more praiseworthy (EE II 11, 1228a11-18).

In Aristotle's view, actions in themselves are not sufficient for determining whether an agent has acted out of virtue. There can be several motives that could lead to the same outcome or to the same action, as the example of a commander shows. One commander may go into battle out of virtue and another out of fear and this marks a difference between the two. Our cowardly commander, who goes into battle because of fear, has probably only chosen the lesser of two evils: punishment is a more fearful consequence than the possibility of dying on the battlefield. Although the deterrent function may be said to have played a role in maintaining the necessary balance, the commander has not really changed his habits. Roughly speaking, punishment has deterred him from leaving his post, not changed him into a courageous person.

The main difficulty when it comes to the possibility of moral reform is represented by Aristotle's pessimistic views about character change. As Aristotle seems to dismiss the idea of genuine moral reform, the definition of punishment as a cure has been dismissed by many commentators, such as R. Sorabji[10] and G. Di Muzio.[11] If a person's character cannot be changed – Sorabji[12] argues – the idea of punishment being a cure must be understood as an *endoxa*, an opinion shared by Aristotle's contemporaries and thus reported by the philosopher, without him actually sharing the view. Di Muzio gives a different account of why the idea of punishment as a cure must be dismissed. The possibility of character change can be upheld by means of persuasion[13] rather than punishment. However, those who are deemed to be incorrigible or incurable (*akolastoi* and *aniatoi*) cannot be persuaded. In this case, it is only exposure to virtue that can help change a man's character. While the *akolastoi* cannot be changed either by persuasion or by punishment, they could still theoretically change their character, as moral agents never lose the possibility to act against their already established states of character. Therefore, by becoming exposed to virtue, little by little they can engage in virtuous actions as virtuous men would, ie with the right disposition and emotions. Having virtuous friends who are ready to help if needed can be the beginning of a process of moral reform, as the bad moral agent will try to emulate them. So whereas in the case of corrigible

[10] R. Sorabji, Necessity, *Cause and Blame, Perspectives on Aristotle's Theory* (Ithaca-NY: Cornell University Press, 1980).

[11] G. Di Muzio, 'Aristotle on improving one's character' *Phronesis*, 45, 205-219 (2000).

[12] R. Sorabji, n 4 above.

[13] G. Di Muzio, n 11 above, 213.

moral agents the process of moral reform will have an external origin, for persuasion is external to the agent, in the case of the incorrigible man, who is not corrigible by external means (persuasion or punishment), the process of moral reform will have its origin within the agent, as it is still possible for the latter to act against her states of character. As Di Muzio argues,[14] high order desires such as self-loathing can lead the agent to desire to change what is bad within her. In this case, exposure to virtue, together with the realization that virtuous people are immune to such feelings, may trigger the process of character change. While giving very different reasons as to why punishment is not effective as a cure, both Sorabji and Di Muzio dismiss the idea of punishment as an efficacious instrument to bring about a change in an agent's character. According to the former, character is simply not changeable; according to the latter, only persuasion and exposure to virtue can trigger and achieve moral reform – even for those who are deemed *akolastoi*, incorrigibles.

On my alternative account, I will try to defend the notion that punishment can be seen as an effective instrument to bringing about a process of moral reform in Aristotle, while at the same time dismissing the idea that change can have an internal origin. Before going into the details of my proposal, it is necessary to carefully consider Aristotle's pessimistic view about moral reform.

Aristotle's skepticism about the possibility of character change is revealed by several passages in his ethical writings and suggests that whereas the idea of an educational role of punishment is perfectly consistent with the Aristotelian project, that of a reformative role of punishment should be approached more cautiously. In the case of an agent whose habits are still in the making, eg a child, laws, with the sanctions they impose, constitute a deterrent from engaging in unjust actions. As already noted, they also present themselves as an instrument to direct one toward the good, happiness and virtue. A child can thus develop his habits in conformity with the laws and, if the laws themselves are good, can be educated in such a way as to achieve virtue and happiness. A different situation emerges when we consider an already mature agent. Aristotle thinks that in this case character change and moral reform are rather difficult, if not impossible, to achieve. Once the agent has acquired his habits, he will act (and feel) in accordance to them.

> In that case it was then open to him not to be ill, but not now, when he has thrown away his chance, just as when you have let a stone go it is too late to recover it; but yet it was in your power to throw it, since the moving principle was in you. So, too, to the unjust and to the self-indulgent man it was open at the beginning not to become men of this kind, and so they are such voluntarily; but now that they have become so it is not possible for them not to be so (NE III 5, 1114a16-21).

[14] ibid 216.

This passage states rather clearly that once an agent has acquired his habits, he cannot cease to be who he has become. A child has the possibility to become a virtuous or bad agent. In this case, as we have seen, the laws contribute to his upbringing by plainly stating which behaviors are to be avoided and which are to be pursued. Indeed, habits are acquired by an agent through the constant repetition of actions of the same sort. An agent will become courageous by acting courageously not just once but constantly and consistently. Conversely, an agent will become cowardly by acting in a cowardly fashion constantly and repetitively. A virtuous agent will find pleasure in acting virtuously and pain in acting poorly. Punishment, inflicting pain, is an educational means, as it helps find pain in bad actions.

But once an agent has already acquired his habits, he will act in accordance with them. A coward will act in a cowardly manner and a brave man will act courageously, just as the sick man will walk as a sick man does on account of his sickness. Once certain habits have been acquired, the possibility of not being that kind of man is no longer open to the agent. The unjust agent had the possibility of not being unjust, when she was still shaping her habits. However, once she has become unjust, it seems she cannot recover from her bad habits, just as it is impossible to take back a stone once it has been thrown. Apparently, the acquisition of character states is considered to be an irreversible process. The parallel is once more with sickness: the sick man had the initial possibility of avoiding sickness but once he has become sick, he cannot cease to be so. The possibility that was admitted at the beginning of the process of habituation, namely to become virtuous or vicious, is no longer available for the agent.

However, Aristotle also argues that this possibility is no longer available to an agent simply through his desire to change.

> Yet it does not follow that if he wishes he will cease to be unjust and will be just. For neither does the man who is ill become well on those terms – although he may, perhaps, be ill voluntarily, through living incontinently and disobeying his doctors (NE III 5, 1114a13-16).

It is specifically the clause 'if he wishes' that can cast light on the subject. Changing our character is difficult – ceasing to be the persons we are is difficult, almost impossible. However, Aristotle says that the man who has become unjust cannot cease to be unjust only by wishing it.

Let's once again consider the case of the sick man. Say this man is affected by a disease he brought upon himself by living unrestrainedly (eg a chronic disease due to an unhealthy lifestyle). The idea envisaged in the passage seems to be the following: this agent has made himself sick and now he cannot just go back to being healthy simply because he wants to. Indeed, he could not possibly recover without following the doctor's advice and changing his lifestyle.

Nonetheless, there are two main difficulties to be considered: the first problem

concerns the desire for change and the second the time of change. Aristotle has stated that what an agent wants or desires, the aim of his actions, does not depend on his reasoning but on his habits. The end, the general end, is not chosen but desired. Desire is, on its part, determined by the agent's habits. So, in order to change his character, the agent would need to desire to be a different person than he is, but this is exactly the possibility that seems to be denied once the desire is assumed to be determined by habits. Habits are constant expressions of desire. It is difficult, therefore, to understand how a desire for complete moral reform could arise in the agent. Although Di Muzio argues that we must distinguish desires from higher order desires, such as self-loathing, there is no passage where Aristotle suggests that an agent may regret his whole moral life. When Aristotle mentions feelings such as regret and remorse, it is always in relation to a single occurrence or to a single action. As the general aim, happiness, is variously interpreted and specified according to different dispositions (an agent, Aristotle argues, may have wealth as a general aim or pleasure, instead of *eudaimonia*),[15] it is difficult to see how a moral agent who already has a certain disposition may desire to radically change his life, his disposition and, most importantly, his ultimate goal or general aim. An agent might, at most, regret a single action, but he never questions his (moral) life as a whole.

Secondly, for an agent the process of habituating himself to be different from how he is requires time and perseverance. In order to change his habits, the agent must acquire new habits, acting constantly and repeatedly according to the new disposition or character states he wants to acquire. But Aristotle has already established that habits give rise to actions of the same quality: in other words, an agent will act in accordance with his own habits. How, then, is it possible for a coward to become courageous if he can act only in accordance with cowardice, ie in cowardly fashion? How can an agent act courageously in a constant way and for an extensive period of time in order to change his habits? The agent seems stuck in his own habits: since out of cowardice he will act cowardly, and in order to become courageous he needs to act courageously, he appears to be caught in a vicious circle. It seems difficult, therefore, to uphold a reformative theory of punishment: possibly, this view should be set aside. However, we can still ask ourselves why the only definition of punishment provided by Aristotle, defines punishment as a kind of cure (*Ret*. 1369b12; *EE* 1214b32; 1220a35; *NE* 1104b16). We are left with an apparent contradiction: on the one hand, character change is so difficult as to be impossible; on the other, punishment is said to be a cure. It seems as though one of the two alternatives must to be dismissed.

One option is to consider the mention of punishment as a cure to be one theory among others, whose soundness and credibility, in Aristotle's view, should be questioned. This is R. Sorabji's perspective. Aristotle was not only familiar

[15] *NE* III 5, 1114a31-1114b3.

with the theories on punishment developed by his contemporaries, but also took them into account. Hence, the idea of punishment as a cure could be understood as an *endoxa*. However, it is also true that the notion of punishment being a cure is the only definition of punishment to be found in Aristotle's practical works. We should ask ourselves, then, if there is another way of understanding the definition without dismissing it.

To avoid having to dismiss the second alternative, it would still be possible to try and reassess the first, ie the absolute impossibility of character change. By looking back again at *NE* III, a different interpretation can be provided in order to admit the view that punishment can also serve a reformative function. This is the aim of section four of the paper.

2. Punishment as a Cure: An Alternative Account

As we have seen, Aristotle is rather skeptical about the possibility of character change. However, nowhere does he argue for the absolute impossibility of the process. On the contrary, he states:

> It is hard, if not impossible, to remove by argument the traits that have long since been incorporated in the character; and perhaps we must be content if, when all the influences by which we are thought to become good are present, we get some tincture of excellence (NE X 9, 1179b16-20).

Aristotle acknowledges the difficulties implicit in changing an agent's character by means of discourses or arguments, once certain habits have been incorporated. So we must be satisfied if only a shade of virtue is generated.

First, attention is drawn to the description of the task of moral reform as something difficult and almost impossible. However, the fact that it is almost impossible does not imply absolute impossibility. Ultimately, Aristotle admits the possibility of generating a shade of virtue in the agent through all the influences and methods that can be used to achieve this good. In book I of the *Eudemian Ethics*, political correction, ie punishment, is once again compared to a medical treatment: 'political correction — for medicine, no less than whipping, is a correction' (EE I 3, 1214b32).

The possibility of moral reform has been shown to be difficult and almost impossible but its possibility cannot be excluded and dismissed *de iure*. In principle, albeit difficult and almost impossible, the possibility of character change must be admitted. The same situation seems to be discussed in the *Categories* (Cat.). When Aristotle speaks about the difference between states (*hexeis*) and conditions (*diatheseis*), he argues that states are harder to change than conditions. As virtue is a state – he explains – it is hard to change one's character. However, in chapter 10 he allows the possibility of moral reform, while at the same time recognizing all the difficulties involved in the process.

> For the bad man, if led into better ways of living and talking, would progress, if only a little, towards being better. And if he once made even a little progress it is clear that he might either change completely or make really great progress. For however slight the progress he made to begin with, he becomes ever more easily changed towards virtue, so that he is likely to make still more progress; and when this keeps happening it brings him over completely into the contrary state, provided time permits (Cat. 10, 13a24-32).

Besides, when Aristotle is talking about the impossibility for the agent to cease being the man he has become, he is talking about a single individual. In this case, not only does the unjust man want to be unjust but even if we were to admit that this man does not want to be unjust, he could still be unable to change his habits because the quality of his action is determined by his character states. He cannot engage in the process of acting courageously merely by wanting to be courageous. Not only that, but the arising or triggering of the process of moral reform is made difficult by his character states. Hence, it is difficult to see how this man will give rise to the desire for a radical life change and even if we admit the possibility of this desire arising, it would still be very hard for him to act contrary to his habits, for habits determine the quality of one's action. Simply put, the man in question is not used to acting courageously.

However, even if we dismiss the possibility of an internal origin of change (*contra* Di Muzio), this is not at all what is required when we define punishment as a cure. Indeed, punishment leads the individual back into the sphere of the community, for it brings another actor into play: the state, the person or the persons in charge of exercising authority by inflicting punishment, ie pain. The changing process is not all internal to the agent – Aristotle clearly excludes this possibility – but it is guided by an external agent. An autonomous and spontaneous process of moral reform from vice to virtue has to be excluded. Punishment, however, does not presuppose an autonomous process of reform but rather the opposite. Furthermore, the importance of a good upbringing and the role of an external influence in shaping one's character and emending some vices is frequently discussed by Aristotle. Moral reform cannot be accomplished suddenly, as it is matter of replacing old habits with new ones.

An agent, being influenced by others' opinions and judgements, by means of pleasure and pain, could start changing under their guidance. This guidance will use pleasure and pain, punishments and rewards, so as to re-educate the agent, since all cures are effective through opposites. A bad man who finds pleasure in vice shall be punished. And the pain of punishment will be maximally opposed to the pleasure the agent finds in his reprehensible behaviors. Moreover, Aristotle distinguishes between citizens to be banished from the city, and who are deemed incurable, and citizen who can be punished, who can be corrected.

One last remark is in order. The possibility of correcting one's character can be interpreted in two ways. One may deserve punishment if there is something

wrong in one's soul that can be corrected. However, correction can mean the sort of punishment merely understood as a socially useful instrument, as a way of conditioning and controlling one's action in a social community. The correction of an agent so as to make her meet social requirements does not directly promote her well being, her *eudaimonia*. As stated with regard to education, since educating someone can be taken to mean 'controlling' or 'conditioning' or 'educating for the good of the learner', the same distinction can be drawn in a reformative theory of punishment. Punishment as a cure, as an effective tool to bring about a change and re-educate the agent, can be seen as instrumental or as already – and always – embedded in an ethical prospective.

In the *Rethoric* Aristotle argues that while vengeance is for the sake of the victim (or the victim's family), punishment is for the sake of the wrongdoer:

> revenge and punishment are different things. Punishment is inflicted for the sake of the person punished; revenge for that of the punisher, to satisfy his feelings (Rhet. I 10, 1369b 12-15).

Punishment, it seems, is not just a way of making the behavior of bad men acceptable to the community: punishment as a cure is directed toward the good of the person punished. Thus interpreted, the theory of punishment as a cure cannot be reduced to an instrumental account, according to which the main aim of punishment is to condition and control deviant behaviors.

In conclusion, although character is described and presented as stable, it is not an inalienable possession. Despite the few references to a possible moral reform process, considered perhaps to be a rare possibility, character change has to be admitted or, at least, not excluded or completely dismissed. The starting point of this process must lie not inside the agent but outside, in the people who, by presenting themselves as teachers and guides, can finally lead the agent to change and replace old habits with new ones.

Short Symposium on the Punishment

Punishment Not War: Limits of a Paradigm

Luc Foisneau*

Abstract

The distinction between punishments and acts of hostility is central to Hobbes's theory of punishment in his three political treatises, but also in the 'Dialogue of the Common-Laws' and 'The Questions concerning Liberty, Necessity and Chance'. Such a distinction is not, as Agamben would have it, the expression of the equivalence between sovereignty and exception, but one dimension of a politics of sovereignty in its international context. The example of the statutes of Provisors as interpreted by Hobbes shows that the cruel punishment inflicted on those coming to England to receive ecclesiastical benefits given to them by the pope, the so-called 'Provisors', testifies mainly to the fierce struggles between the kingdom of England and the papacy. Hobbes invents a new theory of punishment no longer based, as in Suarez, on a metaphysics of free will but on the political consequences of punishing.

I. Introduction

The right to punish in Hobbes is known to be a contested matter: whereas *Leviathan* introduces a justification of the rights of public authority on the basis of a covenant of authorization by subjects, Hobbes says that the right to punish cannot be justified on such a basis. The reason given is that the authorization granted to a sovereign by a subject is so given to him in view of an individual good, which can never be found in a punishment, even though the latter be in conformity with penal laws known in advance. A citizen may authorize his sovereign to punish others – preferably not members of his close family and friends' circle – but not to punish himself. Therefore, the obligation to 'assist him that hath the Soveraignty, in the Punishing of another',[1] is not a sufficient justification of a right to punish, if such a right implies an obligation to be punished. No subject ever conceded to his sovereign such a liberty when that punishment constitutes a direct or indirect threat to his own life. Hobbes's stance did not vary on that point, which is already made, prior to *Leviathan*, in

* Director of Research at French National Center for Scientific Research (CNRS); Head of the Department of Political Studies at École des Hautes Études en Sciences Sociales; Member of Centre d'Études sociologiques et politiques Raymond Aron (CESPRA).

[1] Th. Hobbes, *Leviathan*, *2. The English and Latin Texts (i)*, edited by N. Malcolm (Oxford: Oxford University Press, 2012), XXVIII, 482 (hereafter: *Leviathan*, chapter in Roman capitals, and page).

De Cive.[2] He has it that the right to punish that sovereigns have 'is not grounded on any concession, or gift of the Subjects', but comes from a right that

> before the Institution of the Common-wealth, every man had (…) to everything, and to do whatsoever he thought necessary to his own preservation; subduing, hurting, or killing any man in order thereunto.[3]

The difficulties of such a thesis are: firstly, how can the right to punish be attributed to the public authority at all, if indeed it goes back to a situation previous to the foundation of that authority?; secondly, if there is no right to punish in the state of nature, how can it be said that such a state is the 'door' by which 'the Right, or Authority of Punishing in any case, came in'?[4]

What we would like to do in this paper is to show, first, how Hobbes tries to substitute for a justification of the right to punish by the social contract – since there can be no such justification – a justification by the utility of punishment for the maintenance of a legal state, secondly, that, if Hobbes stresses the distinction between acts of punishment and acts of hostility, it is because the justification of the right to punish by its consequences becomes hard to maintain when a sovereign must defend himself against enemies at home, thirdly, that, for that very reason, there is no need to dive deep into Roman penal law, as Giorgio Agamben had done in *Homo Sacer I*,[5] to understand the reason why the British state had instituted exceptional forms of punishment in the late Middle Ages, and, finally, that what is at stake, philosophically speaking, in this new consequentialist justification of punishment is not so much Roman law as a refutation of the justification of retributivism that was given by Francisco Suarez.

II. A Consequentialist Justification of Punishment

We will start, as Hobbes does in *Leviathan*, with the definition of punishment:

> A punishment, is an Evill inflicted by publique Authority, on him that hath done, or omitted that which is Judged by the same Authority to be a Transgression of the Law; to the end that the will of men may thereby the

[2] Th. Hobbes, *On the Citizen*, translated by R. Tuck and M. Silverthorne (Cambridge: Cambridge University Press, 1998), VI, para 5, 78 (hereafter: *On the Citizen*, chapter in Roman capitals, paragraph and page): 'The right of punishment is recognized to have been given to someone, when each one agrees that he will not go to the help of anyone, when each one agrees that he will not go to the help of anyone who is to be punished. (…). Men generally keep this kind of agreement well enough, except when they or those close to them are to be punished'.

[3] Th. Hobbes, *Leviathan* n 1 above, XXVIII, 482.

[4] ibid.

[5] G. Agamben, *Homo Sacer. Il potere sovrano e la nuda vita* (Torino: Einaudi, 1995); Id, 'Homo Sacer Sovereign Power and Bare Life', translated by D. Heller-Roazen, in Id ed, *The Omnibus Homo Sacer* (Stanford: Stanford University Press, 2017) (Hereafter, *Homo Sacer I*).

better be disposed to obedience.[6]

The meaning of this definition is commented upon by Hobbes in eleven 'inferences', that will help us better understand what theoretical difficulties punishment raises in *Leviathan*.

The definition rests on one central thesis: there is no punishment without a public authority, that is, without an authority capable of establishing laws, and judging what is in conformity with them, or contrary to them. Whatever the form of sovereignty, a sovereign acts as a public authority if his authority can be traced back to the terms of a covenant. The function of the covenant is thus to provide reasons to the citizens to let them know who the sovereign is and for what reasons the laws are to be obeyed. Whenever those conditions are met, citizens know that authorized judges are to be obeyed when they ask that someone be punished for trespassing the law. Those conditions – knowing who the sovereign is, that the laws have his approval, and that the judges judge according to the sovereign's authority – are sufficient for a legal system to function but not sufficient to justify punishment:

> Before I inferred any thing from this definition, there is a question to be answered, of much importance; which is, by what door the Right, or Authority of Punishing in any case, came in. For by that which has been said before, no man is supposed bound by Covenant, not to resist violence; and consequently it cannot be intended, that he gave any right to another to lay violent hands upon his person. In the making of a Common-wealth, every man giveth away the right of defending another; but not of defending himselfe. Also he obligeth himself, to assist him that hath the Soveraignty, in the Punishing of another; but of himself not. But to covenant to assist the Soveraign, in doing hurt to another, unlesse he that so covenanteth have a right to doe it himself, is not to give him a Right to Punish. It is manifest therefore that the Right which the Common-wealth (that is, he, or they that represent it) hath to Punish, is not grounded on any concession, or gift of the subjects.[7]

What are we bound to do then when we have accepted the terms of the social covenant? We are bound to assist the sovereign in punishing others but not to assist him in punishing ourselves. The limits of the sovereign's right is the subject's right to resist violence done to him, since 'no man is supposed bound by Covenant, not to resist violence'.[8] This affirmation raises a major difficulty since the sovereign is said to have received from the covenant, first, the power of 'Punishing with corporal, or pecuniary punishment, or with ignominy every

[6] ibid.
[7] ibid.
[8] ibid.

Subject according to the Law he hath formerly made', and, second, 'if there be no Law made', he is allowed to do

> according as he shall judge most to conduce to the encouraging of men to serve the Common-wealth, or deterring of them from doing dis-service to the same.[9]

The sovereign is thus endowed with a right to punish by covenant, which the subjects have an obligation not to resist, 'except when they or those close to them are to be punished'.[10] The latter remark, made in *De Cive* and repeated in *Leviathan*, shows the limits of the sovereign's right to punish. This limit is also expressed by the justification of the right to punish in relation with the state of nature:

> But I have shewed formerly, that before the Institution of Common-wealth, every man had a right to every thing, and to do whatsoever he thought necessary to his own preservation; subduing, hurting, or killing any man in order thereunto. And this is the foundation of that right of Punishing, which is exercised in every Common-wealth. For the Subjects did not give the Soveraign that right; but onely in laying down theirs, strengthened him to use his own, as he should think fit, for the preservation of them all: so that it was not given, but left to him, and to him onely; and (excepting the limits set him by naturall law) as entire, as in the condition of meer Nature, and of warre of every one against his neighbour.[11]

Since the covenant does not authorize the sovereign to punish the subject who authorized him, the justification of a right to punish based on covenant appears very thin. That is the reason why Hobbes looks for a new way of justifying the right to punish now based on the political consequences of punishing.

In the inferences from the definition of punishment he insists, indeed, on a different strategy for justifying the sovereign's right to punish: that justification has to be looked for in the utility of punishment for a commonwealth. The idea is that there is a common utility to others being punished, even though being punished oneself is always a bad thing for the criminal. From the general definition it can be inferred, first, that punishment is not revenge, that is, that it is not an evil done by a private person to compensate an evil previously done by another private person (Inference 1) – we'll come back to that point in Section 4. My first remark is that the definition also implies that punishment presupposes the existence of law, that is, a rule made by public authority. That characteristic of punishment is essential: it means that a punishment must be inflicted according

[9] Th. Hobbes, *Leviathan* n 1 above, XVIII, 276.
[10] Th. Hobbes, *On the Citizen* n 2 above, VI, 5, 78.
[11] Th. Hobbes, *Leviathan* n 1 above, XXVIII, 482.

to a public rule. To put it another way, when punished, I must be given the reasons why I am to be punished. This element of public justification constitutes a guarantee for the person who is punished, since one cannot be punished by a private authority (Inference 1) nor by a judge not appointed by the public authority (Inference 4), one must know the legal reasons stated by a judge for which she is being punished (Inference 3), the punishment cannot be greater than the one prescribed by the law (Inference 8), and one cannot be punished for an act performed before the law was passed (Inference 9). All those characteristic are part of the rule of law. But the truth is there are other characteristics in Hobbes's inferences from the definition of punishment that go in another direction.

My second remark is that the law, the judgement on the transgression of the law, and the punishment that is prescribed by the law for that particular transgression are not enough, according to Hobbes, to make a penal system work. Something else is required, that he calls the 'finality' of the penal system, which means that a legal system must be assessed according to its output. But what is this output? 'That the will of men may thereby (ie the punishment) the better be disposed to obedience'.[12] The will in question can either be the will of the punished or the will of other men who are aware of the punishment. It follows that a punishment that could not produce a disposition to obedience to the state would not be in conformity with the aim of the penal institution. Hobbes stresses the fact that penal law goes with a particular 'intention, or possibility of disposing the Delinquent, or (by his example) other men, to obey the Lawes'.[13] Punishment, therefore, is not only justified by its conformity to a pre-existing penal law, but also by the way the penal system to which a particular punishment belongs is capable of transforming a political system. Hobbes's main idea is that a penal system aims at producing obedience among citizens, so that contracts, laws, and promises can better be respected. John Rawls takes this Hobbesian idea very seriously, and associates it, in Section 38 of 'A Theory of Justice', with the name of Hobbes:

> By enforcing a public system of penalties government removes the grounds for thinking that others are not complying with the rules. For this reason alone, a coercive sovereign is presumably always necessary, even though in a well-ordered society sanctions are not severe and may never need to be imposed. Rather, the existence of effective penal machinery serves as men's security to one another. This proposition and the reasoning behind it we may think of as Hobbes's thesis.[14]

[12] ibid.

[13] ibid 484. We'll come back to that remark in the third section of the paper.

[14] J. Rawls, *A Theory of Justice* (Cambridge, Mass: Harvard University Press, 1971), Section 38 'The rule of law', 240.

This homage to Hobbes is sufficiently rare in Rawls that we should underline it.[15] But we must also be aware that there is a huge difference between the interpretation of the right to punish based on the consequences and the Rawlsian interpretation based on reciprocity. In a Rawlsian perspective, if I do not abide by a law I should consider it legitimate to be punished. Why is it so? Because of the reciprocity that prevails in the exercise of rights and liberties I can only accept to see others punished by the sovereign if I accept to be punished myself for similar reasons. But in Hobbes there is no reciprocity as to punishment. I may accept that others be punished and even assist the sovereign in exercising punishment of others but won't ever be persuaded to accept being punished myself. Indeed, such a violence against myself is in contradiction with the very reason why I have covenanted with others, that is, to be protected from all violence done to me. That is why, whereas Rawls can easily distinguish the rule of law from the state of war, Hobbes is sometimes at pains to distinguish between a punishment and an act of war.

III. Can an Act of Punishment Be Distinguished from an Act of Hostility?

There is an element in Hobbes's theory that does not find its place in Rawls's reading of Hobbes's theory of punishment, that is, Hobbes's insistence on distinguishing a punishment from what he calls an 'act of hostility', or 'hostile act'. Hobbes repeats the distinction in almost all of the inferences proceeding from his definition of punishment. Inference 4:

> (T)he evill inflicted by usurped power, and Judges without Authority from the Sovereign, is not Punishment; but an act of hostility; because the acts of power usurped, have not for Author, the person condemned; and therefore are not acts of publique Authority.[16]

The situation is the following: if the judges are not real judges, if the power enacting the punishment is a private power in disguise, there is not punishment but an act of war.

Why, however, does Inference 4 have it that the person condemned can be the author of the acts by which he is condemned, when it is said above that no punishment can ever be authorized by the very person condemned? That seeming contradiction may be resolved if we keep in mind that the person condemned has recognized a public authority as the source of punishment but never accepted to

[15] Another homage to Hobbes is to be found in J. Rawls, *Lectures on the History of Political Philosophy* (Cambridge, Mass: Harvard University Press, 2007), 23: 'Why do I begin a course in political philosophy with Hobbes? (…) My reason is that in my own view and that of many others, Hobbes's *Leviathan* is the greatest single work of political thought in the English language'.

[16] Th. Hobbes, *Leviathan* n 1 above, XXVIII, 484.

be punished by that authority.

The question remains: Why is the distinction between a punishment and an act of war so important for Hobbes? Because without it there would be no difference between a sovereign and a powerful enemy that would have taken power over a citizenry. Although not recognized as a legitimate source of punishment, a sovereign must be seen as a public authority, that is, as in charge of implementing penal law in the service of legal order, not as using it as a weapon against his subjects. That a sovereign cannot be subjected to penal law must be interpreted as a characteristic of it being a public authority:

> Hurt inflicted on the Representative of the Common-wealth, is not Punishment, but an act of Hostility; Because it is of the nature of Punishment, to be inflicted by publique Authority, which is the Authority only of the Representative it self.[17]

Another consequence of the distinction is that a subject who would no longer recognize the authority of his sovereign should be considered by him as an enemy, not as a subject to be punished according to penal law, but as a hidden enemy, 'who may lawfully be made to suffer whatsoever the Representative will'.[18] What changes is the perspective of the public agencies in inflicting the punishment – in the case of a normal punishment, the state considers the punished as a subject, in the case of a crime of *Lèse majesté* as an enemy:

> For in denying subjection, he denyes such Punishment as by the Law hath been ordained; and therefore suffers as an enemy of the Common-wealth.[19]

I may not recognize the authority of the sovereign to punish me, but I do recognize the sovereign as a public authority, and that is enough to allow me to be treated as a citizen, and not as an enemy if I do not abide by the law. Before seeing in the next section what a punishment must be like, it was important to understand that the public authority can decide by itself without further justification whether the transgressor must receive a regular punishment or be treated as an enemy.

IV. The 'Manner of Punishment' and the Implementation of the Law

A 'Dialogue Between a Philosopher and a Student of the Common Laws of England', which contains a careful examination of English penal law, is a good place to study the spirit of Hobbes's penal philosophy. The question is no longer to define in general what a punishment is but 'to define and appoint the special

[17] ibid 486.
[18] ibid.
[19] ibid.

manner of punishment'.[20] The philosopher encourages the student of the laws of England to expose the reasons for the variety of punishments – in contrast with the 'Stoics in old time' who thought 'that all faults are equal, and that there ought to be the same punishment for killing a man, and for killing a hen'.[21] The method of the *Dialogue* is different from the method used in *Leviathan* but the result of the enquiry as concerning the function of penal law is quite similar. Indeed, the philosopher of the *Dialogue* wants to persuade the student of the Common Law that the decision concerning the nature of punishments can be found in the reason of the sovereign. The student of the Common Law has a different idea: 'The manner of punishment in all crimes whatsoever, is to be determined by the common-law'.[22] But what does it mean to be determined by the Common Law? As the law of England knows two main sources, the statute, which is an act of Parliament, and the custom, those will also be, according to the student, the two possible sources for determining the nature of a punishment. But when the case is new, the student can see no reason why the judge could not determine the nature of the punishment by his own reason. The answer of the philosopher clarifies the point in question: if reason could settle the matter, there would be everywhere in the world the same penalties for the same offences, which is obviously not the case, therefore, natural reason by itself is not the solution. The absence of a universal legal system is sufficient proof that there is no rational determination of what a punishment should be for a given crime. The destruction of what may be called the natural law thesis in matters of punishment opens up the way for the positive law thesis. The scepticism of the student, who thinks that if no natural reason can determine the order of penalties there will no longer be any valid penal system, is countered by the philosopher who can see no reason for such scepticism. Taking for granted the impossibility to determine a punishment on a rational basis, the philosopher proves that the choice of penalties depends upon the decision of a particular man.

What comes foremost is not the justification of a particular punishment but the fact that the punishment be known in advance by all citizens, and that those citizens know that all punishments shall be implemented. Otherwise, the question is not of the rationality of a given penalty, since there can be many different rationales for the same penalty, but of the efficiency of the penal system as a whole. In order to be efficient, a penal system must be based on the knowledge of what actions are against the law, and of the punishments ordained for a given evil action. The source of rationality is in the law established by a person who

[20] Th. Hobbes, *A Dialogue Between a Philosopher and a Student of the Common Laws of England*, edited by W. Molesworth, VI (London: John Bohn, 1840; reprint, Aalen: Scientia Verlag, 1962), 121. Hereafter: *Dialogue*.

[21] ibid. Cf N. Machiavelli, *The Discourses*, translated by W.A. Oldfather (Cambridge, Mass: Harvard University Press, 1925), Book 1, 7, 30-33.

[22] Th. Hobbes, *A Dialogue* n 20 above, 121.

has the power to implement the law. Therefore, the question is not: how has a given punishment been established, but by whom was it established, and was that person in charge of public authority when she did so?

Hobbes provides an indication of what such an authority must be:

> Now the person to whom this authority of defining punishments is given, can be no other, in any place of the world, but the same person that hath the sovereign power, be it one man or one assembly of men.[23]

It is crucial that public authority be clearly recognizable by citizens, and that the link between an evil action and its punishment be without ambiguity. But why, it may be asked, should the public authority be the sovereign power? It could be envisaged, after all, that a public authority be the authority defining a range of penalties and making those penalties known to all citizens without any other prerogative. In such a case, there would be an autonomy of the penal system, since no one could be submitted to a penalty without a proper legal justification. That solution is not deemed sufficient by Hobbes who has it that the public authority which defines laws and punishment must be the authority that has the power to apply punishments. The difficulty, indeed, is not so much to define penalties for evil actions, nor to have those penalties known, but to have them applied when need be. What Hobbes has in mind is a situation in which several evil doers would unite to oppose public authority:

> For it were in vain to give it (that is, public authority) to any person that had not the power of the militia to cause it to be executed; for no less power can do it, when many offenders be united and combined to defend one another.[24]

Otherwise, the authority to establish punishments would be in vain, had it not conjoined to it the power to have the punishments executed, in particular, in circumstances when wrongdoers get politically organized.

There is no scepticism therefore in Hobbes's theory of punishment, but two very strong arguments: (1) What matters is not so much the *nature* of the punishments themselves, but the fact that they can be established by the authority capable of having them executed; and (2) The only power capable of having punishments executed is the sovereign power. Of course, that power must be recognized as a public authority, and the recognition comes with the process of authorization but, then, what matters is that the citizens abide by the power that has defined penalties, what matters, as said in the definition of punishment in *Leviathan*, is the obedience of the subjects. Although Hobbes does not remind us in the *Dialogue* of the finality of punishment, it is clear that

[23] ibid 122.
[24] ibid 122-123.

the aim of the various punishments discussed by the philosopher and the student of the Common Law of England is to frame the will of the citizenry and to make them abide by the laws of the sovereign.

V. The Statutes of the Provisors

When distinguishing between an act of punishment and an act of hostility Hobbes's real concern is to know what to do in cases of conflict between sovereignties. That aspect of his approach has not always been well seen, notably by Giorgio Agamben who stresses in *Homo Sacer I* the importance of Roman penal law for Hobbes's theory of punishment, whereas what matters is English Statute law. Let's consider for a moment the case of the Statutes of the Provisors, in Latin *de Praemunire* statutes, as understood by Hobbes.

An overly rapid examination of those statutes may conduce us to draw the wrong conclusions, for example, that sovereignty could be defined by its capacity to reduce men to their bare life, or to the fragility of their body when exposed to the violence of others. Agamben's thesis in *Homo Sacer I* is very much concerned by what he sees as a deep link between sovereignty and the 'bare life'. Instead of referring to Roman law he could have drawn on the Statute of Provisors as an even more acute example of a penal law exposing men to the violence of others with the permission of sovereign power. But, first, who are those Provisors? They are the persons who 'provide themselves with benefices', such as bishoprics and abbeys,

> founded and endowed by the Kings and nobility of England (but) bestowed by the Pope upon strangers, or such as with money in their purses could travel to Rome to provide themselves of such benefices.[25]

In section VI of the *Dialogue*, devoted to the punishments established by the Statutes of Provisors, the law student is specific on the content of the punishment for provisors:

> This crime is not unlike to that for which a man is outlawed, when he will not come in and submit himself to the law; saving that in outlawries there is a long process to precede it, and he that is outlawed is put out of the protection of the law. But for the offence against the statute of provisors (…) if the offender submit not himself to the law within the space of two months after notice, he is presently an outlaw. And this punishment (if not capital) is equivalent to capital. For he lives secretly at the mercy of those that know where he is, and cannot, without the like peril to themselves, but discover him. And it has been much disputed before the time of Queen Elizabeth, whether he might not be lawfully killed by any man that would,

[25] Th. Hobbes, *A Dialogue* n 20 above, 111.

> as one might kill a wolf: It is like the punishment amongst the old Romans of being barred the use of fire and water (*Interdictio de aqua et igni*), and like the great excommunication in the papacy, when a man might not eat, and drink with the offender without incurring the like penalty.[26]

This text clearly shows that Hobbes had a good knowledge of the different Statutes of Provisors, and particularly those of Edward III (25 Edward III, st. 1, c. 1, 1351; 27 Edward III), and of Richard II, and of their relation with Roman laws. He was familiar with a type of punishment that was equivalent to treating the guilty man like a wolf. Without doubt, an element of this relates to the mythologeme of the *homo homini lupus*, adding meaning to the connection Agamben makes between the idea of the man who is a wolf to other men, in the state of nature, and the punishment reserved by Provisors in English law.[27] Is it, however, a legal reference that validates Agamben's overall interpretation? Nothing could be less certain.

Before making any further analyses, it is imperative to put those royal writs into their context, which is that of a theory of international relations based on sovereignty. The Statutes of Provisors were adopted over the course of the 14th century in order to prevent the papacy from trespassing on the territory of the English sovereign through ecclesiastical nominations. Thus, the statute II Richard, c. 5, promulgated in 1392, was designed to prevent subjects of the King of England, or foreigners, from obtaining from the Pope the right to enjoy ecclesiastical benefices on English soil. Therefore, the punishment incurred by anyone who violated that law, wishing to assert in England a right that had been granted by Rome, served to sanction a conflict of sovereignty between a spiritual power and a temporal power, and not to reveal the hidden structure of sovereignty, based according to Agamben on the sovereign's capacity to expose men as such to its violence. The statutes of Provisors, already considerably relaxed during Hobbes's time,[28] were a means for the English sovereigns – well before the Anglican split – to establish strict limitations on the ecclesiastical courts' claims to judge *in ordine ad spiritualia*. Certainly, the fact that, on account of this conflict, some men found themselves exposed to prosecution and conviction over the method of legal killing of outlaws constitutes a particularity that should

[26] ibid 110; own italics.

[27] G. Agamben, *Homo Sacer I* n 5 above, 10: 'The protagonist of this book is bare life, that is, the life of *homo sacer* (sacred man), who may be killed and yet not sacrificed, and whose essential function in modern politics we intend to assert.' The relevant pages for the comparison we suggest between the archaic figure of *homo sacer* and the statutes of provisors are to be found in the section 'The ban and the wolf' (ibid 88-92).

[28] The situation described by Hobbes only partly corresponds to the law of his era, which only provided for immediate execution with no other trial in cases of felony, but not in a case involving civil action. Lucien Carrive notes that, in the latter case, 'the outlaw only lost his personal belongings, and even then he was often able to put them in a safe place' (Th. Hobbes, *Dialogue des Common-Laws d'Angleterre*, French translation by L. Carrive (Paris: Vrin, 1990), 139, n 3).

be questioned by anthropologists of punishment, but it is doubtful that it can be made the very paradigm of sovereign power.

Since the return to legally programmed bare life – a life deprived of all legal guarantee – is caused by the coexistence of sovereign States, it is essential to carefully analyse the effects of conflicts of sovereignty on subjects exposed to them. Indeed, contemporary mass slaughters or genocides often testify to the horrid consequences of conflicts of sovereignty. In the *Dialogue*, the extreme violence of the long-lasting conflict between England and the papacy is indeed striking. But the exposure of men to the violence of other men as permitted by the Statutes of Provisors has not much to do with the bare life, although men could be treated in such cases as wolves; it has to do with what a sovereign can do to maintain his sovereignty when threatened by a 'spiritual' power. In matters of sovereignty, Hobbes's logic can be said to be straightforward, as is shown by a critique he addresses to Edward Coke regarding punishment for treason. To the great lawyer who says that a traitor cannot be punished by the king of England without an indictment, a procedure proper to English Common Law, Hobbes answers the following:

> This is not an argument worthy of the meanest lawyer. Did Sir Edward Coke think it impossible for a King lawfully to kill a man, by what death soever, without an indictment, when it is manifestly proved he was his open enemy? Indictment is a form of accusation peculiar to England by the command of some King of England, and retained still, and therefore a law to this country of England. But if it were not lawful to put a man to death otherwise than by an indictment, no enemy could be put to death at all in other nations, because they proceed not, as we do, by indictment.[29]

In order to have an enemy killed, even though the enemy be a former citizen, a sovereign does not need to abide by the laws of his own country: to make this idea still clearer Hobbes goes back to the time of the Conquest, a paradigmatic historical case for considering what a sovereign does when a kingdom is subdued:

> William the Conqueror subdued this kingdom; some he killed; some upon promise of future obedience he took to mercy, and they became his subjects, and swore allegiance to him. If therefore they renew the war against him, are they not again open enemies? Or if any of them lurking under his laws, seek occasion thereby to kill him secretly, and come to be known, may he not be proceeded against as an enemy, who, though he had not committed what he designed, yet had certainly a hostile design?[30]

It is not necessary to press that point any further: Hobbes makes the

[29] Th. Hobbes, *A Dialogue* n 20 above, 74.
[30] ibid.

distinction between punishment and hostility in order to show that punishment is relevant when the sovereign is recognized as the public authority, and that punishment becomes an act of war when the sovereign is no longer acting as the public authority. These are questions that are still relevant to our present situation when it comes to discussing legal cases about terrorists returning to their home country. How should they be treated? As citizens to whom the protection of law is due, or as enemies to the commonwealth?

Establishing the right kind of punishment in those particular cases is part of what may be called a politics of sovereignty. The *Dialogue* shows that Hobbes was very much concerned, for the sake of sovereignty, with the actual forms that the right to punish can take in a commonwealth. To be sure, Hobbes's concern is more for having a multitude obey the law than for the State to respect the principle of liberty:[31] he does not follow here a Rawlsian path by anticipation. But it would be a greater mistake to follow Agamben's suggestions, since Hobbes is willing to preserve the guarantees of a legal system. The fear that goes with the state of nature accompanies, certainly, his political ideas on punishment, but the aim of this fear is not to make the state of nature rule in civil society, or to reduce human life to the bare life. Why so? Because (1) a Hobbesian politics of punishment goes with a strict respect for the forms of the law – no punishment without a prior law establishing the judicial reasons of the punishment; and because (2) such a politics does not aim at developing among the citizenry the terror that goes with a politics of exception but to give us supplementary reasons to abide by the law.

VI. Examining Punishment in the *Questions*: What is Punishment Good for?

Asking who can be punished obliges one to reflect, as we have done in the previous section, on the fact that punishing is not equivalent to waging war. However, such a distinction does not tell us what punishment can be good for in a society. Why is it that the public authority wants to punish its citizens who have trespassed the limits determined by law? Hobbes's answer in *Leviathan* is: 'to the end that the will of men may thereby the better be disposed to obedience'.[32] That is also the answer that he gives in the *Questions*:

> The intention of the law is not to grieve the delinquent for that which is passed and not to be undone; but to make him and others just, that else would not be so: and respecteth not the evil act past, but the good to come. Insomuch as without this good intention of future, no past act of a

[31] On Rawls and Hobbes on the state of nature, cf L. Foisneau, 'Sortir de l'état de nature (Rawls)', in Id ed, *Hobbes. La vie inquiète* (Paris, Gallimard, series: folio/essais, 2016), 464-503.

[32] Th. Hobbes, *Leviathan* n 1 above, XXVIII, 482.

delinquent could justify *his* killing in the sight of God.[33]

What is new in the *Questions* is the setting of this particular part of the definition of punishment within the context of a larger discussion, no longer on the common law, but on our relations to the law of nature, and the right of nature. We'll try and show how the latter debate can help us better understand the intellectual context of Hobbes's discussion of punishment. The truth is that the debate with Bramhall is a source of fruitful objections – Bramhall might be the anonymous source of other famous objections, those that led Hobbes to write extremely helpful remarks in the second edition of *De Cive*.

A first objection conduces Hobbes to justify the end he attributes to punishment: 'But you will say, how is it just to kill one man to amend another, if what was done were necessary'?[34] Hobbes gives here, in a nutshell, what he considers to be the spirit of Bramhall's objections: without free will, and the concomitant capacity to choose between good and evil, rewards and punishments are undeserved, and, therefore, perfectly unjust. That objection corresponds to a theory of natural law that can be found in Suarez, to which Hobbes's natural law theory offers, obviously, a sharp contrast. In Bramhall's Suarezian theory, what matters is that the agent be capable of choice: the punishment will be justified by the fact that the agent did not make an appropriate use of his liberty, using it instead to perpetrate an evil act. The punishment inflicted can be therefore understood as a response to the evil use of liberty: it is a way to make the culprit pay for what evil he has done. The problem to be solved is to find the appropriate equivalent between the evil acted and the punishment inflicted. That kind of theory belongs to retributivism. In contrast, Hobbes's response to Bramhall's objection shows how his theory is distinct:

> To this I answer, that men are justly killed, not for that their actions are not necessitated, but that they are spared and preserved, because they are not noxious; for where there is no law, there no killing, nor any thing else can be unjust. And by the right of nature we destroy, without being unjust, all that is noxious, both beasts and men.[35]

In a somewhat surprising manner, Hobbes discards the criterion of free will to replace it by the criterion of noxiousness: to establish the reason of a punishment there is no need to prove the liberty of the agent before the action was done, since there is no such liberty, but one must know if there were a law, or no law, established by the state, and judge the action accordingly. When there is no law, that is, in the state of nature, one cannot say that a killing or any

[33] Th. Hobbes, *The Questions Concerning Liberty, Necessity and Chance*, edited by W.Molesworth (London: John Bohn, 1841), V, 152.

[34] ibid.

[35] ibid.

act is unjust. Hobbes rejects the idea, present in Suarez, that the problem is to establish the freedom of the agent; the question is, according to him, to determine whether there is, or not, a law. In the absence of a law, there is no reason to say that an act was unjust, even though that act was 'necessitated'. It is not the liberty of the agent that matters, understood as a liberty to act according to the injunction of natural law but the consequences of his actions in his relationship to others. If there is no public authority, there is no limit to what can be done in order to achieve self-preservation among other animals but also among other human beings. One characteristic feature of the arguments advanced in the controversy with Bramhall is the fact that Hobbes takes animals into consideration, and opens up the limits of the state of nature, making it somewhat like a Darwinian state of nature *avant la lettre*: 'And for beasts, we kill them justly, when we do it in order to our own preservation'.[36] The truth is that 'justly' has no place in the state of nature, since justice appears only within the limits of a civil state. What Hobbes means is that animals can be killed by men without men having to feel guilty about the killing in the absence of a law of nature. The change introduced by him in natural law theory has consequences in various fields: in our relations with other men in the state of nature, but also in our relations with other species. But why is it that men can't be punished for their way of treating other animals? The reason is that there can only be punishment when there is a civil law forbidding or commanding a given behaviour. It is therefore public authority that gives meaning to the act of punishment: contrary to what Locke says, Hobbes considers it absurd to speak of punishment in the state of nature since the laws of nature have no application there.

But what makes it so important for a public authority to be able to implement punishments?

> For men, when we make societies or commonwealths, we lay down our right to kill, excepting in certain cases, as murder, theft, or other offensive actions. So that the right which the commonwealth hath, to put a man to death for crimes, is not created by the law, but remains from the first right of nature, which every man hath to preserve himself; for the law doth not take that right away, in case of criminals, who were by law excepted.[37]

Despite its new theory of authorization, *Leviathan* had kept elements of the theory that was already there in the *Elements of Law*, and the *Questions*, just quoted, are still keeping the same line of argument. Punishing a criminal relies on the fact that the public authority has not been given a new right when it was established by the covenant, but has maintained the right the natural person of

[36] ibid.
[37] ibid 153.

the sovereign had in the state of nature. Answering Bramhall's Suarezian objection – that 'Men are (…) put to death or punished, for that their theft proceedeth from election'[38] – conduces Hobbes to be more specific in establishing the scope of punishment. His thesis can be called therefore exemplarism, and as such opposed to Bramhall's retributivist theory. Why so? Because in retributivism, punishment endeavours to repair the evil done by a free person, whereas in exemplarism, punishment aims to create an example that will reinforce the motivation of citizens to abide by the law. Not considering so much the nature of the will, Hobbes proposes a social theory of punishment in which the example of the criminal being punished is supposed to deter others from imitating him:

> Men are not therefore put to death or punished, for that their theft proceedeth from election; but because it was noxious and contrary to men's preservation, and the punishment conducing to the preservation of the rest: inasmuch as to punish those that do voluntary hurt, and none else, frameth and maketh men's wills, such as men would have them.[39]

The debate with Bramhall makes it clear that a theory of punishment is central to a politics of sovereignty aiming at the preservation of its citizens. If establishing the right kind of punishment is an essential part of such a politics, it is because it contributes to motivating citizens to abide by the law. The *Dialogue of the Common Laws* shows, as we have seen in the previous section, that Hobbes was very much concerned with the actual forms that the right to punish can take in a commonwealth; the *Questions* show that the politics of punishment in Hobbes accompanies the refutation of a natural law theory based on a metaphysical theory of free will, since what is central to the whole question is not so much the liberty of subjects when they commit mischiefs as the political consequences of punishment.

To be sure, Hobbes does not follow a Rawlsian path by anticipation: his concern is more for having a multitude follow the law than for the State to respect the rule of law. But it would be a mistake to follow Agamben's suggestions in our reading, since Hobbes is willing to preserve the guarantees of a legal system, and does not want to have the state of nature reappear in the civil state. The fear that goes with the state of nature accompanies, surely, the politics of punishment, but the aim of this fear is not to make the exception become the rule, or to reduce human life to the bare life, because (1) a Hobbesian politics of punishment goes with a strict respect for the forms of the law (no punishment without a prior law establishing the reasons of the punishment); and because (2) such a politics does not aim at developing among

[38] ibid.
[39] ibid.

the citizenry the panic that goes with a politics of exception but to give citizens additional reasons to abide by the law.

A last remark in guise of conclusion: the form taken by Hobbes's theory of the finalities of punishment can also be explained by the fact that he knew from direct experience of a civil war that it is sometimes very difficult for a public authority to maintain a clear cut distinction between punishments and acts of hostility. Indeed, punishments can sometimes be seen, and rightly so, as acts of hostility, that is, as attempts to separate enemies from the rest of the citizenry. Hobbes could have been more explicit on the reasons that prevent us from thinking that all punishments are expressions of the hostility of the state. Going back to the 'right to everything' (*jus in omnia*) as a foundation for a right to punish is probably not the best way to obtain such a clarification. Some cases analysed in the *Dialogue,* the Statutes of Provisors in particular, are good examples of a penal law in which a war-like politics can be read, a politics in which the distinction between punishment and acts of war is hard to maintain. These statutes are, indeed, the translation in the realm of legality of a long-standing struggle between the kings of England and the Papacy. That bellicose dimension can explain the cruelty of the penalty, and the fact that persons so condemned could be killed by whomever found them on English land, without further legal proceedings.[40] But the situation is still very different from the situation described by the ancient Roman texts concerning *homo sacer*: the ambiguities or frailties of Hobbes's theory are also the expression of the frailty of the British state of his time, confronted by civil war.

[40] There is no apology for political violence in Hobbes, as his awareness of men's cruelty shows sufficiently: 'Thomas Hobbes said that if it were not for the gallowes, some men are so cruell a nature as to take a delight in killing men more than I should to kill a bird. I have heard him inveigh much against the Crueltie of Moyses for putting so many thousands to the Sword for Bowing to the Golden Calf' (J. Aubrey, *Aubrey's Brief Lives*, edited by O. Lawson-Dick (London: Mandarin paperback, 1992), 157).

‘Public Enemy’? Difficulties in Rousseau’s Theory of Punishment

Francesco Toto*

Abstract

This article focuses on references to the issue of punishment disseminated in the *Social Contract*. Through the analysis and contextualization of these references, it aims primarily to frame Rousseau’s theory of punishment within the broader context of his political theory. It focuses in particular on the apparent tension between conceiving the criminal as a citizen on the one hand, and as a public enemy, external to the State and to the legal guarantees reserved to its members, on the other. Finally, it attempts to highlight the underlying coherence of Rousseau’s discussion by showing that not just any criminal is a ‘public enemy’, but only the political criminal, that is, the usurper or the despot.

Punishment is certainly a peripheral subject in the *Social Contract*; it is very quickly discussed in just two contexts and does not reappear except in passing in a handful of passages. Its marginality, however, does not prevent it from challenging some fundamental elements of Rousseau’s theory. In fact, Rousseau’s treatment of the right/duty to punish raises questions both for the central problem of the essence of the State and for the problem, no less crucial, of the codependence of the apparently opposed elements of force and law, obedience and freedom, general and particular will, legislative and executive power. Moreover, considering the moral presuppositions of punishment reveals both the friction between two potentially conflicting conceptions of it – one of a legal-moral nature, the other of a so-called ‘political’ nature – and the tension between the absoluteness of sovereign power and the legitimacy of resistance. To test the significance and coherence of Rousseau’s discussion of punishment, it may then be useful to start from a passage taken from the concluding chapter of Book II, in which different types of laws are defined.

Leaving aside custom, which Rousseau calls the ‘most important of all’ kinds of law, there are three types of law listed in the *Contract*: political laws, which ‘constitute the form of government’; ‘civil laws’, which regulate the relations of citizens with each other and with the State; and ‘criminal laws’, which regulate the relations between ‘disobedience and penalty’, and therefore

* Assistant Professor, Roma Tre University, Department of Philosophy, Communication and Performing Arts.

represent 'not so much a particular type of law as a sanction for all others', that is, the sanction of political laws and civil laws.[1] There is something enigmatic about this definition of criminal law.[2] What kind of law could by definition not be 'a particular kind of law'? The enigma deepens as soon as we consider that these laws are meant to regulate the relationship between 'disobedience and punishment', whereas for Rousseau 'the essence of the body politic lies in the harmony of obedience and freedom'.[3] The opposition both between obedience and disobedience and between freedom and punishment allows us to perceive an invisible link between the essence of the State and punishment, but it does not yet allow us to clarify the nature of this link. Certainly, one can suppose that if disobedience is opposed to obedience, then punishment, as a repression of disobedience, fulfils the function of safeguarding the integrity of the State, of promoting the obedience which constitutes its essence. But how could punishment consisting in repression at the same time preserve that freedom which should also constitute the State's essence, that is, how could it enforce obedience without damaging at least the freedom of the criminal, and possibly the freedom of those who obey only from fear?

In order to clarify the first aspect of the enigmatic definition of criminal law, it is necessary to focus on the role of force. As is well known, for Rousseau the general will expressed in the laws is not necessarily the 'will of all': 'indeed, every individual can, as a man, have a private will contrary to or differing from the general will he has as a citizen', because 'his private interest can speak to him quite differently from the common interest', 'his absolute and naturally independent existence can bring him to view what he owes to the common cause as a free contribution'.[4] If the establishment of society is made 'necessary' by the 'opposition of private interests' and possible by the 'agreement of these same interests', then it is also true that man is not erased by the citizen. Particular wills and interests are not eliminated by the general will and interest, nor is conflict eliminated by commonality and agreement.[5] In the framework of this codependence of apparent opposites, between general and particular,

[1] J.J. Rousseau, *Du contrat social ou principes du droit politiques*, texte établi et annoté par R. Derathé, in Id, *Euvres complètes*, sous la diréction de B.B. Gagnebin et M.M. Raymond (Paris: Gallimard, 1959-1995), III, 1964, II, 12; English translation, *Social Contract*, in *The Collected Writings of Rousseau* (Hanover-London: University Press of New England, 1990-2010, 1994), IV, 164-165.

[2] B. Bernardi, 'Le droit de vie et de mort selon Rousseau: une question mal posée?' *Revue de Métaphysique et de Morale*, 89, no 97 (2003), rightly observes that criminal laws are not a subject of frontal examination in the *Contract*. This observation is intended to indicate, again rightly, the secondary role played by the question of punishment in his reading of CS II, 5. This does not mean, however, that the question of punishment in general and the death penalty specifically did not interest Rousseau during the writing of the *Contract*.

[3] J.J. Rousseau, n 1 above, III, 13, 190.

[4] ibid II, 3 and I, 7, 147, 140-141.

[5] ibid II, 1, 145. The common interest 'is formed by opposition to that of each person', which in turn opposes it (ibid II, no 3, 147).

between identity and difference, between agreement and conflict, the particular interests that, on the one hand, contain the general and are therefore able to find in the law a point of mediation and convergence are the same that, on the other hand, continue to disturb the unity of the political body by opposing themselves to one another and to the general interest expressed in the law. Each individual is indeed received as an 'indivisible part of the whole' by a body politic which acquires 'absolute power over all its members', yet at the same time it continues to exist as a 'perfect and solitary whole', endowed with an 'absolute and naturally independent existence', with instincts, appetites, and a 'private will (which) tends by its nature towards preferences'; as such, it cannot have any 'lasting and unchanging' agreement with the general will.[6] In this framework, it seems that the political body can exist as a 'union of its members' only to the extent that the gap between particular and general will is filled by the intervention of a 'repressive force' capable of unifying, if not the first will to the second, at least the external conduct of citizens to the obligations imposed by law.[7] This hint of a force 'that can prevail over the resistance'[8] seems to offer a solution to the problem of the status of criminal law. The 'public force' is the 'guarantee' of the solidity of the political body, because it represents a means to ensure that citizens 'fulfil (their) duties' and 'to be assured of their fidelity'.[9] As Rousseau explains,

> in order for the social compact not to be an ineffectual formula, it tacitly includes the following engagement, which alone can give force to the others: that whoever refuses to obey the general will shall be constrained to do so by the whole body.[10]

Even if Rousseau does not establish any explicit connection between punishment and force, it is reasonable to assume that punishment constitutes an essential expression of force.[11] 'Just as nature gives each man absolute power over all his

[6] ibid I, 6, II, 4, II, 7, I, 7, II, 1, 139, 148, 155, 141, 145.
[7] ibid II, 4 and III, 1, 148 and 168.
[8] ibid I, 6, 138.
[9] ibid II, 4 and I, 7, 150 and 140-141.
[10] ibid I, 7, 141.
[11] This connection is suggested, among others, by G. Silvestrini, 'Fra diritto di guerra e potere di punire: il diritto di vita o di morte nel *Contratto sociale*' *Rivista di Storia della Filosofia*, 125-142 (2015). Silvestrini connects the right to punish with the ability of the legitimate State force to 'compel to be free'. However, a clarification is needed. The right to compel is not in fact identical to the right to punish: it is one thing to say 'the law prescribes X, therefore the sovereign has the right to force citizens to do X', and another to say 'the law prescribes X, therefore the sovereign has the right to punish citizens who do not do X'. Whoever is punished is not forced to obey the violated law: he is forced to suffer what the criminal law prescribes. Thus, I believe an intermediate step is required to connect the right to punish with the right to coerce: the threat of punishment forces those exposed to it to obey the law. When Rousseau speaks of coercion to be free, then, he has obedience in mind, and not disobedience: coercion makes one free because it reconciles the particular will with the general will expressed in the law. In order to understand whether the

members, the social compact gives the body politic absolute power over all its members', but for this purpose 'it must have universal, compulsory force to move and arrange each part in the most convenient way to the whole', and this coercive force seems to be closely connected with the capacity to punish citizens for violating the law.[12] If criminal law is not a 'particular kind of law' but the sanction of all laws, this is because there is no law that is not either applied by citizens or enforced by the executive power. But just as enforcement by citizens finds a decisive motivation in the threat of punishment, so too enforcement by the executive recognizes punishment as its essential expression. Though, criminal law merely regulates the punishment of political or civil crimes, it is at the same time the condition under which all other laws can be more than a vain 'collection of formulas'.

The solution to the first difficulty raised by the definition of the criminal law brings up the second aspect of the enigma. The 'coercive force' of punishment promotes obedience, but is obedience ensured by the threat or exercise of force compatible with freedom? Does it not jeopardize the combination of obedience and freedom which constitutes the essence of the State? 'The right of the strongest', says Rousseau, is at most a 'right (…) taken ironically', because 'force is a physical power' whose effects are devoid of any morality.[13] 'Yielding to force', obeying, is in fact 'an act of necessity, not of will', and therefore neither free nor obligatory, because where there is no freedom there are no duties.[14] One is certainly 'obliged to obey legitimate powers', but since 'force does not make right' it has no legitimacy.[15] Indeed, 'if it is necessary to obey by force', Rousseau writes, 'one need not obey by duty and, and if one is no longer forced to obey, one is no longer obliged to do so', to the point that 'as soon as one can disobey without punishment, one can do so legitimately'.[16] If those with sufficient strength can escape punishment, the body politic can exercise its essential right to punish only with a force capable of overcoming all resistance. But the only relationship that force can establish is that between servant and master, that is, the opposite of a political relationship, in which obedience cannot be separated from freedom. Wherever there are 'master and slaves' there can be an aggregate, never a society, a 'body politic': 'the moment there is a master (…) the political body is destroyed', because the 'people (…) is (in) no way obligated toward (their) master'.[17] From this it seems that we are at an impasse with no

coercion carried out by means of punishment can represent a form of liberation of the criminal, it will therefore be necessary to ask whether it can be desired by the criminal himself. As we shall see, it can be the subject of the criminal's will because it is governed by a law which the criminal himself, as a citizen, has approved.

[12] J.J. Rousseau, n 1 above, II, 4, 148.

[13] ibid I, 3, 133.

[14] ibid

[15] ibid I, 3, 134.

[16] ibid I, 3 and 4, 133-134.

[17] ibid I, 3, II, 1, I, 4, 137, 145.

way forward. On the one hand, the State cannot exist without the capacity of exercising force sufficient to ensure that laws and punishments be maintained all possible resistance, but on the other hand, it must cease to exist as soon as it resorts to that force, thereby actualizing the converse of a political relationship. To save the distinction between political power and tyrannical or despotic power, it is necessary to indicate the conditions under which the 'coercive force' of punishment can be compatible with freedom in a republican State. How can punishment guarantee the legitimate interests and freedom of all citizens, including evildoers?

The challenge of the *Contract* is to reconcile security and legitimacy, interest and law, utility and justice. The necessity of these alliances is revealed by general anthropological assumptions. On the one hand, the nature of man requires him 'to attend to his own preservation' and 'his first cares' are those 'he owes himself'.[18] On the other hand, 'to renounce one's freedom is to renounce one's status as a man' and is thus 'incompatible with the nature of man'.[19] The duality of interest and freedom, therefore, can only be apparent: 'common freedom' – that is, being 'master of oneself' – consists precisely in being the only 'judge of the proper means of preserving (one)self', that is, of one's own interest.[20] The theory of punishment is called to fit into the framework outlined by these assumptions. The link between punishment and the general interest can be grasped by paying attention to the utilitarian aspect of Rousseau's theory. The 'collaboration of many' carried out by the political institution represents, for each individual, the only way to preserve himself.[21] As a 'sum of forces' capable of getting the upper hand over both internal and external threats to the security of citizens, it embodies 'a form of association that defends and protects the person and the goods of each associate with all the common force'.[22] 'Since all are born equal and free, they only alienate their freedom for their own utility', but this utility is paradoxically implicit in the totality of alienation itself:

> since each one gives his entire self, the condition is equal for everyone, and since the condition is equal for everyone, no one has an interest in making it burdensome for the others.[23]

Through the laws 'everyone necessarily submits himself to the conditions he imposes on others', and this equal submission is enough to ensure 'an admirable agreement of interest and justice'.[24] 'Formed solely by the private individuals

[18] ibid I, 2, 132.
[19] ibid I, 4, 135.
[20] ibid I, 2, 132.
[21] ibid I, 6,
[22] ibid.
[23] ibid I, 2 and I, 6, 132 and 138.
[24] ibid II, 4, 149.

composing it', in other words, the sovereign 'does not and cannot have interests contrary to theirs': the sovereign may therefore need 'guarantees' from his subjects, but they have no need of 'guarantees' on his part.[25] Every sovereign decision, that is, every law, is not only 'equitable, because it is common to all', but also 'useful, because it can have no other object than the general good'.[26] Inasmuch as it is governed by a law, then, punishment can and indeed must be imposed in the interests of citizens: it is precisely by means of it that the State carries out its mission of protecting their person, their security and their rights. The link between punishment and freedom understood as 'perfect independence' of 'each citizen (...) from all others' is made clear by Rousseau when he argues that 'only the force of the State creates the freedom of its members', because the only guarantee against the arbitrariness and violence of their fellow men is represented by the 'full vigour' of the laws, and among these in particular by the strength of the criminal laws that fix the 'sanction' without which 'the laws of justice are ineffectual among men'.[27] As Rousseau states in a note that rephrases an anecdote already discussed by Hobbes,

> in Genoa we read on the entrance of prisons and on the irons of convicts this word, *Libertas*. This application (...) is beautiful and right. In fact, in every State it is only the evildoers who prevent the citizen from being free. In a country where all these people were in prison, they would enjoy perfect freedom.[28]

Yet this explanation of the link between punishment and freedom, understood as independence or protection of private individuals from the violence of their peers, still leaves some questions open.[29] It is still unclear how the same individual who will face punishment can commit his own strength and freedom by giving the sovereign the right to punish him, and this 'without harming himself and without neglecting the cares he owes to himself'.[30] It is also unclear how Rousseau can avoid the conclusion that respectable citizens pay for their protection from the violence of their equals by sacrificing not only of the

[25] ibid I, 7, 140.
[26] ibid II, 4, 150.
[27] ibid II, 12 and IV, 7, 164, 215 and 152.
[28] ibid, note to IV, 2, 200-201.
[29] It is this first meaning of freedom that Frederick Neuhouser has in mind when he states that 'universal compliance with the general will effectively safeguards citizens from personal dependence and that this protection from dependence is so bound up with their freedom that obedience to the general will be said to make them free, even when their obedience is not voluntary in the ordinary sense of the term': universal compliance with the general will, ie obedience to the law, is a necessary and sufficient condition of a kind of freedom defined as independence from personal ties, and the fact that this obedience is not 'voluntary in the ordinary sense of the term' means that it can be imposed on those who have committed or would like to commit a crime by force or threat of force. See F. Neuhouser, *Foundations of Hegel's Social Theory* (Cambridge: Harvard UP, 2000), 63
[30] J.J. Rousseau, n 1 above, I, 6, 138.

freedom of the evildoer, but also, after all, their own freedom before the State. Speaking of total alienation, did not Rousseau really go further than Hobbes, who at the very least excluded the right of resisting the violence of punishment from the rights ceded by his subjects to the sovereign?

These doubts can be clarified by focusing on the marginal but exemplary case of the criminal: the reasons why punishment is compatible with the interests and freedom of the criminal are the same as those which, a fortiori, it is compatible with the interests and freedom of the innocent. The sense in which the penalty may be in the interest of the same person who will be punished as a criminal is clarified by analogy. Citizens protected by the State may find themselves risking their own lives in order to defend it in war, and in doing so they 'give back to the State what they have received from it': 'whoever wants to preserve his life at the expense of others should also give it up for them when necessary', and accepting this risk does not involve renouncing life at all, but merely putting it at stake 'to preserve it'.[31] Similarly, says Rousseau, 'in order not to be the victim of a murderer' one can agree 'to die if he becomes one', and in so doing he does not dispose of his life, but 'only thinks of guaranteeing it'.[32] Even the death penalty, therefore, can be considered useful to all citizens, including those who will suffer it, because it does not contradict the primacy of care due to one's own preservation but rather constitutes a tool to exercise this care. The law that establishes the death penalty is, in fact, the very law that allows citizens to live in safety, at least until they violate the laws for which it provides the sanction. Reconciling punishment with the freedom of the criminal is a more complex matter. For Rousseau, the social pact constitutes 'the most voluntary act in the world', because 'by its nature (it) requires unanimous consent', and its validity also depends on the fact that every citizen, 'uniting with all', continues to obey 'only himself' and to remain 'as free as before'.[33] Since he

[31] ibid II, 4 and 5, 150-151.

[32] ibid II, 5, 151. On the whole Chapter V of Book II of the *Contract* see B. Bernardi, n 2 above. Bernardi clarifies the meaning of the chapter starting first of all from its genesis: he begins from a manuscript annotation that discusses the difference between *péril*, *danger* and *risqué*; he follows the first draft presented in the *Geneva Manuscript*; he explains the most significant variations introduced by the *Contract* and makes explicit the close relationship with the problem of the limits of sovereign power dealt with in Chapter IV. In particular, Bernardi shows that the unity of the Chapter does not lie in the discussion of the right to punish or the death penalty but in the discussion of the right to life or death of the legitimate State with respect to its citizens, and underlines the difference between the logic of risk, according to which one can put one's life on the line to guarantee one's security and freedom, and the logic of sacrifice, in which one meets certain death for an interest greater than one's own.

[33] J.J. Rousseau, n 1 above, IV, 2 and I, 6, 200 and 138. The freedom that marks the constitution of the sovereign people, that is 'the act by which a people is a people', continues to characterize political and civil life. The 'natural freedom', which for the individual consists in 'unlimited freedom to everything that tempts him' and therefore has no other 'limit than (his) forces', must certainly be distinguished from 'moral freedom' or 'civil freedom', which consists in 'obedience to the only law one has given oneself' and is limited instead by 'general will' and the law.

violates in others the rights that he hopes others will continue to respect in him, the criminal puts his own particular interest before the general one, denying the reciprocity that constitutes the heart of the social pact and thus committing 'an injustice whose spread would cause the ruin of the body politic',[34] that is to say, of the instrument that he himself had created by signing the social pact as a guarantee of his own security and freedom. Within this framework, in which Rousseau re-elaborates the Hobbesian theory of injustice as a self-contradiction of the will,[35] it makes no sense to ask 'how one is free yet subject to the laws', including criminal laws, 'since they merely record our wills', nor 'how (…) a man (can) be free and forced to conform to wills that are not his own', such as laws to which he is opposed, because in reality 'the Citizen consents to all the laws, even to those passed in spite of him'.[36] 'As long as subjects are subordinated only to such conventions', 'even to those that punish (them) when (they dare) to violate one of them', they 'do not obey anyone, but solely their own will': they obey only 'the general will which is theirs', that is, 'the general will (they have) as citizen(s)'.[37] Therefore, 'to ask how far the respective rights of the Sovereign and of Citizens extend is to ask how far the latter can engage themselves', where of course no one can be unjust toward themselves.[38] Slaves can lose everything once they are in chains, even the desire to break free of them, in which case they love 'their servitude as the companions of Ulysses loved their brutishness'.[39] But the citizen who gives himself laws capable of forcing him into obedience and

The civil liberty won in the instant of the pact represents not only and not so much the denial of natural liberty, but its realization on a different level. See ibid, I, 6 and I, 8.

[34] ibid I, 7, 141.

[35] See L. Foisneau, 'Punishment Not War: Limits of a Paradigm' in this issue.

[36] ibid II, 6 and IV, 2, 153 and 200. C. Brettschneider, 'Rights within the Social Contract: Rousseau on Punishment', in A. Sarat et al eds, *Law as Punishment/Law as Regulation* (Stanford: Stanford UP, 2011), 50-76 rightly observes that in Rousseauian theory the legitimacy of a State act depends on the unanimous consent to which it is subject, and questions the sense in which the subject can allow his own punishment. It seems to me that Rousseau's position is greatly weakened by the reduction of this consensus to the 'hypothetical consensus' that in a kind of thought experiment any citizen could have given at the time of the pact, and by the consequent assimilation of Rousseau's theory to contemporary theories such as that of Rawls. Rousseau's aim is not to build an ideal standard, but to highlight the internal logic of law: the consent is real, and not hypothetical, insofar as it is implicit not in a hypothetical pact, but in the consensus given to the laws and in the daily recognition of oneself as a citizen of a State. For a discussion of the claim that every citizen also allows laws that pass against his vote, along with the difficulties it faces, see C. Brooke, 'Aux limites de la volonté générale: Silence, exil, ruse et désobéissance dans la pensée politique de Rousseau' *Les Études Philosophiques*, 425-444 (2007).

[37] J.J. Rousseau, n 1 above, II, 4, IV, 2 and I, 7, 150, 200 and 140.

[38] ibid II, 4, 150. The most intriguing problem posed by these passages is that of personal identity in relation to time. Only on the condition that the two are considered the same man can the Socrates who approves the social pact or law commit himself for the Socrates who then, having violated them, finds himself suffering the punishment. The will of the 'second' Socrates is not really 'his own', because the only will which really is Socrates' own is the one he expressed responsibly beforehand, depriving himself of the possibility of expressing a contrary one afterwards.

[39] ibid I, 2, 133.

punishing him if he violates them is comparable to Ulysses, who, as *Emile* reminds us, asks his companions to chain him to the mast to force him to resist the sirens' song.[40] The laws, like the bonds that hold Ulysses back from disaster, are of course forms of coercion, but of a legitimate coercion precisely because freely chosen.[41] The social compact, Rousseau says in a rightly famous chapter, 'tacitly includes the following engagement, which alone can give force to the others: that whoever refuses to obey the general will shall be constrained to do so by the entire body; which means only that he will be forced to be free'[42]. This reconciliation of freedom with force and coercion shows that, beyond any logic of sacrifice, punishment can and must be understood for Rousseau as a guarantee of the security and freedom of all citizens, including the evildoer: both punishment and the threat of it can enforce obedience, yet just like the chains that prevent Ulysses from yielding to the charm of the sirens, this coercion is not a denial, but a realization of freedom.[43]

Rousseau's argument is certainly disconcerting, at least for the contemporary reader. Just think of the following famous statement: 'When the Prince has said

[40] J.J. Rousseau, *Émile, ou de l'éducation*, texte établi par C. Wirz et annoté par Burgelin, in Id, *Euvres complètes* n 1 above, IV, 1969, English translation *Emile or on Education*, introduction, translation and notes by A. Bloom (New York: Basic Books, 1979), 326.

[41] It should be remembered here that the famous incipit of the *Contract* does not merely state that 'man was born free, and is everywhere in chains' (J.J. Rousseau, n 1 above, I, 1, 131), but also identifies the purpose of the work as explaining what can make *this* change legitimate. In a sense, the theme of punishment is at the heart of the entire work, because punishment is precisely the form by which chains are transformed from a symbol of slavery to an instrument of freedom.

[42] ibid I, 7, 141. J.F. Spitz, 'Rousseau et la tradition révolutionnaire française: une énigme pour les républicains' *Les Études Philosophiques*, 445-461 (2007), proposes an alternative reading of this constraint to be free and its relationship with punishment, aimed at safeguarding Rousseaus's discourse from the totalitarian implications attributed to it by the 20th century liberal tradition since Isaiah Berlin. On Spitz' view, conceiving the freedom to which one can be forced through punishment as the freedom that would result from adherence to the general will conceived as one's true will actually lends itself to liberal criticism. Indeed, Spitz argues that individual freedom should be understood here above all as the condition in which all others have a duty to respect in him the same rights that he respects in them. From this point of view, in the very act of violating the rights of others the individual deprives himself of all freedom, because by violating his duties to others he frees them from their duties to him. Punishment would restore the criminal's freedom because it obliges him to respect the rights of others and, in so doing, gives him back the rights that others are obliged to respect, and so the freedom he was deprived of. This reading is to some extent clarifying, but it does not seem to me an alternative to the one on which freedom is conceived as obedience to a law which is in turn the expression of the general will, and therefore of one's own will. Rights are fixed by law, and the passages Spitz draws on show well that according to Rousseau even the criminal, who puts his particular interest before the general interest, wants to enjoy rights, that is, he wants the law or the general will to be respected.

[43] R. Dérathé, 'Jean-Jacques Rousseau et le progrès des idées humanitaires du XVIe au XVIII siècle' *Revue internationale de la Croix Rouge*, 523-543 (1958), states that freedom, like life, is an essential gift of nature of which no one may be deprived except by law, as a punishment for a crime. It can be said that punishment limits or denies the 'natural freedom' of the condemned criminal, but not his 'civil' or 'moral' freedom, which, if anything, it helps to achieve. In this sense, the penalty is an essential element in the transition, described in Chapter VIII of Book I, from the state of nature to civil state: from instinct or physical drive to morality and justice.

to him: 'it is expedient for the State that you should die', as in the case of the death penalty given to the murderer, then the citizen simply 'ought to die' because 'his life is no longer merely a favor of nature, but a conditional gift of the State'.[44] The concept of 'total alienation' allows the *Contract* to transform that consideration of life as a conditional gift of the sovereign, which the second *Discourse* still denounced as an unfounded claim of despotism,[45] into an established truth. But the ruthless face of justice must not make us forget the theoretical framework that gives meaning to punishment in general and circumscribes its legitimacy. In the first place, the legitimacy of punishment is bound to the criterion of public utility and subordination to the general interest. 'The Sovereign', in fact, 'cannot impose on the subjects any burden that is useless to the community. It cannot even will to do so'.[46] In this sense Rousseau immediately corrects the provocative formulation according to which the citizen must die whenever the State wants, and affirms that 'one only has the right to put to death, even as an example, someone who cannot be preserved without danger'.[47] Bearing in mind that 'there is no such thing as a wicked man who could not be made good for something',[48] it is perhaps not so absurd to say that the logic of the argument inclines Rousseau towards a view on which the re-educational value of punishment is in fact combined with a rejection of the death penalty. With the statement that 'in a well-governed State there are few punishments' and 'frequent (…) punishment' is therefore 'always a sign of weakness or laziness in the Government', he seems to favor a process of prevention and decriminalization.[49] Secondly, the legitimacy of punishment is subordinate to an apparently paradoxical criterion of self-determination or legality. The retribution of any crime with a penalty, including the death penalty, is conceivable only in accordance with a universal law, equal for all, approved by the citizens in view of the protection of their common interest and common freedom.

[44] J.J. Rousseau, *Du contrat social* n 1 above, II, 5, 151.

[45] In the *Discourse* Rousseau admits that it is still 'with wisdom' that the 'advantages of a political constitution' were identified, and that they are therefore 'wise' to understand that it was necessary to sacrifice 'one part of their freedom' for the preservation 'of the other'. See J.J. Rousseau, *Discours sur l'origine et les fondemens de l'inégalité parmi les hommes*, texte établi et annoté par J. Starobinski, in Id, *Euvres complètes*, n 1 above, III, 1964, English translation in Id, *The* Discourses *and Other Political Writings*, edited by V. Gourevitch (Oxford: Oxford University Press, 1997), 173. The partial character of this 'sacrifice' is a consequence of the fact that life and freedom appear in this work as 'essential gifts of nature' which the individual is not allowed to divest himself at will (ibid 179). The theory according to which individuals can alienate these essential rights is the ideological support of despotic power, its claim to be the absolute master of its subjects and all their possessions, a master who 'dispenses justice when he despoils them' and a 'grace' or 'favor' whenever he does not deprive them of life or goods (ibid 178). The challenge of the *Contract* is precisely that of transforming these 'essential gifts of nature' into 'conditional gifts of the State' without contradicting the central assumption of the *Discourse*, that is, without making life and freedom something that the individual can 'dispose of'.

[46] J.J. Rousseau, *Du contrat social* n 1 above, II, 4, 148.

[47] ibid II, 5, 151.

[48] ibid.

[49] ibid II 5, 151-152.

> The social treaty has the preservation of the contracting parties as its end. Whoever wants the end also wants the means and these means are inseparable from some risks, even from some losses,[50]

but the only arbiter of these means is and remains the sovereign people, composed of the totality of citizens, and therefore also of the future criminal: preventive consent and compliance with the law transform punishment into a choice, prisons into the place of the liberation of the prisoners from the heteronomous inclinations that led them to betray in the first place themselves, death into a kind of suicide. Thirdly, the exclusion of any violence as a sanction of a crime other than that provided for by law tends to 'exempt a guilty man'[51] from any possible arbitrariness on the part of the authority in charge of judgement and its execution. We know that executive power is no less essential than legislative power. Just as in man

> every free action has two causes that combine to produce it: one moral, namely the will that determines the act, the other physical, namely the power that executes it,

so in the State one can distinguish between 'force and will (…), the latter under the name *legislative power* and the former under the name of *executive power*', and 'nothing is or should be done without their cooperation'.[52] If 'the legislative power is the heart of the State', 'the principle of political life' in the absence of which the political body is 'dead', the executive constitutes 'its brain, giving movement to all its parts', and doing 'for the public person what the union of the soul and the body does in man'.[53] The essentiality of the executive – and therefore of that form of execution which is realized in the trial and in the administration of justice – does not, however, remove its structural subordination to the legislative. The 'trustees of the executive power', including of the power that today we would call judicial, 'are not the masters of the people, but its officers', and 'the people can establish and depose them when it pleases'.[54] Their power can never be exerted 'except by virtue of status and the laws'.[55] Their will 'is not or should not be anything except the general will or the law'.[56] The assignment and execution of punishment can reconcile the criminal with that general will which is his or her own no less than of all the other members of the State, and

[50] ibid II, 5, 151.
[51] ibid.
[52] ibid III, 1, 167.
[53] ibid III, 11 and III, 1, 188 and 166. On the relationship between legislator and executive, and in particular the metaphors mentioned, see B. Bachofen, 'La notion d'exécution chez Rousseau. Une psychopathologie du corps politique' *Revue Française d'Histoire des Idées Politiques*, 275-298 (2011).
[54] J.J. Rousseau, *Du contrat social* n 1 above, III, 18, 196.
[55] ibid II, 11, 162.
[56] ibid III, 1, 169.

therefore with the whole of which he or she is a part, only to the extent that the executor is reduced to a mere 'officer', 'commissioner', 'employee' of the sovereign people, and the execution to a merely technical event.

As I mentioned, Rousseau's discussion of punishment is not free of difficulties. Let's begin with an anecdote which although apparently insignificant is actually useful to explain some general premises. One day, the anecdote has it, 'certain drunkards of Samos defiled the Tribunal of the Ephors' and 'the following day, a public edict gave the Samians permission to be filthy'. Commenting on this episode, Rousseau states that 'a real punishment would have been less severe than such impunity'.[57] Setting aside of the spirit of provocation typical of Rousseau's prose, what is evoked in this episode is the ghost of a real desire to be punished, a desire to flee public scorn and to reconcile themselves through punishment both with themselves and with the political community. This desire makes a first difficulty visible, because while it expresses that radical conformity of the particular will to the general will which is the very definition of perfect virtue, this conformity depends on various assumptions. The first assumption is of a cognitive order.

> As long as several men together consider themselves to be a single body, they have only a single will, which relates to their common preservation and the general welfare.[58]

To affirm that the will is 'one', however, means to presuppose a clear elision of the plurality of particular wills and of the conflict between them. The conformity also relies on an assumption of a practical order, namely that institutions promote concern for public affairs and the related eclipse of particularistic concerns. In a well-ordered State, in fact, 'public service' is the 'main business of the citizens', and not only do 'public affairs dominate private ones in the minds of the citizens', but more generally 'there is less private business': 'common happiness' provides 'a larger portion of each individual's happiness', and 'the individual has less to seek through private efforts'.[59] At the limit, one can imagine a State in which public affairs are the *sole* concern of citizens and 'common happiness' replaces individual happiness completely. This elimination of particular wills, which no longer have any space outside of public affairs to develop, is reflected in all areas of civil and political life. Thanks to it

> all the mechanisms of the State are vigorous and simple, its maxims are clear and luminous, it has no tangled, contradictory interests; the common good is clearly apparent everywhere, and requires only good

[57] ibid IV, 7, 215.
[58] ibid IV, 1, 198.
[59] ibid III, 15, 191-192.

sense to be perceived.[60]

First, the virtue of citizens as members of the sovereign body makes the very few laws that are needed an expression of general interest and will, and their 'necessity (…) universally seen'.[61] Second, the virtue of citizens as public officials makes it impossible for the judiciary to abuse the laws: in Republican Rome, for example, morals 'made many precautions superfluous that would have been necessary in other times'.[62] Finally, the virtue of the citizens is embodied in the 'love of the law' which makes for a State with 'few criminals' and 'few punishments'.[63]

When he poses the problem of the practical and cognitive presuppositions of a virtue able to reign over all spheres of civil life, Rousseau seems to find the solution in a double *deus ex machina*: the civil religion and the Legislator. First of all, it is the civil religion that is called to make 'each citizen (…) love his duties', and to promote with its cult 'the love of laws (…), making the fatherland the object of citizens' adoration' and teaching that 'to die for one's country is to be martyred, to violate the laws to be impious'.[64] Thus, it is religion that can dispose the citizen to that acceptance of the risk of death, if not of certain death, which was at the centre of the discussion of the right to punish. It is not by chance, moreover, that the creation of religions appears in the *Contract*, like customs and opinions, as a 'part unknown to our political thinkers, but on which the success of all the others depends', and 'to which the great legislator', like Moses or Mohammed, 'attends in secret'.[65] The task of the semi-divine figure of the legislator, as is known, is to lay the very foundations of civil life:

> one who dares to undertake the founding of a people should feel that he is capable of changing human nature, so to speak; of transforming each individual (…) into a part of a larger whole from which this individual receives, in a sense, his life and his being; (...) of substituting a partial and moral existence for the physical and independent existence that we have all received from nature;

of obliging individuals 'to conform the will to reason and (the public) to know what it wants'.[66] This idyll, in which we better understand what Rousseau meant when he redefined life as a 'conditional gift of the State' and he substituted the State for God, makes manifest an internal tension in his theory of punishment which we must now face. Punishment presupposes not only the divergence of the will of individuals from the general will expressed in the law, without which

[60] ibid IV, 1, 198.
[61] ibid.
[62] ibid IV, 6, 213.
[63] ibid IV, 8 and II, 5, 219 and 152.
[64] ibid IV, 8, 222, 219-220.
[65] ibid II, 12, 165.
[66] ibid II, 7 and II, 6, 154-155.

there would be no crime, but also the conformity of the former with the latter, without which the law could not be just, the executive power would be autonomous by usurping the legislative one, and the evildoer would not be ready to accept the punishment as an act of justice. This conformity of the particular will to the general will and the 'social spirit' in which it is embodied are, however, at once the 'cause' and the 'effect', the premise and the 'result' of the institution, of the patient and invisible work of the legislator.[67] But if this is right, then what can we say about punishment in imperfect States, where the work of building a free and virtuous people has not been completed and the prerequisites of civil life have not been adequately laid down? What is the rationale of punishment when citizens present particular wills which differ from the general will, do not recognize laws as an expression of their will and do not love them, and therefore do not wish to be punished when they violate them? It is in the light of this question that we can fully grasp a seemingly dissonant element of Rousseau's discourse.

Consider the following passages. The first is taken from a chapter we have already encountered:

> Every offender who attacks the social right becomes through his crimes a rebel and a traitor to his fatherland; he ceases to be one of its members by violating its laws and he even wages war against it. Then, the State's preservation is incompatible with his own, so one of the two must perish; and when the guilty man is put to death, it is less as a citizen than as an enemy. The proceedings and judgment are the proofs and declaration that he has broken the social treaty, and consequently is no longer a member of the State. Now, as he had acknowledged himself to be such, at the very least by his residence, he ought to be removed from it by exile as a violator of the compact or by death as a public enemy. For such an enemy is not a moral person but a man, and in this case the right of war is to kill the vanquished.[68]

The second passage is taken from the final chapter of the whole work dedicated to civil religion, whose importance we have already seen. After arguing that there is 'a purely civil profession of faith, the articles of which are for the Sovereign to establish, not exactly as religious dogmas, but as sentiments of sociability without which it is impossible to be a good citizen or a faithful subject', Rousseau continues in these terms:

> Without being able to obligate anyone to believe them, the sovereign can banish from the State anyone who does not believe them. The sovereign can banish him not for being impious, but for being unsociable; for being incapable of sincerely loving the laws, justice, and of giving his life, if need

[67] ibid II, 7, 156.
[68] ibid II, 5, 151.

> be, for his duty. If someone who has publicly acknowledged these same dogmas behaves as if he does not believe them, he should be punished with death. He has committed the greatest of crimes: he lied before the law.[69]

These passages present several noteworthy elements. Consistent with the idea that the citizen does not need 'guarantees' against the State, the sovereign's judgment is not based on evidence but is itself evidence. Moreover, it seems here that not only murder but every crime must be punished by death, as every evildoer assaults 'social law' or, if one prefers, acts 'as if he does not (…) believe' those articles of faith without close adherence to which good citizens and the maintenance of the social pact would be impossible. To act against the law means in any case to have contravened the covenant and 'lied against the law', and the fact that this lie constitutes 'the greatest of all crimes' nullifies any distinction between one crime and another, and so any proportionality of punishment. What is required of the citizen, after all, is not simply to observe but to 'sincerely love' law and justice, to be willing no longer simply to risk but to 'sacrifice' one's own interests and even one's life at the altar of the general interest. Exteriority of conduct is a consequence of interiority of conviction and affections, and the reference to sincerity, not by chance, calls into question the 'heart of the citizens' on which is engraved that fourth type of law, consisting of opinions and customs, which is the legislator's main object in his secret dealings with religion: the type of law which makes

> the genuine constitution of the State; (…) which, when other laws age or die out, revives or replaces them, preserves a people in the spirit of its institution.[70]

Beyond these elements there is a particularly problematic knot: the idea that the evildoer would be punished not as a citizen, but as a 'public enemy' and therefore in the name of the right of war. The problematic nature of this thesis lies less in its apparent conflict with Rousseau's theory of war, previously defined as a relationship that can only happen between States,[71] than in the friction between

[69] ibid IV, 8, 222-223.

[70] ibid II, 12, 164.

[71] This first difficulty was noted by R. Dérathé in J.J. Rousseau, *Oeuvres complètes* n 1 above, III, 1460 (n 377), and following him by all those who have dealt with the subject of punishment in Rousseau. Though she accepts the majority reading, which sees a 'strident contradiction' between the foundation of the right to punish on the right of war and the conception of war as war between States, G. Silvestrini, n 11 above, 138, notes how the limitations of the right of war introduced in the *Contract* echo in Rousseau's discussion of the right to life or death. This observation places the Silvestrini in the paradoxical situation of having to implicitly admit that Rousseau, in dealing with the death penalty, was well aware of what he had already said about the right of war but ended up forgetting precisely the decisive point of that discussion. This paradox is not insurmountable, however, and it is precisely by overcoming it that the contradiction so widely denounced by readers can be resolved. In one context, Rousseau states that when the prince finds himself making a

this interpretation of the 'offender' as an enemy and the rest of his theory.[72] So far we have seen how the legitimacy of punishment lies in its ability to protect the interests and freedoms not only of the injured parties but of the evildoer himself, reconciling him with that general will which is also his own and with the totality of which he is, as a citizen, an indivisible part. If the relationship between the evildoer and the State is one of war, however, then political power comes perilously close to the master-slave relationship, because in such a case punishment becomes mere force, a coercion to which it may be necessary, but

particularistic use 'of the public force of the State' there are, 'so to speak, two sovereigns, one by right, the other in fact' and instantly 'social union (vanishes) and the political body (is) dissolved' (J.J. Rousseau, n 1 above, III, 1, 169). In another context we read that 'wherever the clergy constitutes a body, (...) it is master and legislator in its domain. There are, therefore, two powers, two sovereigns' (CS IV, 8, 218). As we shall see better shortly, that of the prince who exceeds the limits set by law and that of religious sedition are the main cases of political crime taken into consideration by Rousseau. What is interesting, here, is that this reference to the 'two sovereigns', which re-actualizes the Hobbesian reading of the civil war as the effect of factions operating each as a 'State within the State', Rousseau can consider the conflict opened by political crime, with all its specificities, as a real war, which unlike duels consists in opposition not between individuals but between collective entities. The contradiction that is usually attributed to Rousseau is thus overcome. As a member of the State, in fact, the political criminal continues to be a citizen, but as long as he takes up arms against the State, he is for all intents and purposes a 'public enemy'. Moreover, 'the act of declaring war' (CS II, 2, 146) is not an act of sovereignty, because it is not a law, but is the responsibility of the executive power in accordance with the law, and for precisely this reason, the death penalty falls under the limits imposed by the laws and – as Silvestrini pointed out – to those internal to the logic of the law of war. This overcoming seems to me to find a strong confirmation in the case of the repression of the Catilina's sedition by Cicero. First of all, it should be taken into account that the conclusion of CS, II, 5 on grace in general and its exercise in Republican Rome suggests that in speaking of the death penalty Rousseau had in mind precisely the case of Catiline, in which the theme of grace is essential in a twofold sense. On the one hand, Cicero is rightly condemned because he would have deprived the conspirators of their right, guaranteed by the law, to the *provocatio ad populum*, to request the grace; on the other hand Cicero's salvation is the only case of grace to which Rousseau makes explicit reference. Now, it is true that the conspirators are political criminals, that the salvation of the homeland depends on their repression, and that they are therefore condemned to death as public enemies. But it is also true that Cicero's just condemnation can be read in light of the limits which, as we have seen, are placed on punishment not only by the principle of legality, which prevents penalties from being imposed against the law, but also by the subordination of the right to punish to the internal logic of the right to war. If it is true that it is possible to win a war 'without killing even one' of the enemies, as Rousseau's discussion of the right of war hypothesizes, it is also true that 'one only has the right to put to death, even as an example, someone who cannot be preserved without danger', as outlined in the chapter on the right to life or death, and Rousseau suggests that Cicero could have spared the 'blood of citizens' (CS II, 5 and IV, 151 and 214). On Rousseau's treatment of the right to war, in relation to that of previous or contemporary theorists such as Suarez, Vitoria or Vattel, see R. Dérathé, n 43 above. On the tensions between this rootedness of the right to punish in the law of war and the previous treatment of this same right, see S. Labrusse, 'Le droit de vie et de mort selon J.J. Rousseau ou la politique de l'homme infaillible' *Annales J.-J. Rousseau*, 122-123 (2001). On Hobbes' conception of political factions as a 'State within a State', see F. Toto, 'Fazioni e sedizioni. Aspetti della teoria hobbesiana dei sistemi' *Studi Filosofici*, 49-70 (2018).

[72] The tension between the effort to reintegrate the criminal into civil society and his punishment as an external enemy of that society has been observed several times. See eg S. Labrusse, n 71 above, 127-128.

never obligatory, to submit. Unlike the citizen, in fact, the enemy is

> outside of the State,

and what is expressed in the law and manifested in the judgment is 'a will that is foreign for (him)', and not that general will which is his own will:

> 'a relationship between the whole and its parts is formed which makes of them two separate beings, one of which is the part and the other is the whole minus that part. But the whole minus a part is not the whole, and for as long as this relationship lasts, there is no whole, but rather two unequal parts'.[73]

By grounding the right to punish in the right of war, do we not risk depriving punishment of all its essential features and thus eliminating the distinction between what should be an act of justice and naked violence?[74] Furthermore, and bearing in mind that power relations are not moral relations and exclude any obligation or duty, do we not risk by authorizing the killing of the criminal in the name of the right of war – the right of the strongest – the concomitant authorization of the killing of the State by the offender? Can we account for the punishment of the public enemy consistently with the rest of Rousseau's discourse without denying its importance?[75] In order to answer these questions, it is necessary to more precisely identify the profile of the public enemy that Rousseau had in mind, and to return to the question we previously left open: what happens in imperfect States, when particular wills make their voice heard at various moments of civil life?

The detachment of particular wills from the general will tends first and foremost to affect the legislative process, and with it the very heart of the State. Indeed, when 'the private interests start to make themselves felt', then

> the common interest changes and is faced with opponents, (...) the

[73] J.J. Rousseau, *Du contrat social* n 1 above, II, 6, 152-153.

[74] This passage has produced multiple difficulties for interpreters. Rousseau states that 'quand on fait mourir le coupable, c'est moins comme citoyen que comme ennemi'. In this statement, 'citizen' and 'enemy' can be read both as incompatible properties and as properties that one and the same subject can possess to different degrees. Curiously enough, G. Coqui, 'Le 'droit de vie et de mort' est-il un droit de punir?' (Sur Rousseau, *Du contrat social* II, V)' *Corpus*, 156-176 (2012), states in one place that Rousseau justifies at least certain punishments as 'acts of war' and, in another, that the deatrh penalty is not a punishment because it aims at the suppression of the enemy and does not address the citizen (ibid 166, 172).

[75] Such denial can be found in C. Brettschneider, n 36 above, 61, who rightly highlights the limits that the criminal's rights impose on the State right to punish, but who concludes, in patent contradiction with the text, that 'even those guilty of the worst crimes, for Rousseau, are not regarded as exiles from the social contract; rather, they are still considered citizens within the contract'. Rousseau's reference to the public enemy presupposes precisely this relationship of reciprocal exteriority. On these limitations see also S. Labrusse, n 71 above, 110-112.

> general will is no longer the will of all; contradictions and debates arise and the best advice is not accepted without disputes,

so that 'the social tie begins to slacken and the State grow weak'.[76] At the end of this process, when everyone is 'guided by secret motives' and 'the State close to its ruin continues to subsist only in an illusory and ineffectual form', the general will 'becomes mute', and 'iniquitous decrees whose goal is the private interest are falsely passed under the name of laws'.[77] That is, when the general will ceases to animate the decisions of the majority and the laws 'cease', because the 'decrees' that take their place are like them only in 'name', 'there is no longer any freedom regardless of the side one takes'.[78] In addition to the problems regarding the production of laws, the particularism of interests tends to undermine the transparency of their application. 'Just as the private will acts incessantly against the general will', says Rousseau, 'so the government makes a continuous effort against Sovereignty'.[79] But when the 'dominant will of the Prince' ceases to be 'the general will or the law', and under the pressure of his particular will makes use of the 'public force concentrated in him' to 'produce some absolute and independent act', one ends up with 'two Sovereigns, one in law and the other in fact': 'the bond tying the whole together begins to loosen' and in the long run 'the social union would vanish and the body politic would be dissolved'.[80] Finally, the multiplication of conflicting interests poses a threat to the very tenacity of what we might call public order: 'when the State declines', as happens when the content or execution of laws become unfair, the 'high number of crimes guarantees their impunity'.[81] My hypothesis is that the rootedness of the right to punish in the law of war assumes full significance only within this framework of social collapse – that the coincidence of the punishment with what Hobbes would have called an 'act of hostility' is premised on a context in which the law and judgment are no longer the expression of the general will, the social bond is dissolved, and the union of the political totality gives way to fragmentation.[82] To be sure, this hypothesis faces

[76] J.J. Rousseau, *Du contrat social* n 1 above, IV, 1 198.

[77] ibid IV, 1, 198-199. The problem of these decrees that are law in name only is noted by C. Brooke, n 36 above, but in his review of possible responses to this institutional degeneration (exile, cunning, civil disobedience) he never considers the possibility of a violent and legitimate revolt.

[78] J.J. Rousseau, n 1 above, IV, 2, 201.

[79] ibid III, 10, 186.

[80] ibid III, 1, 169.

[81] ibid II, 5, 152.

[82] One often tends to oppose Rousseau to Hobbes, even on the subject of punishment. The essential divergence between the two would lie in the fact that in Hobbes the rights of individuals precede the covenant and no one can give someone the right to life and death over himself, while in Rousseau the rights derive from a covenant of total alienation. It is worth remembering, however, that although in Chapters XXVII and XXVIII of *Leviathan* the right to punish seems to be entirely based on the right of war, punishment, which proceeds from public authority, is distinguished from an act of hostility. Acts of hostility proceed from usurped power, are inflicted in the absence of a law

major difficulty: in the *Contract*, the assimilation of crime to a reactivation of war, to a laceration of the social fabric, appears to be formulated in the most general terms, and so seems to be valid for every type of crime that may occur both in a failed State and in a well-ordered one.[83] Nevertheless, it is perhaps not insignificant that in its only other occurrence, the concept of 'public enemy' is referred to the prince who prevents the periodic meeting of the popular assemblies, and who, in the very act of doing so, openly declares himself a 'violator of the laws and enemy of the State'.[84] In fact, the full title of Rousseau's work is *The Social Contract, or Principles of Political Law*, and it is entirely consistent with the program explained in the subtitle that after having reviewed the four species of laws Rousseau warns the reader that 'among these various classes, political laws (…) are the only ones relevant to my subject'.[85] Finally, it is no coincidence that the crimes mentioned in the *Contract* are all political crimes. These three features of the text seem to suggest a more precise profile of the enemy, and to mobilize with it a meaning of punishment that is no longer simply legalistic-moralistic, but more harshly political. It seems to suggest, that is, that not every evildoer must be punished as a public enemy, but only the one who becomes in the strict sense 'rebel and traitor of the homeland': he who 'makes war on it' and whose 'preservation is incompatible' with the preservation of the

prohibiting the crime or of a previous public condemnation, do not take into account the possibility of inducing the criminal or others to obey the laws, and exceed the penalties provided for by the law. It is equally curious to note that, although for Hobbes the right to punish coincides with the original *jus in omnia* which the subjects leave to the sovereign, Hobbes alludes in various contexts to the consent of the subject to his own punishment: he affirms that he who performs an action accepts all the known consequences, and that the power to punish, like any other power of the sovereign, was granted to him by the consent of each of his subjects in order to be protected. In this sense, punishment can be imposed as such only to the extent that the subject accepts it. Not recognising the punishment prescribed by law means knowingly and deliberately denying the authority of the State. The right to punish derives from the original right to war, but only those who carry out acts of hostility against the current structure of the State, traitors who in one way or another stains the *crimen lesae majestatis,* are punished as enemies of the State. Uncovering in Hobbes a possible source of Rousseau's reflection on the right to punish and its limitations helps to illuminate the problem of punishment of the 'public enemy': the fact that in Rousseau, as in Hobbes, he is not just any criminal, but one who rejects public authority.

[83] In this regard, it is necessary to note a certain discrepancy between the French text of Rousseau's work and its translations. Rousseau's French states that 'tout malfaiteur attaquant le droit social devient par ses forfaits rebelle et traître à la patrie'. One might be tempted to translate it this way: 'every evildoer, attacking social law, becomes, with his misdeeds, a rebel and a traitor to the homeland'. The same passage, however, could also be translated as follows: 'every evildoer who attacks social law becomes rebel and traitor to the fatherland through his misdeeds'. Unlike the first translation, this translation leaves open the possibility that not 'every evildoer' attacks 'social law' and therefore becomes a 'traitor and rebel' – that public enemies are only those who frontally oppose the social pact,, which for Rousseau is the basis of that 'social order' which is a 'sacred right, serving as a basis for all others', and whose violation leaves everyone their original natural freedom, which coincides with the right to war.

[84] J.J. Rousseau, *Du contrat social* n 1 above, III, 18, 197.

[85] ibid II, 12, 165.

State.[86] In other words, the punishment concerns one whose suppression is for the State a matter of life or death:[87] not the assassin but the usurper, he who silences the general will, promulgating unfair decrees under the name of 'law' and applying them arbitrarily.

In order to prove that what we are inclined to consider a general theory of punishment is better understood as a theory of the repression of political crime, of the violation not of any law but of the political laws that 'constitute the form of government', we can review the examples of political crime scattered in the text and try to grasp the latent thesis that offers a unified key to Rousseau's whole discourse. In a first passage, Rousseau targets the detractors of the people, stating that though one can well smile imagining 'all the nonsense of which a clever swindler or an insinuating talker could persuade the people of Paris or London', one must not forget that

> Cromwell would have been condemned to hard labor by the people of Bern and the Duc de Beaufort sentenced to the reformatory by the Genevans.[88]

With this reference to Cromwell, this first passage, which stages the failure of the ambitious, recalls a second, which instead evokes the success of the skilful impostor and his deception. Rousseau here imagines that, unfortunately, in a hypothetical society of good Christians there appears 'a single ambitious man, a single hypocrite – a Catiline, for example, or a Cromwell', and concludes inexorably that 'these will certainly get the better of his pious compatriots':

> Christian charity makes it hard to think ill of one's neighbor. As soon as he has learned the art of how to trick them through some ruse and seize part of the public authority for himself, he will be a man of constituted dignity; it is God's will to respect him. Soon he is powerful: it is God's will to obey him. Does the depository of this power abuse it? It is the rod with which God punishes his children. It would be against conscience to chase

[86] ibid II, 5, 151.

[87] B. Bernardi, n 2 above, notes with acuity the tension that generated in the *Contract* between two requirements: that of protecting the security and freedom of all citizens, including criminals, and that of preserving the State itself, a requirement that in case of conflict takes precedence over the first to the extent that the preservation of the State constitutes the condition of the possibility of protecting citizens. With the same acuity, he notes that the discussion of the death penalty must be read against the background of this tension: that the criminal is not put to death because of his guilt, but because of the danger he represents for the community. At the same time, Bernardi does not note that although the murderer is the only example explicitly taken into account in CS II, 5, all the points he highlights apply much more clearly to the political criminal, and in particular to the Prince who ceases to confine himself to applying the law and attempts to direct public force against the State itself. This is an eminent case in which the conservation of the criminal or his freedom is incompatible with that of the State.

[88] J.J. Rousseau, *Du contrat social* n 1 above, IV, 1, 198.

> out the usurper: it would be necessary to disturb the public tranquillity, use violence, shed blood. All of that is inconsistent with the gentleness of a Christian. And after all, what is the importance of being free or serf in this vale of tears? The essential thing is to go to heaven, and resignation is but an additional means of doing so.[89]

For its reference on the one hand to Catiline and on the other to Christianity, this passage recalls two others. If Cromwell's controversial figure is brought up only in counterfactual reasonings, as an example of the ambitious against which a well-ordered society such as that of Geneva – but not a society of true Christians – would be endowed with sufficient protection, Catiline also appears in a context in which Rousseau reasons in more historical terms. Rousseau recalls here that a dictator with unlimited authority would have easily thwarted Catiline's conspiracy, but Cicero lacked such authority, and, 'in order to act effectively, (he) was constrained to exceed' the power conferred on him 'in a crucial respect', so that 'the first explosions of joy gave approval to his conduct', but 'later on he was justly called to account for the blood of the citizens spilled against the laws'.[90] Finally, the last citation I would like to consider is that in which Rousseau unmasks Christianity as a 'religion of the priest'.[91] Rousseau comes here almost to approve of the persecutions of the first Christians. The pagans, in fact,

> always regarded the Christians as true rebels who, beneath a hypocritical submissiveness, were only awaiting the moment to become independent and masters, and to usurp adroitly the authority they pretended to respect out of weakness.[92]

Even if Rousseau cautiously attributes this image of Christians to the Romans, strategically distancing himself from it, it is clear that it should not have appeared too far from reality: the philosopher himself, in fact, admits that early Christianity was transformed over the centuries, with the papacy, into 'the most violent despotism in this world'.[93]

The general sense of this game of references becomes clearer if we consider the differential treatment that Rousseau gives to different cases. Even taking into account the more nuanced approach reserved for the case of Cicero, who on the one hand loves his glory 'more than his land' but whose violation of the law, on the other hand, succeeds in 'saving the State',[94] it is clear that all the figures portrayed in these passages are figures of usurpers (or aspiring usurpers). Equally

[89] ibid IV, 8, 221.
[90] J.J. Rousseau, *Du contrat social* n 1 above, IV, 6, 214.
[91] ibid IV, 8, 219.
[92] ibid IV, 8, 217.
[93] ibid IV, 8, 218.
[94] ibid IV, 6, 214.

clear is that these individual figures emerge in contexts of institutional and moral degeneration. Cromwell's England is not only the same England in which kings have only nominally assumed the role of heads of the church, actualizing the nightmare of divided sovereignty, it is also the England in which 'the (…) people (…) think it is free', but 'it greatly deceives itself'.[95] The country is free, in fact, 'only during the election of the members of parliament; as soon as they are elected, it is a slave, it is nothing'.[96] Similarly, late-Republican Rome is not only the Rome in which the division between patricians and plebeians has already generated the 'abuse of the aristocracy'[97] and civil wars, but also the Rome in which customs and virtue begin to lose their vigour, and desires – as witnessed by Cicero's lust for glory – begin to focus on particularistic goals. Finally, Imperial Rome, in which early Christianity exerted all its subversive force, is an exemplary case of the tendency of governments to degenerate: territorial expansion brought about not only the widening of the gap between custom and law, the intensification of repression and the concentration of power in one set of hands, but also to the degeneration of law itself from the legitimate monarchical form that was still proper to Augustus to the despotic form that it acquired with Tiberius. It is precisely within this context of degeneration of the institutional and moral framework that the theme of political crime and the rootedness of its punishment in the right of war assumes all its weight. As is easy to see, the theme of punishment circulates in all the passages we have read. Rousseau believes that Cicero was rightly punished and imagines both Cromwell and the Duke of Beaufort punished as well – one in a prison and the other in a house of correction. In the same way, one can suppose the violence and bloodshed that Rousseau wished to see as a reward for usurpation can be understood as a case of death penalty. It is equally easy, however, to see that all the passages focus on examples of civil war, rebellion or seditious movements. The threat to the institutions can come from above (as in the case of Cromwell, Beaufort, etc) or from below (as in the case of Catiline, false Christians, etc.). The usurper can be content to extend the public authority of which he is already the depositary beyond the limits drawn by the law (as in Cicero's case) or go so far as to act as a kind of punctual reversal of the figure of the 'great Legislator'.[98]

[95] ibid III, 15, 192.

[96] ibid.

[97] ibid III, 10, 187.

[98] A semi-divine figure endowed with an extraordinary intelligence, the Legislator should see all the passions of men without feeling any of them, enjoy a happiness that does not depend on men, but nevertheless desire to take care of them by protecting them from the seductions of particular wills and taking care of religion as an instrument capable of conforming wills to reason. By contrast, the diabolical figure of the usurper is endowed with an extraordinary intelligence, which is not, however, free of the passions, but enslaved to ambition; he does not enjoy any happiness independently of others, and precisely for this reason he must not enlighten them, but deceive them, seduce them using religion as a useful instrument to camouflage his particular interest and enslave everyone to his own whims.

The assault on public authority, finally, can be either unsuccessful (as in the cases of Catiline and of the rebellion of the first Christian communities) or victorious (as perhaps in the case of Cromwell, or of the hypocrite who takes power in an imaginary Christian republic). In the laceration of the social and institutional fabric inflicted by war, in any case, the only punishment that appears as an act of justice is that of Cicero, who continues to recognize himself as a citizen in the not yet completely degenerated context of late Republican Rome. In all other cases the punishment goes beyond the legal framework to take the form of the elimination of an internal enemy. In this sense, the case of the impostor who takes over public authority in an ideal Christian republic is paradigmatic. As long as one remains captured within the theological-political imagination mobilized by the usurper (or the would-be usurper), one can believe that he really manages to conquer public authority, and thus that resistance constitutes a criminal form of disturbing the public peace, and that repression constitutes punishment in the strict sense of the word. In reality, however, the exact opposite occurs. The usurper silences the laws at the very moment he claims to apply them: the relationship between general will and particular will degenerates into a naked relationship of forces, and what is presented as punishment is in reality nothing more than an act of war. Now, it is quite possible that the usurper will meet no resistance, as happens in the case of the republic composed of true Christians. For Rousseau, however, the people who hypothetically evade the despot's abuses and shed his blood (such as those who put Cromwell or the Duke of Beaufort in prison or in a house of correction) would do so rightly. This consideration allows us to see the deeper coherence in Rousseau's apparently ambivalent treatment of Christianity. On the one hand, true Christians are criticized precisely for their indifference in the face of abuse and usurpation of public power, for their resignation in the face of slavery and the misery into which usurpers throw entire peoples, and for the gentleness that prevents them from responding to violence with violence – the violence of prison, that of bloodshed, and if necessary that of persecution. On the other hand, the first Christians are not without reason persecuted as rebels because they wish to seize the power to dictate law and aspire to the establishment of the worst of despotism, representing an exemplary case of the violation of public faith that, as he makes clear in the chapter on civil religion, Rousseau hopes will be punished by death. The greatest of crimes, we have seen, is *lying* before the law, being incapable of *sincerely* loving the laws, but the deference to law of Christians was just a 'hypocritical submission' by which they waited for an opportunity to seize the authority they 'pretended to respect'.

The treatment of this range of different cases finally gives rise to a clear thesis. The case of failed usurpation is the simplest one: while Cicero is tried as a citizen, because his violation of the law does not constitute a real usurpation and does not endanger the life of the State, Catiline and the first Christians can

be put to death as public enemies in the name of the right to war, because by ceasing to recognize themselves in the institutions and to consider themselves as parts of the whole they break the social pact, they stop being members of the State. In short, they come to represent an 'outside' that undermines the State from the inside. The case of successful usurpation is certainly more problematic. Of course, in a well-ordered State Cromwell and Beaufort would be imprisoned. But what happens in a context of institutional degeneration in which Cromwell or the priests succeed, as in fact happened in revolutionary England or Christian Europe? Rousseau carefully avoids getting into the complications posed by concrete examples (although a passage in which he states that very few States have laws and are therefore legitimate political formations allows us to understand what his thinking was).[99] On a more theoretical level, however, the move made by Rousseau is clear and unambiguous, and consists in recasting the apparently successful usurpation as a kind of failed usurpation: the despot and the tyrant can never seize authority, but only exert force. If we remember how 'the act of declaring war' is not an act of sovereignty,

> since each of these acts is not a law but merely an application of the law, a particular act which determines the case of the law,

we can conclude that the death sentence of the political criminal, the criminal who attacks the social rights by becoming 'a rebel and a traitor', and in particular of the despot and the tyrant, is precisely an application of those political laws that can tentatively be labelled 'fundamental laws'.[100] What will resurface, in its potentially terrible form, is therefore that original democracy which is for Rousseau the source of all other power, and in the face of which all constituted powers vanish like a puff of smoke. Popular sovereignty is reaffirmed in its absoluteness through the terrible punishment and the revolt against the spectre of the State, its so-called laws, and the arbitrariness of the powerful. It is true that 'the established government should never be touched until it becomes incompatible with the public good', and that

> it is impossible to be too careful about observing all the requisite formalities, in order to distinguish a regular, legitimate act from a seditious tumult and the will of an entire people from the clamours of one faction.[101]

It is also true, however, that when power becomes 'incompatible with the public good' the formalities disappear, and what the illegitimate power calls a riot is stated by force as a 'regular (and) legitimate act'. Certainly, 'when civil machinery is worn out' seditions can destroy the State without revolutions being able to restore it,

[99] J.J. Rousseau, *Du contrat social* n 1 above, III, 15, 193.
[100] ibid II, 2 and II, 12, 146 and 164.
[101] ibid III, 18, 196.

because 'freedom can be acquired but can never be recovered'[102]. There are, however, 'violent periods' in which

> revolutions have the same effect on peoples as do certain crises on individuals; the horror of the past is equivalent to amnesia, and the State, set afire by civil wars, is so to speak reborn from its ashes and resumes the vigour of youth by escaping from death's clutches.[103]

It is precisely here, against the background of this civil war in which those who would be entrusted with its protection reveal themselves to be 'enemies of the State', that the sense of punishment of the criminal as a 'public enemy' is manifested: a sense that goes beyond that the simply juridical-moral because what Rousseau presents as the restoration of the juridical-moral order broken by crime is not so much the confirmation of the established order, which in reality has plunged into disorder, but the institution of a new order that reduces even the memory of the old one to ashes. In this case it is not a question of guaranteeing the interest and freedom of the criminal, of making him or her good for something, of reconciling him or her with the whole of which he or she is a part, thus recomposing the tear introduced by crime into civil life. 'Everything that destroys social unity', says Rousseau, 'is worth nothing'.[104] Each citizen, in the same way, 'is nothing'[105] outside the bond of solidarity with others. The usurper who stops conceiving himself as 'part of a greater whole' to try to become its master and lord is precisely this very dangerous nothing, in the relationship with which no mediation or reconciliation is possible, but only victory or slavery. That's why he must be punished not as a citizen, but as a public enemy.

[102] ibid II, 8, 158.
[103] ibid.
[104] ibid IV, 8, 219.
[105] ibid II, 7, 155.

How to Punish?
The Deontology of Punishment in the Enlightenment Philosophy

Dario Ippolito*

Abstract

The article faces the penal problem in the Enlightenment philosophy, proposing a three-step approach: 1) detection of the normative principles elaborated in the debate on the right to punish; 2) clarification of the theoretical foundation and the political scope of such principles; 3) examination of the relationship between these principles and different types of penalties.

I. Introduction

The 'Beccaria moment' is now! With these words, Gianni Francioni highlighted the recent flourishment of studies on Beccaria's work and its historical importance.[1] Looking from a different perspective, we could argue that the 'Beccaria moment' has returned. Indeed, in recent decades, discussions of the power to punish have reached the same level of radicalism, critical vigour and reformatory impetus as in the Enlightenment era. Today, as then, the penal problem puzzles us as it imposes itself upon the philosophical culture, the legal science and the civic conscious with the urgency of its intrinsically political set of questions. How can we conciliate liberty and security? How do we prevent crimes without infringing rights? Where do we draw the line between the powers of penal agents, on the one hand, and the faculties that shape the individual's legal sphere, on the other hand? On what principles – and based on which rules – ought we model a procedure aimed at eliminating – to the extent possible – an innocent man's fear of conviction and the guilty man's expectation of impunity? Finally, what sanctions should be imposed on those who violate prohibitions?

Criminal law is today – just as much as it was in Beccaria's times – a field of tensions. Its balance is frail; its physiognomy deformed by contrasting impulses,

* Professor of Philosophy of Law, Legal Logic and Argumentation, Department of Law, Roma Tre University. This article was translated by Matija Žgur.

[1] See G. Francioni's talk at the symposium 'Illuminismi. Attualità e limiti dell'età del Lumi' (Milan, Casa della Cultura, 16 February 2016), available at: https://tinyurl.com/yh7zzuwj (last visited 30 June 2021).

requirements and aims. Its concepts and institutes spark public debate and lead irreconcilable visions of society to clash – the securitarian ideologies vs liberal axiologies, the models of restorative justice vs incitements of revenge, the zero-tolerance policies vs recommendations for *cautela in poenam*.

In the current state of these conflicts, deprivation of freedom through incarceration – the typical punitive mechanism in contemporary systems – emerges as a fundamental and decisive issue. On the one hand, its apologists argue for its necessity, affirm its irreplaceability and advocate for its expansion; its critics, on the other hand, point out its inefficacy, denounce its injustice and foreshadow its overcoming. In short, the prison is today at the centre of a crucial dispute over the civility of law – just as the *supplice* was in the second half of the Eighteenth-century, when the *custodes iuris* and the *sacerdotes iustitiae* rallied in defence of the traditional criminal order against the reformatory demands of the Enlightenment movement.

In my view, this historical parallel renders the analysis of the Enlightenment thinkers' juridico-political discourses on the topic of penal sanctions especially interesting. In particular, it allows us to reflect upon the way in which they were able to discredit the dogmas of the dominant culture, contributing in this way to the extinguishing of the 'splendour of the supplice'. It seems useful to begin this investigation by analysing the original deontology of the punishment developed during the 'Beccaria moment'. A deontology that is complex, variegated and irreducible to a uniform paradigm, but which is – in all of its normative declensions – nevertheless polemically aimed at delegitimizing the punitive system of the time. Not, therefore, a deontology, but rather a set of doctrines on just punishment, characterized by a plurality of principles aimed at limiting and constraining the power to decide on how to punish.

Which are these principles? What are their doctrinal foundations and what their pragmatic function? Although highly relevant, these questions have not attracted the appropriate attention in penal Enlightenment studies. In principle, we know the Enlightenment philosophers' answers to the question 'what to prohibit?': their struggle for the secularization of criminal law and their defence of individual freedoms against the despotic prohibitions are well known even beyond the borders of specialist studies. We also have a fairly clear understanding of their theses on 'how to adjudicate': in the antithesis of the inquisitorial model, they directed the criminal trial towards the protection of innocence and the search for truth by way of a procedure based on the principles of publicity and orality, equality and contradictoriness between the parties, as well as the impartiality of the judge. However, what do we know about their doctrines on penal sanction? What types of punishment did they accept as legitimate? Within what limits did they restrict the power of the State to inflict harm upon those who violated some legal prohibition? In short, how did they respond to the question 'how to punish'?

II. How to Punish?

1. The Enlightenment Thinkers' Reply According to Michel Foucault

Someone might object that we actually know the answer to this question as well. After all, Foucault explained it more than four decades ago in *Discipline and Punish*, on the pages dedicated to the new philosophy of punishment that was diffused in the second part of the Eighteenth century. The objection appears founded. Therefore, it seems worthwhile to return to Foucault's influential lesson, briefly illustrate its interpretative structure and evaluate the issues pertaining to the topic of our interest.

According to Foucault's reading, the Eighteenth-century penal reformism is a cultural expression of the disciplinary society. It promotes the development of a new policy of combating delinquency, one based on the criteria of preventive efficiency and repressive capacity. Foucault invites us to consider the reformers' discourse beyond their humanitarian rhetoric. Their goal was 'not so much to establish a new right to punish based on more equitable principles',[2] but rather to extend and strengthen the grip of punitive power on the social body, making it more effective, orderly, precise, pervasive and less costly.

> (T)o make of the punishment and repression of illegalities a regular function, coextensive with society; not to punish less, but to punish better; to punish with an attenuated severity perhaps, but in order to punish with more universality

– these are, argues Foucault, the 'primary objectives'[3] of the theorists of penal modernization – the strategists and architects of the 'generalized punishment'.[4]

Within the framework of this original representation of reformist ideology, the problem of the penal sanction stands out. Underlying the 'technology of the power to punish'[5] is a sophisticated 'semio-technique'[6] of the *modus puniendi*. Examining its normative scope, Foucault presents its 'major rules'.[7] The first is the rule of minimum quantity, according to which the harm threatened by the criminal law must be the minimum necessary to dissuade from the criminal act. The second is the rule of sufficient ideality, on the basis of which the legislator must pursue the goal of the effectiveness of prohibitions, relying more on 'the 'pain' of the idea of 'pain'' than on its 'corporal reality'.[8] The third is the rule of lateral effects, for which the sanction associated with the infringement must

[2] M. Foucault, *Discipline and Punish* (Middlesex: Penguin Books, 1991), 80.
[3] ibid 82.
[4] This is the meaningful title of the first chapter of the second part of M. Foucault, *Discipline and Punish* n 2 above.
[5] ibid 89.
[6] ibid 94.
[7] ibid 94.
[8] ibid 94-95.

aim at general deterrence. The fourth is the rule of perfect certainty, which prescribes a clear legal determination of punishments and requires that they be meticulously imposed upon those guilty of respective crimes. The fifth is the rule of common truth, which – regarding the discipline of evidence in the criminal trial – goes beyond the thematic perimeter of our discussion. The sixth is the rule of optimal specification, which requires (in addition to the codification of the types of crimes) the 'individualization of sentences' on the basis of 'particular characteristics of each criminal'.[9]

To overcome the confusion due to this unusual terminology and to provide a clearer picture of this new 'economy of the power to punish'[10] Foucault speaks of, let us attempt to order the elements of this catalogue. On the basis of its principles, (1) the punishment must be (1a) established by law and (1b) endowed with intimidating force; (2) the legislator must pursue the goal of having his prohibitive norms respected, not by increasing the punishment's cruelty, but through (2a) the maximization of their capacity of representing harm, (2b) the optimization of punitive reactions and (2c) the minimization of the area of impunity. Finally, (3) the punishment must be adjusted to the personality of the offender.

Before considering further features of 'the punitive city'[11] designed by the reformers, let us reflect on the characteristics of this layout. Certain features of the deontologies of the punishment inspired by Beccaria can undoubtedly be found therein. Indeed, with the international spread of the ideas contained in *On Crimes and Punishments*, the principles of legality in criminal law and of punitive economy had become the *tòpoi* of the debate on the right to punish. The same can be said of the thesis according to which the purpose of general prevention – ie the dissuasion of other members of society from committing crimes – is not achieved by increasing the severity of sanctions, but rather by increasing the effectiveness of sanctioning norms. In other terms, by increasing the efficiency of criminal justice.

On the contrary, the requirement of the individualization of the punishment appears utterly out of place in this cultural framework. Moreover, on a closer look, the heuristic procedure followed by Foucault in framing the rule of the optimal specification results somewhat distorted. The said requirement is related to two doctrinal positions: a) the idea that the law must establish punishments on the basis of subjective statuses, considering that the same threat of harm will not have the same deterring effect on individuals in different positions (for example, a shaming punishment may scare a nobleman, but not a commoner); b) the idea according to which in sanctioning the same kind of crime, the judge must impose punishments of different severity in view of the

[9] ibid 99.
[10] ibid 99.
[11] ibid 113.

different degree of guilt of the offender (eg one who steals due to hunger does not deserve the same punishment as one who steals out of greed).

It would be misleading to invoke these two claims in order to affirm that besides the requirement for a 'codification of the offences-punishments system'[12] there also appears that of the 'modulation of the penalty' in relation to the 'defendant himself, to his nature, to his way of life and his attitude of mind, to his past'.[13] The thesis *sub* a) does not relate the type of punishment to the individual identity of the offender, but rather to his social identity. It is rather a utilitarian justification of penal inequality in the context of a society of orders and not the germ of the doctrine of the individualization of punishment. Likewise, the thesis *sub* b) does not concern a 'precisely adapted code', in which the taxonomy of crimes and punishments is complemented by the 'modulation of the criminal-punishment'.[14] As clearly emerges from Marat's writings on which Foucault underpins his discussion, this thesis rather establishes a principle of equitable justice, related to the judicial conception of crimes:

> the judge must always be mindful of the circumstances in which the culprit is placed; depending on these circumstances, a crime can be more or less serious.[15]

It should also be noted that neither of the two theses can be considered fully representative of the reformist ideas circulating in the age of Enlightenment. The first is quite clearly a conservative thesis that endorses a typical feature of the *ancien régime*'s law, namely the importance of social differences in the legal regulation of social relations and of individual behaviour. This does not mean that it did not find support even among the exponents of the Enlightenment movement. However, it is precisely in the Enlightenment debate that its anti-thesis is affirmed – namely, the principle of equality before the criminal law. '(T)he punishments', writes Beccaria, 'ought to be the same for the highest as they are for the lowest of citizens'.[16] Pierre Louis de Lacretelle – whose *Discours sur le préjugé des peines infamantes* Foucault invokes precisely in relation to the connection between status differences and criminal laws – takes a resolute position with regard to the same: 'Where citizens are not equal before the

[12] ibid 99.
[13] ibid 99.
[14] ibid 99.
[15] J.-P. Marat, *Plan de la législation criminelle* (Paris: chez Rochette, 1790), 34. On this Marat's work, see J. Llobet Rodríguez, 'Jean-Paul Marat y la Ilustración penal' *Revista CENIPEC*, 275-306 (2005); and the book by L. Prieto Sanchís, *La filosofía penal de la Ilustración* (Lima: Palestra, 2007), 163-174.
[16] C. Beccaria, *On Crimes and Punishments* (Cambridge: Cambridge University Press, 1995), 51. On Beccaria's philosophy of punishment, see the fundamental studies by G. Francioni, 'Beccaria filosofo utilitarista', in S. Romagnoli e G.D. Pisapia eds, *Cesare Beccaria tra Milano e l'Europa* (Milano-Bari: Cariplo-Laterza, 1990), 69-87; and Ph. Audegean, *La philosophie de Beccaria. Savoir punir, savoir écrire, savoir produire* (Paris: Vrin, 2010).

criminal law, there can be neither safety nor happiness'.[17]

Although the meta-judicial principle *sub* b) expresses a widespread sensibility in the Enlightenment literature, it nevertheless conflicts – *de iure condendo* – with other contemporary issues which are far from negligible. These include the Beccarean doctrine which attaches the gravity of the penalty to the gravity of the social harm caused by the crime, excluding from consideration the subjective element of guilt; or the principle of the strict subjection of the judge to the law which – strengthened by the general preference of the reformers for the establishment of jury trials – was laid down in the first penal code of revolutionary France.[18]

These latter observations highlight what, in my view, is the greatest limitation of Foucault's analysis. By disregarding the dialectic polyphony of the Enlightenment movement, Foucault furnishes a unitary, compact and consistent image of the Eighteenth-century reform movement. His hermeneutical proposal, according to which '(t)he reform of criminal law must be read as a strategy for the rearrangement of the power to punish'[19] is translated into the construction of an ideological canon – a canon obtained by the selection and combination of principles held to be useful for the goal of 'regularizing, refining, universalizing the art of punishment'.[20]

Moreover, this mixture of principles is not always successful. Even when it does succeed, it produces uniformity at the price of distorting reality. An example for each of the two aporias can be useful for the further development of our discussion.

a) Incommixturable ingredients. Foucault includes two antinomic theses within the canon of the reformist thought. He expresses the first on the basis of Gaetano Filangieri's words:

> The proportion between the penalty and the quality of the offence is determined by the influence that the violation of the pact has on the social order'. But this influence of a crime is not necessarily in direct proportion to its horror; a crime that horrifies the conscience is often of less effect than an offence that everyone tolerates and feels quite ready to imitate. There is a scarcity of great crimes; on the other hand, there is the danger that everyday offences may multiply. So one must not seek a qualitative relation between the crime and its punishment (…) One must calculate a penalty in

[17] L. de Lacretelle, *Discours sur le préjugé des peines infamantes* (Paris: Cuchet, 1784), 143-144.

[18] Cf L.M. Le Peletier, 'Rapport sur le projet de code pénal fait au nom des comités de Constitution et de législation criminelle' *Archives parlementaires*, t. XXVI (the session of Monday, 23 May 1791), 322.

[19] M. Foucault, n 2 above, 80.

[20] ibid 89.

terms not of the crime, but of its possible repetition.[21]

Apart from the surprising discrepancy between Filangieri's thesis (according to which

> the crime that violates a pact more relevant to the social order must receive (…) a more severe punishment than the crime that violates a less relevant pact)[22]

and Foucault's interpretation of it (according to which the gravity of the penalty must be commensurate with the influence of the crime – or rather, to the imitative effect it produces), it is interesting to highlight the conclusion of the latter's reasoning: the rejection – attributed to the promotors of the new *ars puniendi* – of the principle of qualitative correspondence between the crime and the punishment.

I emphasize this conclusion because in the continuation of Foucault's examination this principle is rather included among the maxims of the new *ars puniendi*. 'To derive the offence from the punishment is the best means of proportioning punishment to crime',[23] writes Jean-Paul Marat echoing Montesquieu. 'Exact relations are required between the nature of the offence and the nature of the punishment',[24] affirms Louis-Michel le Peletier before the *Assemblée nationale*. Foucault summarizes: 'The punishment must proceed from the crime'.[25]

Whereas the idea that we should 'calculate a penalty in terms not of the crime, but of its possible repetition' has no place in the *République des Lumières*, both the affirmation of the principle of typological correspondence between the punishment and the crime as well as its rejection, are present in the debate on the right to punish. If the rejection of that principle is very rare, its affirmation in widely shared: it is easy to find it both in the natural law doctrine and in the psychological conjectures of utilitarians. This confirms the inadequacy of a homogenizing representation of the Enlightenment.

b) Distortion as the price of uniformity. In the 'techno-politics of punishment'[26] – to which Foucault reduces the science of penal legislation fostered by the reformers – there is no place for values different from those

[21] ibid 92-93.

[22] G. Filangieri, *La Scienza della legislazione*, in V. Ferrone ed (Venezia: Centro di Studi sull'Illuminismo europeo 'G. Stiffoni', 2009), Book III, Part II [1783], XXXIX, 141. Indispensable with regard to Filangieri's criminal doctrine are F. Berti's, *La ragione prudente. Gaetano Filangieri e la religione delle riforme* (Firenze: Centro Editoriale Toscano, 2003) 473-502; and 'Diritto penale e diritti dell'uomo: il garantismo di Gaetano Filangieri', in D. Ippolito ed, *La libertà attraverso il diritto. Illuminismo giuridico e questione penale* (Napoli: Editoriale scientifica, 2014), 115-147.

[23] Quoted in M. Foucault, n 2 above, 105.

[24] ibid 105.

[25] ibid 106.

[26] ibid 92.

inherent in an orderly government of society: efficiency, legality, discipline, obedience etc. Once established that the goal of the reforms is 'to insert the power to punish more deeply into the social body',[27] any axiological criterion that is *prima facie* foreign to the logic of this governmental project must be cleansed of its rhetorical cosmetics.

Hence, the Enlightenment struggle against inhumane punishment is interpreted by Foucault in a purely utilitarian perspective and explained in function of the interests of the one exercising social domination:

> The pain that must exclude any reduction in punishment is that felt by the judges or spectators with all the hardness of heart that it may bring with it, all the ferocity induced by familiarity, or on the contrary, ill-founded feelings of pity and indulgence (…) What has to be arranged and calculated are the return effects of punishment on the punishing authority and the power that it claims to exercise.[28]

In short, cruel punishments are counterproductive due to the effects they generate – that is, they are counterproductive for the administrators of the public order. 'It is this 'economic' rationality that must calculate the penalty and prescribe the appropriate techniques'.[29] It is here that – due to the necessity to regulate the exercise of power in order to ensure its regulative capacity – the principle of the humanisation of punishment is rooted: ' 'Humanity' is the respectable name given to this economy and to its meticulous calculations'.[30]

Demystified and re-semanticised, the evaluative reference to humanity is thus reabsorbed into the ideology of punitive optimization. Punishing cruelly is not worthwhile as it accustomises to violence and produces connivance. As reform theorists have repeated time and again, cruel punishments are harmful and criminogenic. However, we ought to ask ourselves whether they keep repeating this to increase the performances of the penal enterprise ('to punish better')[31] or to support a demand for justice (the mitigation of punishments)?

In Foucault's interpretation, he who says 'humanity' is a skilful manager of the 'theatre of punishments'.[32] However, it seems to me that to look at the reformist discourse as a sort of *speculum principis* is to obliterate the whole political discourse – *ex parte civium* – which goes from Beccaria – according to whom 'a society cannot be called legitimate where it is not an unfailing principle that men should be subjected to the fewest possible ills'[33] – to Kant, whose retributivist rigor, at the moment he justifies the capital punishment, requires

[27] ibid 82.
[28] ibid 91.
[29] ibid 92.
[30] ibid 92.
[31] ibid 82
[32] ibid 106.
[33] C. Beccaria, n 16 above, 48.

that its execution 'be freed from any mistreatment that could make the humanity in the person suffering it into something abominable'.[34] The humanity, that is, of the one found guilty, not of his judges; of the one suffering the punishment, not of those assisting in the penal spectacle.

These principles of the deontology of the punishment are not considered in the ideological canon constructed by Foucault. Instead, what finds its place therein is the whole penal arsenal advocated by François-Michel Vermeil in his *Essai sur les réformes à faire dans notre législation criminelle*, which is a typical expression of the new 'technique of punitive signs':[35] from humiliation for crimes of pride, the stakes for arson, all the way to the punishment devised for parricide:

> blinded, locked in a suspended cage, naked (…), exposed to all the rigors of the seasons, (…) covered in snow, (…) burned by the sun, he would be fed on bread and water until the end of his life.[36]

Putting in the same ideological framework the creator of this 'torment', expressly conceived as a 'prolongation of a painful death',[37] and writers like Filangieri and Marat is an inadmissible operation. Not all reform proposals are aimed at the same goals. In order to understand, we must first distinguish. Thus, if we wish to know the Enlightenment thinkers' reply to the question 'how to punish', we cannot be satisfied with Foucault's interpretative paradigm.[38]

2. The Enlightenment Thinkers' Reply According to Tarello

Even after more than forty years following its publication, Giovanni Tarello's *Storia della cultura giuridica moderna* continues to tower above other studies on the juridico-political doctrines of Enlightenment thinkers.[39] By combining an extraordinary knowledge of the sources with interpretative rigour and conceptual clarity of analytical legal philosophy, the founder of the so-called 'Genovese school' was able to shed light on the reasons for the emergence of the Eighteenth-century 'penal problem' as well as on the multiplicity of its aspects. Among the series of questions in which the penal problem is articulated, the question of punitive sanctions occupies a prominent position alongside the problem of the qualification of crimes. According to Tarello, the normative theses corresponding to these questions stem from three main doctrines of punishment: namely, utilitarianism, humanitarian ideologies and the

[34] I. Kant, *Metaphysics of Morals* (Cambridge: Cambridge University Press, 1991), 142.

[35] M. Foucault, n 2 above, 94.

[36] F.-M. Vermeil, *Essai sur les réformes à faire dans notre législation criminelle* (Paris: Savoye et Delalain, 1781), 148-149.

[37] ibid 149.

[38] For an acute critique of the foucaultian paradigm, see A. Punzi, 'Sorvegliare per non punire? Note su sicurezza e garanzie' *Rivista internazionale di filosofia del diritto,* 621-636 (2014).

[39] G. Tarello, *Storia della cultura giuridica moderna* (Bologna, il Mulino, 1997).

proportionalist ideology. While the former primarily concerns prohibitions, the latter two refer directly to the punishment.

'Humanitarian' is the adjective attributed to the ideology of the mildness of punishments. As Tarello remarks, 'it appears to derive, in part, from the utilitarian ideology'.[40] Indeed, the latter prescribes to the sovereign not to 'impose punishments more severe than required by utility'[41] as well as to abolish punishments that are useless by nature. 'Proportionalist', on the other hand, is the name ascribed to the ideology according to which 'the penalty must be commensurate (…) with the crime'.[42] Tarello sees its origins in two distinct traditions of thought: one – philosophical – dating back to the 'Pythagorean doctrines of retribution'[43]; the other – religious – rooted in the Old Testament. In other terms, we are dealing with a 'retributivist conception of punishment', whose Enlightenment renewal is related to the 'Eighteenth-century legal rationalism'. Accordingly, 'if the punishment is the (exact) retribution',[44] then for each crime its measure has to be (exactly) established by the legislator.

With regard to the legitimacy of prohibitions, Tarello demonstrates how both utilitarianism and retributivism advocate for the exclusion

> of matters regarding the conscience and religion, from the sphere of issues worthy of penal sanction by the secular sovereign.[45]

On the one hand, the utilitarian argument in favour of the secularization of criminal law affirms the indifference of religious attitudes regarding the goal of safeguarding the civil order. The retributivist argument, on the other hand, is found in the thesis according to which sanctions for offences against God belong to God itself, not to the State.

As for the deontology of the punishment, Tarello highlights how the three ideologies are attuned in challenging punitive cruelty. In the face of 'terrible 'deterrents' ' of the *ancien régime,* the different components of legal Enlightenment marched together under the same banner, fighting for the moderation of repressive orders.

In his effort of historical understanding, Tarello arrives at a solid conclusion:

> The opposition between the three schools is in the principles, not in their historical expressions along the Eighteenth century; it is only at (…) the beginning of the Nineteenth century that the principles will be discussed in their crystalline purity, and conflicts will emerge on a theoretical level; nevertheless, at that time, those principles had already been absorbed into

[40] ibid 388.
[41] ibid 388.
[42] ibid 388.
[43] ibid 388.
[44] ibid 388.
[45] ibid 389.

the institutional realities and the conflicts were mainly academic.[46]

However, not even this conclusion fully satisfies the interests propelling our investigation. It is, therefore, necessary to first understand this conclusion's limits to then overcome them. To do so, we should first note that just as disputes about principles and theoretical conflicts developed in the Nineteenth-century criminal law are foreign to the Enlightenment culture, the image of the 'three schools' or 'ideologies (…) that characterize the Eighteenth century'[47] also appears to be an anachronistic distortion of reality. Indeed, if we can find no doctrinal contrasts between utilitarian and proportionalist conceptions of punishment in the philosophy of the Enlightenment thinkers, it is because those conceptions are not constructed as ideologies and even less so as schools. They are simply conceptions that – for the most part – coexist within the same normative discourse.

The notion of 'ideology' appears inadequate in relation to the conceptions of punishment distinguished by Tarello. We may, of course, argue that this is only a matter of semantics. However, if by ideology we refer to a set of evaluative and normative theses expressing a particular value system, and use it to characterize the position of the one professing them in the face of the reality to which they refer to, then neither penal utilitarianism nor the proportionalist conception of punishment can properly be called ideologies. For instance, Cesare Beccaria's condemnation of the death penalty is utilitarian, and so is its apology by Ferdinando Facchinei. Can we thus argue that Beccaria and Facchinei are exponents of the same ideology? On closer inspection, it rather appears that utilitarianism and retributivism cut across ideological camps of Eighteenth-century criminal law, as they are two axiologically flexible conceptions of criminal law. They can be grounded in different value systems and lend themselves to the support of the most diverse normative theses.

On the other hand, the humanitarian doctrine of the mildness of punishments is ideologically saturated. Unsurprisingly, it is here that we find one of the main issues dividing Beccaria's supporters and his opponents. Based on this simple observation, we may highlight another aporia in Tarello's analysis. Since utilitarianism does not, of itself, imply any mandate regarding the *modus puniendi* – neither concerning the type nor the extent of punishment – it is an error of perspective to represent the issue of penal mitigation as its derivative. The *philosophe* who employs utilitarian arguments in the debate against cruel punishments does so on account of an ethico-political option – namely, a moral attitude marked by a set of values which the Enlightenment vocabulary epitomizes in the word 'humanity'. It is precisely humanitarianism that directs the utilitarian discourse in favour of the mildness of penalties. Thus,

[46] ibid 390.
[47] ibid 387.

the 'derivation scheme'[48] proposed by Tarello should actually be inverted. The Enlightenment renewal of penal utilitarianism depends on the diffusion of unprecedented humanitarian sensitivity and the promotion of the person as an end in itself.

Put briefly, in our attempt to understand the Enlightenment juridico-political doctrines, we must commit ourselves to experiment with new instruments of observation and analysis. In this direction, I wish to propose a three-step approach to the problem of the criminal harm in the Enlightenment philosophy: 1) The discovery of meta-legislative principles related to the punishment, present in the reformist debate on the right to punish; 2) The clarification of theoretical foundations and of the normative scope of the identified principles; 3) A review of the relationship between these principles and the accepted or rejected punitive typologies.

III. Crimes and Punishments

1. The Canons of Criminal Justice

Two fundamental documents in the history of modern legal culture facilitate the realization of a first, approximate outline of the normative principles underlying the deontologies of punishment developed in the age of Enlightenment. Following their chronologic order, we being by opening – on its last page – Beccaria's 'miraculous booklet',[49] where the author argues:

> In order that punishment should not be an act of violence perpetrated by one or many upon a private citizen, it is essential that it should be public, speedy, necessary, the minimum possible in the given circumstances, proportionate to the crime, and determined by the law.[50]

On Crimes and Punishments ends with this connotation of the just punishment. To appreciate its revolutionary effects, we should travel from Milan to Paris – from the *Accademia dei Pugni* to the *Assemblée nationale*. It is here that on 23 May 1791, Louis-Michel Le Peletier presented the draft of the penal code, drawn up by the Constitutional and the Criminal Legislation Committees. Summarizing the list of principles of the 'penalty theory'[51] underlying the text, he declared:

> The punishments must be humane, properly graduated, in an exact relation to the nature of the offence, equal for all citizens, free from any

[48] ibid 389.
[49] P. Calamandrei, 'L'inchiesta sulle carceri e sulla tortura' *Il Ponte*, 229 (1949).
[50] C. Beccaria, n 16 above, 113.
[51] L.M. Le Peletier, n 18 above, 321.

> judicial arbitrariness. It is necessary that they not be distorted in the manner of their execution. They must be repressive, mainly by embarrassment and deprivation, (...) by their publicity, by their proximity to the place where the crime was committed. It's important that they correct the moral inclinations of the condemned, through the habit of work. (...) Finally, they must be temporary.[52]

The confrontation of these two passages clearly reveals the growing scope of the Enlightenment debate on the right to punish. In the summary of the canons of the punishment's legitimacy with which Beccaria concludes his *pamphlet*, we find only one regulative principle regarding the correlation between the offences and punishments, namely that of proportionality. On the other hand, in Le Peletier's compendium of the penal code's principles, we find two, namely, that the punishment must be a) 'properly graduated', b) 'in an exact relation to the nature of the offence'.[53] The normative connotation *sub* a) corresponds to the principle asserted by Beccaria – that is, that the correct graduation of the punishment consists in 'establishing the proportion between severity of the punishment and the seriousness of the crime'.[54] On the other hand, the directive *sub* b) prescribes that the type of punishment must depend on the type of crime committed –

> physical pain will punish the offences causes by ferocity; hard work will be imposed on the culprit whose crime has its source in laziness; infamy will punish actions inspired by an abject and degraded soul.[55]

Here, then, we find the first important difference between the two principles: on the basis of the first, a quantitative relationship between crimes and punishments must be established, whereas on the basis of the second, the relationship is qualitative. Both principles apparently satisfy the requirements of retributive justice. However, to reduce them both, *sic et simpliciter*, to the sphere of penal retributivism would be to misunderstand and confuse them.

If we analyse them individually, it becomes easy to see their different philosophical consistency as well as their different normative implications. In order to avoid ambiguity, let us first of all name and define them. I propose to call 'the principle of homogeneity' the canon of legislation according to which the punishment must typologically correspond to the crime. Also, I propose to restrict the extensional meaning of the expression 'principle of proportionality' to the prescriptive thesis according to which the harshness of the punishment must be commensurate to the gravity of the crime.

[52] ibid 323.
[53] ibid 323.
[54] ibid 322.
[55] ibid 322.

2. Utilitarianism and the Principle of Proportionality

The affirmation of this principle of proportionality in the Enlightenment discourse rarely rests on the retributivist moral postulates. Rather, for the most part, it pertains to the utilitarian logic of general prevention. All too known are the pages of the *On Crimes and Punishments*' Chapter VI to lay them out as proof of this observation. All too numerous are the repetitions of the Russian example – with which Montesquieu denounced the harmful consequences of the lack of proportion between crimes and punishments[56] – to be able to mention them without resulting tedious. Rather, it is worthwhile to return to Le Peletier, whose *Rapport* to the Constituent Assembly results exemplary in this respect as well. Says Le Peletier,

> The effectiveness of a penalty depends less on its severity than on its proper place in the scale of penalties. It's important that each crime is punished in proportion to the punishments associated with the other crimes. It's important that there is a fair relationship between the various degrees of the scale.[57]

Substantiated in this way, the principle of proportionality is the result of a utilitarian conception of punishment: *punitur ne peccetur*, and not *quia peccatum est*. However, precisely for this preventive purpose, *punitur proportionabiliter ad peccata*. In short, the respect for the principle in question is recommended to the legislator as the key factor of deterrence:

> If a great distance separates the punishment for one crime from the punishment for another crime, the bad man who cold-bloodedly meditates upon a bad action will stop where a great danger begins for him.[58]

On the other hand, the absence of an escalation in the sanctioning reaction – one corresponding to the scale of offensiveness of the prohibited actions – incentivizes the delinquent to maximize profit:

> It was a great absurdity of our laws to punish the thief on the high road (…) with the same punishment as the murderer. The law itself invited (one) to murder since murder did not aggravate the punishment (…) and could suppress the proof of the crime'.[59]

We can, therefore, say that in the context of Eighteenth-century reformism,

[56] 'In Russia, where the punishment of robbery and murder is the same, they always murder (footnote omitted). The dead, say they, tell no tales.' See Montesquieu, *The Spirit of the Laws* (Kitchener: Batoche Books, 2001), 107.

[57] L. M. Le Peletier, n 18 above, 321.

[58] ibid 322.

[59] ibid 322.

a utilitarian doctrine of a normative relation between crimes and punishments, based on the criterion of quantitative proportion, begins to take shape. The clearest confirmation of the removal of proportionalism from retributivism comes from Jeremy Bentham's philosophy of punishment.[60]

In its utilitarian version, the principle of proportionality appears as a norm on the production of norms. As such, it is clearly an extra-legal norm, addressed at the legislator. More specifically, as far as it affects the relationship between prohibitions and sanctions, it is a norm on (the production of) penal norms. It is, therefore, a substantive meta-norm, seeing how it concerns the content of normative production. If it has been specified that this deontic configuration regards the utilitarian version of the proportionality principle, it is because the latter can be – within a retributivist framework – loaded with a different prescriptive meaning. Whoever believes that we ought to punish by looking at past behaviour is led to conceive the canon of proportionality between crimes and punishments as a principle of equity; that is, of justice in a concrete case and, thus, as a meta-jurisprudential principle. In this perspective, it is up to the judge to proportion the crime to the punishment.

It is with the innovative doctrines of penal deterrence, proposed by Enlightenment thinkers, that the principle of proportionality emerges as a prescription addressed at the legislator. It imposes upon the latter the duty to define a scale of punishments tailored to the hierarchy of the goods protected by prohibitive norms. This – in consequence – also requires the criminal law to be ordered in a unitary and systematic *corpus* of norms. In this sense, the principle of proportionality is constantly invoked in the struggle for codification.

Moreover, within the scope of this struggle, the principle of proportionality reinforces the requirement for the mitigation of punishments. Demonstrating the usefulness of proportionality against a repressive system that introduced severe punishments even for minor crimes and indiscriminately imposed death for deviant behaviour of diverse gravity, allowed for a defence of penal humanitarianism – not in terms of philanthropy, but of political rationality. Utilitarianism (*punitur ne peccetur*), proportionalism (*punitur iuxta gravitatem criminis*) and humanitarianism (*punitur temperate*) thus formed the rings of a single argumentative chain.

3. Natural Law and the Principle of Homogeneity

'Civil liberty flourishes when the laws deduce every punishment from the peculiar nature of every crime'. Citing this maxim from Catherin II of Russia's *Nakaz,* Foucault points out that the empress takes 'almost word for word' 'Beccaria's lesson on the specificity and variety of penalities'.[61] This remark,

[60] J. Bentham, *Traités de législation civile et pénale* , in J. Bentham ed, *Œuvres de Jérémie Bentham* (Bruxelles: Société Belge de Librairie, 1840), Book I, Part III, Chapter II, 156.

[61] M. Foucault, n 2 above, 117.

however, is philologically and conceptually inaccurate. As is well known, that maxim derives from Montesquieu's *opus maior*[62] and expresses a principle that cannot be related to Beccaria's penal philosophy – namely, the principle of homogeneity.

The idea of a 'natural correspondence between punishment and crime' – writes Luigi Ferrajoli – is the 'oldest answer' to the question 'how to punish' and is 'closely linked to the penal retributivism and the doctrines of natural law'.[63] Its Enlightenment reformulation fully confirms this connection. Montesquieu elevates the moral postulate – according to which 'an intelligent being who has done harm to another intelligent being deserves the same harm in return' – to the level of 'relation of justice antecedent to the positive law'.[64] The power to punish is thus justified in its retributive function. The legitimacy of its exercise is conditioned to the respect for the principle of homogeneity, which imposes the qualitative conformity of the punishment to the crime.

As the basic principle of criminal justice, *lex talionis* is arguably the most rigorous and consequential expression of the above-mentioned requirement. A prominent example is found in the Antonio Genovesi's philosophico-juridical thought. In what is considered to be the first natural law treaties in Italy,[65] this admirer of Montesquieu, proposes the identification of a naturally-derived rule which renders a punishment just.[66] The normative proposal of his cognitive investigation is a complex penology,[67] based on the talionic principle:

> whoever violates a right, loses one himself, and of the same kind;[68] every punishment that is equal to all violated rights, is always a *talion*. If it is not a *talion*, it is not equal, and therefore not just; in consequence, it is a crime punishable by another talionic law.[69]

Those who position penal Enlightenment within the philosophical framework of utilitarianism cannot but leave outside of it those Enlightenment thinkers who, following in Montesquieu's footsteps, accepted and revitalized the idea of a punishment that is congenerous, isomorphic, or 'equal' to the crime. Those who see the talion as nothing more than an archaic criminal rule, supported by

[62] Cf Montesquieu, n 56 above, 207-208: 'Liberty is in perfection when criminal laws derive each punishment from the particular nature of the crime'.

[63] L. Ferrajoli, *Diritto e ragione. Teoria del garantismo penale* (Roma-Bari Laterza, 2011), 384.

[64] Montesquieu, n 56 above, 1.

[65] I. Birocchi, 'Genovesi, Antonio', in Id et al eds, *Dizionario biografico dei giuristi italiani (XII-XX secolo)* (Bologna: il Mulino, 2013), I, 964.

[66] A. Genovesi, *Della Diceosina o sia della filosofia del giusto e dell'onesto*, in N. Guasti ed (Venezia: Centro di Studi sull'Illuminismo europeo 'G. Stiffoni', 2008), Book I, XX, para IV, 270.

[67] On Genovesi's philosophy of punishment, see D. Ippolito, *Diritti e potere. Indagini sull'Illuminismo penale* (Roma: Aracne, 2012), 105-127.

[68] A. Genovesi, n 66 above, 273-274.

[69] ibid I, 274-275.

religious authority, will be surprised to see it elevated to the position of the supreme principle of justice by the exponents of the cultural movement that symbolizes modernity and secularism. Some may even transform their amazement into a perplexity about the representation here provided.

The facts, however, are well-established and it would be superfluous to linger in exemplifications. It was already Michel Foucault, who in his 1972-1973 *Collège de France* course, recognized the talion as one of the 'models of punishment' proposed by the Eighteenth-century reformers. He explains its 'reappearance' in relation to the desire to prevent any abuse of power: indeed, where the penalty is, in its nature and strength, exactly correlative with the offence itself,[70] the arbitrariness of the penal agents is annihilated. In the Foucaultian lesson, it is precisely Beccaria that becomes the champion of *ius talionis*.

This paradoxical result (partially corrected in *Discipline and Punish*) is, in my view, the consequence of a fallacious identification between the regulative idea of retaliation (the paradigmatic version of the homogeneity principle) and a different model of penal normativization – namely, the principle of analogy.

With this, I mean the thesis according to which the legislator must establish punishments able to reflect the crime to which they are associated. The same as with the principle of homogeneity, the prescription here regards the relationship between the type of violation and the type of punishment. However, unlike the natural law and retributivist versions of homogeneity principle, the analogy criterion is dictated by a utilitarian *ratio* – it is in order to increase the deterrent efficacy of penal norms that the punishments have to be tailored to the nature of crimes:

> This sort of fit – argues Beccaria – greatly eases the comparison which ought to exist between the incentive to crime and the retribution of punishment, so that the latter removes and redirects the mind to ends other than those which the enticing idea of breaking the law would wish to point out.[71]

By reinforcing the infringement-punishment ideas, 'his criminal mimesis'[72] helps to dissuade potential offenders. By acting on the dynamics of psychological movements, it transforms the impulses into inhibitions; it activates – against themselves – passions that prompt one to commit a crime; it pits – with the utmost urgency – the idea of advantages achievable through a crime against the mirror image of the consequences that follow it.[73]

[70] M. Foucault, *The Punitive Society. Lectures at the Collège de France* (Basingstoke: Palgrave MacMillan, 2015), 69.

[71] C. Beccaria, n 16 above, 49.

[72] ibid 59.

[73] See the illuminating pages by Ph. Audegean, 'Dei delitti e delle pene: significato e genesi di

At this point, one might object that I am contradicting the axiom of analytical economy according to which *entia (et nomina) non sunt moltiplicanda sine necessitate*. After all, aren't the normative contents of the theses under examination identical? Are we not perhaps dealing with the same meta-legislative principle justified with different doctrinal arguments? Indeed, this is what I also thought until recently.[74] On the contrary, I now believe it is necessary to emphasize the difference between one absolute principle of justice, which is ontologically based on the natural order – ie the principle of homogeneity – and a different principle of criminal law policy – the principle of analogy – whose value is relative to its instrumental function. This is not a negligible difference. Indeed, it is a crucial one, as the Beccarian delegitimization of the death penalty proves.

The doctrines of isomorphic punishment – particularly the one on the identity between the criminal harm and the punitive harm (*talio esto*) – may appear to us as archaic and repugnant. Contemporary criminal guarantism dismisses them as the 'result of an illusion',[75] pointing to its 'deleterious effects' in obstructing the 'process of formalization and legalization' of sanctions and in the support for 'corporal and capital punishment'.[76] Only by casting aside these theoretical and moral evaluations, can we properly understand the normative implications of the homogeneity principle in the Enlightenment penal philosophy.

First of all, the requirement for the legal limitation of political power in light of individual freedom finds its expression precisely in this principle. When 'the type of punishment' derives from 'the nature of offences' – Marat echoes Montesquieu – liberty triumphs with justice, 'since the punishment comes not from the will of the legislator, but from the nature of things; thus, it is not man who does violence to man'.[77] In this version of criminal natural law, the punitive arbitrariness is annihilated as sanctioning norms are subtracted from the sovereign's will. The naturalization of punishment protects individuals from the excessive use of force and the risk of despotic oppression.

Another normative implication of the principle of homogeneity ought to be emphasized. In the penal order of the *ancien régime*, built on the dogma of the connection between the harshness and intimidating vigour of the punishment, death was foreseen as a punishment for a vast and heterogeneous series of crimes, such as crimes against divinity, the sovereign, property etc. Raised to the level of a fundamental canon of natural justice, the homogeneity principle required the

un pamphlet giuspolitico', in D. Ippolito ed, *La libertà attraverso il diritto. Illuminismo giuridico e questione penale* (Napoli: Editoriale scientifica, 2014), 71-92.

[74] D. Ippolito, *Lo spirito del garantismo. Montesquieu e il potere di punire* (Roma: Donzelli, 2016), 37-42.

[75] L. Ferrajoli, n 63 above, 384.

[76] ibid 385.

[77] J.P. Marat, n 15 above, 31-32.

illegitimacy of the existing sanctioning system to be denounced; the ideological buttresses of the penology of deterrence cruelty to be overthrown, and the ambit of application of the capital punishment to be drastically restricted. Tarello is therefore right when he observes that, in the Enlightenment re-elaboration, 'the ancient idea of punishment as retribution (...) appears profoundly innovative and humane'.[78]

IV. Criminal Law Principles and Types of Punishment

In the previous section, I sought to achieve the goals – limited to the principles of proportionality and homogeneity – indicated in point one and two of the work plan proposed at the end of the first section. I conclude my analysis by dealing with point three of the plan, namely the problem of the relationship between the theorized principles and the punitive methods.

First, a preliminary observation of a general nature: not all canons of the punishment's legitimacy that characterize Enlightenment doctrines involve indications regarding the composition of the penal arsenal. The principle of legality (*nulla poena sine lege*), for example, is compatible with any kind of sanction. The same can be said of the principles of equality, (temporal) promptness and (spatial) proximity. Not even the principle of necessity – or punitive economy – provides an answer to the question of 'which punishment'.

We should add that in the Enlightenment penal deontologies, the indicated sanctions are not always consistent with the professed principles. Take, for example, one of the most interesting characterizations of the just punishment, established by Le Peletier:

> A punishment must be and remain what the fairness of the law has made it, not what the severity or leniency of the executor of the judgment makes of it.[79]

Note that the prescription does not regard the relationship between the punishment established by the legislator and the one imposed by the judge, but rather the one between the penal norm and the penal execution. According to this perspective, the law ought to provide sanctions that, 'once imposed by the judge, cannot be distorted by the arbitrariness of those who execute them'.[80] Well, the prison does not pass the legitimacy test of this principle.[81]

[78] G. Tarello, n 39 above, 389.

[79] 'Any punishment which due to its nature can be either aggravated or attenuated according to the disposition of the one who imposes it on the convicted person, is essentially bad.' (L.M. Le Peletier, n 18 above, 322).

[80] ibid 323.

[81] See L. Ferrajoli, 'Il carcere: una contraddizione istituzionale', in L. Ferrajoli ed, *Il Paradigma garantista. Filosofia e critica del diritto penale*, D. Ippolito and S. Spina eds (Napoli: Editoriale

As for the homogeneity principle, what we already determined with regard to its importance in the defence of life against the abuse of the State's power to kill, ought to be here integrated with considerations regarding the flip side of the coin. Planted upon the ontological structures of law, the retributivist canon of *malum passionis propter malum actionis* corroborates and prescribes the capital punishment. Whoever 'has committed murder, (…) must die', writes Kant, for '(h)ere there is no substitute that will satisfy justice'.[82] 'Life is the only good without equivalent', argues Marat, '(…) thus, justice wants that the punishment for murder be capital'.[83]

Capital punishment aside, the principle of homogeneity justifies any kind of sanction that is equivalent to the type of infringement. If the punishment is 'the loss of a right' for a violation of a right, then – argues Filangieri – 'different kinds of rights will indicate (…) different kinds of punishment':[84]

> Life, honour, real property, personal property and the prerogatives depending on citizenship are the general objects of all social rights. We shall thus have (…) capital punishments, infamy punishments, pecuniary punishments, punishments depriving or suspending personal liberty, punishments depriving or suspending civic prerogatives.[85]

Only by abandoning this strictly retributivist conception of punishment does it become possible to conceive the punitive system *qua* political artifice, *qua* legal institution, *qua* the product of decisions for which men bear full responsibility. From this point of view, the distance between the natural law principle of homogeneity and the utilitarian principle of proportionality can be fully grasped. It is among the theorists of the latter that the awareness of the conventional character of criminal law and of the inexistence of natural relationships between crimes and punishments begins to grow in the Enlightenment period.

> As there is no relation between the pain of the punishment and the malice of the action, it is obvious that the distribution of the punishments, relative to the greater or lesser gravity of the offence, is an arbitrary matter,[86]

writes Diderot in the *Encyclopédie*. The modulation of the severity of punishments along the scale of crimes is therefore based upon a legislator's choice. 'There is always a first punishment, which is arbitrary; once this is fixed, it conditions all the others'.[87]

scientifica, 2016), 179-187.

[82] I. Kant, n 34 above, 142.

[83] J.P. Marat, n 15 above, 63

[84] G. Filangieri, n 22 above, Book III, Part II, Cha XXVIII, 17.

[85] ibid 18.

[86] D. Diderot, 'Châtiment' *Encyclopédie*, 1753, III, 250.

[87] D. Diderot, 'Osservazioni sull'Istruzione dell'Imperatrice di Russia ai deputati per

Compared to the strict principle of homogeneity, the proportionality criterion thus expands the power to decide on how to punish. How to, and how not to punish. If the punishments do not derive from natural law, it is then possible to criticize them, reform and abolish them. Thus, in this way, Diderot can challenge Montesquieu's retributivism (via his comment of Chaterine II's *Nakaz*), stating that one need not criminalize 'acts contrary to (…) good customs', as punishing them with infamy would indeed 'be a terrible atrocity'.[88] In this way, free from the idols of natural justice, Beccaria begins his fight against capital punishment.[89]

l'elaborazione delle leggi', XXII, in Id ed, *Scritti politici*, F. Diaz ed (Torino: UTET, 1967), 390. This work, never published by the author, was first printed in 1920.

[88] ibid 394.

[89] Cf Ph. Audegean, n 16 above, 152-170; P. Costa, 'Beccaria e la filosofia della pena', in R. Davis and P. Tincani eds, *Un fortunato libriccino. L'attualità di Cesare Beccaria* (Milano: L'Ornitorinco, 2014), 33-50; D. Ippolito, 'Contratto sociale e pena capitale. Beccaria vs. Rousseau' *Rivista internazionale di filosofia del diritto,* 580-620 (2014); G. Francioni, 'Ius e potestas. Beccaria e la pena di morte' *Beccaria. Revue d'Histoire du Droit de Punir,* 13-50 (2016).

Short Symposium on the Punishment

Kant on Punishment: Between Retribution, Deterrence and Human Dignity

Francesca Fantasia*

Abstract

This article aims at offering an organic understanding of different elements of the Kantian philosophical-juridical conception of punishment. After analyzing Kant's argument in favour of the legitimacy of the punishment, I will single out two distinct levels of analysis: on the one hand, that of the conditions of punishability in general, where the function of punishment as retribution is outlined; on the other, that pertaining to the identification of a criterion to adjudicate the severity of a punishment. Particular attention is paid to the different functions performed, in a juridical context, by the concept of humanity as a *sui generis* human right: either drawing the boundaries of what can be object of punishment, or imposing limitations to the punishments a criminal can undergo. Finally, a long-overlooked element of Kantian theory is considered: his acknowledgment of a preventive-specific role of punishment, albeit limited to a pragmatic sphere.

I. Introduction

For a long time, Kant has been interpreted as proposing an inflexible, absolute theory of punishment. Only in recent years, thanks to closer attention paid to his theory (and to German Idealism in general) by studies in penal theory, it has become possible to obtain a more comprehensive picture of a much discussed, criticized, and often misunderstood – at least in its single component elements – theory. A misunderstanding that has often engendered a more generalized confusion regarding the systematic relationship that holds, in Kantian moral philosophy, between ethics and law. To offer just a few examples of misunderstood concepts, we could mention Kant's definition of penal law as a 'categorial imperative', his employment of the *ius talionis*, his ideas about capital punishment, and the role played, in a juridical context, by the concepts of humanity and human dignity. That these ideas have been object of such divergent interpretations can be explained by the fact that Kant failed to offer a linear and exhaustive presentation of his conception of punishment.[1]

* PhD, Humboldt Universität/University of Palermo.

[1] The main sources to understand Kant's positions on these topics are: the paragraph 'On the Right to Punish and to Grant Clemence' in his *Doctrine of Right,* which *exclusively* explores the question of the criterion for punishment (hence not to be considered the only source for Kant's penal theory); parts of the *Introduction to the Metaphysics of Morals*, where Kant demonstrates

We can discern four different levels of analysis in Kant's penal theory. First of all, it briefly examines the questions of the legitimacy of punishment, considered by Kant to be closely linked to the concept of law itself (I). Secondly, it analyses the concept of punishability in its juridical meaning of *criminal liability* (II). At this level of analysis, Kant answers the question of whether someone should be punished (the 'if' of the punishment) by establishing the conditions for punishability. In this context, he states his well-known proscription of instrumental punishments (II.1) – which is closely linked to the idea of the human being as an end in itself (II.2) – and he further defines penal law as a categorical imperative (II.3). Thirdly, Kant asks how criminals should be punished (the 'how' of the punishment) (III), identifying in the *ius talionis* (III.1) the sole *a priori* criterion applicable to establish the appropriate severity of a punishment, and to draw its boundaries – particularly those imposed by the idea of humanity (III.2). It is in this context that Kant outlines his refutation of Beccaria's theses against the death penalty (III.3). From a Kantian perspective this question, so central to the post-Enlightenment debate which led to radical juridical reforms in many European states, will be formulated from the fundamental principles of law alone. When evaluating the magnitude of the punishment there is also a fourth level of analysis, albeit secondary in importance: the pragmatic one (IV). Here Kant's acknowledgment – long ignored by the secondary literature – of a preventive-specific function of punishments emerges clearly. In this case, philosophical analysis steps beyond the boundaries of pure practical reason in order to consider a kind of pragmatic reason (or practical ability) as a fundamental pre-requisite of reforms in a juridical-historical context.

Of particular relevance for Kantian critical philosophy is the (doubly) limiting function of the idea of humanity. Far from being a concept like any other, the bond imposed by humanity is, for Kant, the very foundation of moral obligation. The constraint applied on the agent by his or her own humanity – to recognize him or herself as a moral (both ethical and juridical) subject – and the injunction to consider one's intrinsic worth as a limit in the relationships with others and oneself, constitute both the origin and the guiding principles for

the foundation for the authorization to punish; the *Reflexionen zur Moralphilosophie* and the *Reflexionen zur Rechtsphilosophie,* as well as other, mostly unpublished, texts where the preventive-specific function of punishment is clearly delineated. References to the works of Kant follow volume and page of the German Academy edition (AA), I. Kant, *Kants Gesammelte Schriften* (Berlin: Reimer, then de Gruyter, 1902), 1-29. Among recent studies on Kant's penal theory we should mention R. Brandt, 'Gerechtigkeit und Strafgerechtigkeit bei Kant', in G. Schönrich and Y. Kato eds, *Kant in der Diskussion der Moderne* (Frankfurt/M.: Suhrkamp Verlag, 1996), 425-463; B.S. Byrd and J. Hruschka, *Kant's Doctrine of Right. A Commentary* (Cambridge: Cambridge University Press, 2010), 261-278; O. Höffe, 'Vom Straf- und Begnadungsrecht', in O. Höffe ed, *Immanuel Kant. Metaphysische Anfangsgründe der Rechtslehre* (Berlin: Akademie Verlag, 1999), 213-233; G. Mohr, 'Nur weil er verbrochen hat. Menschenwürde und Vergeltung in Kants Strafrechtsphilosophie', in H. Klemme ed, *Kant und die Zukunft der europäischen Aufklärung* (Berlin-New York: Walter de Gruyter, 2009), 469-499; D. Tafani, 'Kant e il diritto di punire' *Quaderni fiorentini per la storia del pensiero giuridico moderno*, 55-84 (2000).

the normativity of practical action as a whole.[2]

In a juridical context, this right to humanity gives rise to a right to juridical humanity: a presupposition of a system of jurisprudence that also appears in several passages of the *Rechtslehre*, as a limiting condition for juridical measures.[3] Functioning as a guarantee against any possible obstacle to freedom, humanity is at once a right and not fully a (juridical) right, while still being able to constrain the juridical field.[4] It can impose a number of proscriptions, like that, for a State, to use its citizens as instruments in the event of war,[5] the obligation to dissolve, in time, all permanent armies, the obligations shared by states,[6] and – what is most pertinent to the current discussion – the proscription of using punishment as an instrument for the sake of some other good, or with the aim of demeaning the dignity of the criminal as human being.

II. The Legitimacy of the Punishment

In contrast to moral law, which regulates internal freedom and is aimed at the individual's moral intention (*Gesinnung*), legal law pertains to the form of the external relations between free wills, and excercises binding power over pathological motives of determination of the will. More precisely, it addresses motives arising not so much from the inclinations of the individual as much as from conflicts with other subjects. Unlike the ethical field's case, the motive for compliance with legal laws depends on a heterogeneous element that is added to external laws, to guarantee their observance. Now, Kant maintains, such motive is analytically included in those very laws:

> In all lawgiving (…) there are two elements: *first,* a law, which represents an action that is to be done as *objectively* necessary, that is, which makes

[2] About this see *Reflexionen* no 7862 in 19:538. This doesn't mean that right is grounded (and dependent) on ethics: from the Kantian standpoint ethics and right needs to be stritcly separated. (cf 6:93-100). On the relationship between right and ethics in Kant see M. Baum, 'Recht und Ethik in Kants praktischen Philosophie', in J. Stolzenberg ed, *Kant in der Gegenwart* (Berlin-New York: Walter de Gruyter, 2007), 213-226; B. Dörflinger, D. Hüning and G. Kruck eds, *Das Verhältnis von Recht und Ethik in Kants praktischer Philosophie* (Hildesheim-Zürich-New York: George Olms, 2017).

[3] Cf 4:431. As it was thoroughly demonstrated by Ponchio, the right to humanity is not, strictly speaking, part of right, because this requires an external constraint; it doesn't belong to ethics either, because this requires that duty be elected as a motive (*Triebfeder*) for action. See A. Ponchio, *Etica e diritto in Kant. Un'interpretazione comprensiva della morale kantiana* (Pisa: Edizioni ETS, 2011), 198-214.

[4] The move from right to humanity imposes on the subject a preliminary action as condition for the very existence of right, a kind of *Vor-Leistung,* a stance taken towards the relationships with others and oneself, as a holder of both rights and duties. Cf O. Höffe, *Königliche Völker. Zu Kants kosmopolitischer Rechts- und Friedenstheorie* (Frankfurt am Main: Suhrkamp, 2001), 157-160.

[5] 6:345.

[6] 7:345.

> the action a duty; and *second*, an incentive, which connects a ground for determining choice to this action *subjectively* with the representation of the law.[7]

By affirming the immediate inclusion of the motive in the law, Kant draws an analytical connection between the concept of right and the authorization to coerce. If, according to the *Introduction to the Doctrine of Right*,

> Right is (…) the sum of the conditions under which the choice of one can be united with the choice of another in accordance with a universal law of freedom[8]

or even

> the possibility of a fully reciprocal use of coercion that is consistent with everyone's freedom in accordance with universal laws,[9]

it follows that the legitimacy of the use of coercion coincides with the right to punish:

> There is connected with right by the principle of contradiction an authorization to coerce someone who infringes upon it.[10]

Right and authorization to use coercion (*Zwangsbefugnis*) 'mean one and the same thing'.[11] So in the *Further Discussion of the Concept of the Right to Punish* (Section V of the *Doctrine of Right*), Kant writes that the right to punish has its foundation in the very concept of public right:

> The mere idea of a civil constitution among *human beings* carries with it the concept of punitive justice belonging to the supreme authority.[12]

In this context, the penal institution represents the supreme power's legal instrument used for producing constraint, and, from an inverse perspective, the legal motive that conditions the free will of a potential offender towards the observance of public laws.

On the basis of this close link between external law, legal motive and law/coercion, the transition from the state of nature to the state which, according to Kant, individuals are obliged by the postulate of public law, is conceived as the transfer by individuals to the state, not of the content – even to

[7] 6:218.
[8] 6:230.
[9] 6:232.
[10] 6:231.
[11] 6:230.
[12] 6:362.

the slightest degree – of each person's rights (a right already present in a provisional form of private right) but rather of the very authorization to use coercion, which in the state of nature belongs to the subjective law of each person. Kant then sees the transition to the juridical state as a modification of the form of coercion: moving from violence (*violentia*) to distributive justice (*austeilende Gerechtigkeit*), in which 'what belongs to each can be secured to him against everyone else (*lex iustitiae*)'.[13] According to the postulate of public right, in fact,

> you ought to leave the state of nature and proceed with them into a rightful condition, that is, a condition of distributive justice. The ground of this postulate can be explicated analytically from the concept *of right* in external relations, in contrast with *violence* (*violentia*).[14]

Hence, public justice is also defined as 'the formal condition under which this is possible in accordance with the idea of a will giving laws for everyone'.[15] The State, therefore, takes on as its primary task – as a State – that of guaranteeing the rights of each individual; and precisely this guarantee represents the reason that pushes the individual to enter, willingly or unwillingly, into the juridical state.

Criminal law is therefore based on the right to punish, and aims at the creation of a legal motive. The state's right to punish is, in turn, based on the need to ensure that juridical law is actually binding.[16] The principle from which criminal laws derive consists in the establishment of a mechanism which, by associating punishment with the violation of the law, produces an act in accordance with the law. Here we recognize in Kant what today we consider to be a clear *utilitarian foundation* of criminal law, as well as an evident general-preventive function of punishment.[17] This preventive function is not mentioned in the passages of the *Doctrine of Right* concerning criminal law: there, Kant's primary interest is that of conducting a systematic criticism of the (*exclusive*) preventive penal theories. However, this general-preventive function emerges in other writings too. In a reflection on the philosophy of right (*Reflexionen* no 8026), Kant writes:

> *Strafe ist das Zwangsmittel, den Gesetzen Achtung zu verschaffen. Laesionen einer Person werden abgewehrt aber nicht bestraft in statu*

[13] 6:237.
[14] 6:307.
[15] 6:306.
[16] On the legitimacy of penal punishment in Kant, see W. Enderlein, 'Die Begründung der Strafe bei Kant' *Kant-Studien*, 303-327 (1985); H.G. Schmitz, *Zur Legitimität der Kriminalstrafe. Philosophische Erörterungen* (Berlin: Duncker & Humblot, 2001).
[17] Cf B.S. Byrd, 'Kant's Theory of Punishment: Deterrence in its Threat, Retribution in its Execution' *Law and Philosophy*, 151-200 (1989).

naturali, weil da kein äußeres Gesetz ist,[18]

and in *On the Common Saying*, too, Kant writes that external laws are

> *public coercive laws,* by which what belongs to each can be determined for him and secured against encroachment by any other.[19]

III. The Question of the 'if' of Punishment

1. The Neutral Meaning of Retribution (*Vergeltung*)

Based on the model of the separation of the State's three powers, which guarantees the reciprocal limitation of the exercise of power, Kant distinguishes between three different aspects of punishment. The first, from a legislative point of view, concerns punishment as an intended effect of the promulgation of a criminal law; the second, from a judicial point of view, concerns the punishment that is inflicted on the criminal through a sentence; the third aspect, from an executive point of view, concerns the execution of the punishment, namely 'the right a ruler has against a subject to inflict pain upon him because of his having committed a crime'.[20]

Kant systematically rejects any kind of preventive theory of criminal justice, wanting to demonstrate that a punishment can never be justified by its purpose. To construe the punitive institution as useful for a certain purpose would entail, for Kant, that the State – represented by the three bodies: legislative, judicial, and executive – would be entitled to eliminate or modify the law or its application according to the circumstances. At the judicial level in particular, the punishment would risk being sanctioned, in an individual case, for purposes that go beyond the crime in itself, and the person who committed it. Kant provides the conditions for the application of judicial punishment (*poena forensis*), by distinguishing it from natural punishment (*poena naturalis*), formulating here the famous prohibition of a punishment assigned for the sake of something else:

> *Punishment by a court* (*poena forensis*) (…) can never be inflicted merely as a means to promote some other good for the criminal himself or for civil society. It must always be inflicted upon him only *because he has committed a crime*.[21]

[18] 'Punishment is a means of coercion to establish respect for the law. In the state of nature the crimes towards a person are rejected but not punished, because here there is no external law' (19:585).

[19] 8:289.

[20] 6:331.

[21] ibid

From an anti-utilitarian perspective, a punishment must be imposed on the offender first and foremost on the basis of his or her having actually committed the offence. It follows that the condition of punishability of the accused is the sole condition for the sentencing to a punishment: the punishment cannot be imposed by the judge in order to pursue any aim other than that of judging an already performed act. The principle that a crime must be matched by an adequate punishment must override any other utilitarian consideration that may arise: the defendant

> must previously have been found *punishable* (*strafbar*), before any thought can be given to drawing from his punishment something of use for himself or his fellow citizens.[22]

It follows that although the punishment can certainly be considered useful, this can only be relevant at a later stage and on a secondary level of reflection. By prohibiting an instrumental understanding of punishments, Kant reaffirms the characteristic principle of the liberal juridical state, aimed at satisfying the need to guarantee legality: the indispensable premise of any legal punishment is the ascertainment of the actual perpetration of a crime.

Two consequences derive from this argument. First, as a *post-factum* sentence, the imposition of a punishment to an individual case can only refer to the past, to the *fait accompli*: this is a characteristic element of every absolute theory of punishment, with exclusive reference to the past (*quia peccatum est*). Secondly, the attestation of punishability implies that the crime must be proven: as it is pointed out in a reflection of moral philosophy (*Reflexionen* no 7491) '*Es kan niemand gestraft werden als nach bewiesenem Verbrechen*' ('no one can be punished except on the basis of a proven crime')[23] – a claim which might lead us to consider Kant as a precursor of the presumption of innocence.

Kant therefore presents a neutral and restrictive meaning of retribution (*Vergeltung*), one judging a punishment legitimate only where the subject has *voluntarily* committed a *crime*. This implies, on the one hand, a subjective constraint concerning the accused person and, on the other, an objective constraint concerning the gravity of the act committed.

On the one hand, the person must be recognised as being liable, ie his or her capacity to act as the person responsible for his or her own actions must be verified. The actor must possess free will and a healthy intellect, thus being an autonomous subject capable of imputation and, as such, depositary of rights and duties. It follows that in the case, for example, of an accused whose mental health is doubtful, the judge is required, first of all, to ascertain his or her sanity with the competent authorities. In the case of an inherited and incurable illness,

[22] ibid
[23] 19:413.

the judge will be required to assume that the subject may have been hindered in his or her actions by internal factors or mental disorders, and thus to hold the accused as not (or only partially) responsible for his or her actions.[24] As a recent critical study has pointed out, this close link between criminal law and psychiatry represents a significant contribution of Kantian philosophy to modern forensic psychiatry.[25]

On the other hand, the object (the act performed) must have legal significance, that is to say it must be recognised as an act that can be evaluated by the law. It must be judged by a court as the result of an action freely carried out under public law.[26] In other words, the fact must fall under the definition of a crime as that 'which violates the security a state gives each in his possession of what is his'.[27] This is a so-called 'public crime' (*crimen publicum*) which – unlike a private crime (*crimen*) defined as damage to a person and therefore judged by civil justice (such as an abuse of trust) – represents damage to the common body and involves the removal of civil status (such as embezzlement and business fraud) and is thus judged by criminal justice. Thus, only an act that does not conform to the freedom of others constitutes an offence worthy of penal action: an external and particularly serious action that endangers the life of citizens. It follows that a punishment will be legitimate only and exclusively when falling under the jurisdiction of a public law, and therefore only in the context of a lawful State. This means that the principle holds: there can be no punishment without law.[28] For these reasons, while an injustice will be punishable by law according to the very concept of right, not having at all transitioned into the legal state is to be considered the greatest injustice: a life without public laws knows no punishments because there is nothing that can guarantee individual rights.

2. Retribution (*Vergeltung*) and Humanity

At the basis of this doubly restrictive condition, determined by the gravity of the crime on the one hand, and by the person who committed it on the other, there is another more original constraint, imposed by the very nature of a

[24] See the Kantian exposition of various forms of mania (*gestörte Gemüth*) in the *Anthropology from a Pragmatic Point of View*, 7:213-214.

[25] Cf A. Mooij, 'Kant on Criminal Law and Psychiatry' *International Journal of Law and Psychiatry*, 335-341 (1998).

[26] Cf 6:227.

[27] 6:362.

[28] The sources for Kant are likely *Romans* 4:15 ('Where nolaw is, there is no transgression') and T. Hobbes, 'Leviathan', in C.B. Macpherson ed (Harmodsworth: Penguin, 1968), chapter 27 ('Where there is no law there is no sin (…) If civil laws cease, crimes cease'). The first appereance of the latin formula '*nulla poena sine lege*' can be found in P.J.A. Feuerbach, *Lehrbuch des gemeinen in Deutschland gültigen peinlichen Rechts* (Giessen: G.F. Heyer, 5th ed, 1812), 22. On Kant and Feuerbach cf J. Hruschka, *Kant und der Rechtsstaat und andere Essays zu Kants Rechtslehre und Ethik* (Freiburg im Breisgau: Karl Alber, 2015), 89-114.

human being as an end in itself (*Zweck an sich selbst*)[29] who, in virtue of his or her freedom – the moral foundation of right and ethics – is above all a responsible and imputable being.

Humanity, considered as an end in itself – as per the second formulation of the categorical imperative demonstrated in the context of the foundation of morals:

> *So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means*[30]

– is for Kant, in the context of criminal law, the reason for the ban on the instrumentality of punishment. The idea of humanity has a very precise and restrictive function in the juridical-penal sphere, and enters into the argument in a negative sense, and by refuting the theories of preventive justice and invalidating their possible consequences:

> (…) a human being can never be treated merely as a means to the purposes of another or be put among the objects of rights to things: his innate personality protects him from this, even though he can be condemned to lose his civil personality.[31]

The human's innate personality functions as the guarantee of a fair punishment, even when such punishment has the effect of depriving the accused, *qua* criminal, of his civil personality or legal status as an acquired right. By virtue of the innate capacity to consider him or herself as an end in him/herself and to relate to others as such, every human being deserves to have the unique and primordial right to innate freedom –[32] the precondition for any determination of the concept of right – understood as

> *Freedom* (independence from being constrained by another's choice), insofar as it can coexist with the freedom of every other in accordance with a universal law.[33]

Referring here to the so-called humanity formula of the categorical imperative, Kant invokes the principle of humanity as an end in itself – which defines the

[29] 4:435. By virtue of his morality the human being is a subject of the moral law (cf 5:87; 5:132), subject of pratical-moral reason, (cf 6:432; 434) and therefore subject of all (possible) ends (cf 4:430-431; 437), having 'the capacity to realize all sorts of possible ends, so far as this is to be found in the human being himself' (cf 6:392), to be the end of his own existence, and to determine his ends by the employment of reason (cf 5:431).

[30] 4:429. Cf also 8:107-108 and 113-114.

[31] 6:331.

[32] The moral law as a *ratio cognoscendi* of human nature – a freedom grounding the normativity of practical action – is one and innate.

[33] 6:237-238.

capacity of each human being (unlike things or animals) to determine a particular form of interpersonal and self-related relationship – as a distinctive feature of human nature. As a moral capacity underlying any relationship, whether ethical or legal, the special type of determination of relations between humans expressed by the idea of humanity as an end in itself, is explicitly distinguished from any other practical determination of inter-subjective relations. At the basis of the distinction made here by Kant between innate and civil personality there is that between a subject as a free subject, considered in 'his personality independent of physical attributes (*homo noumenon*)', and the same subject as 'affected by physical attributes, a human being (*homo phaenomenon*)'[34] – a distinction that will come to play a central role in Kant's discussion of Cesare Beccaria's theses, as we will see shortly. The person as a legal entity should also always be considered as different from the person seen from the standpoint of his or her civil status, that is to say as a human being as such.[35]

Being a *natural* possession, superior to any other acquired right, humanity is considered by Kant to be a regulative principle operating in the field of law: a *sui generis* human right, unique and innate, just as unique and innate as the preliminary right to freedom. Deriving from the wider field of morality, this right is the supreme limiting condition for any exercise of freedom (both internal and external):

> *Das Recht der Menschheit (ist) dasienige, was alle freyheit durch nothwendige Bedingungen einschränkt* (The right of Humanity (is) that which limits all freedom by necessary conditions).[36]

This is as much an ethical as it is a legal requirement, that both legislations are called upon to enforce in their different areas of competence. Within the framework of law, humanity constitutes such an *internal* limit of *external* freedom in intersubjective relations. It is the limiting condition for the external freedom of the individual in relating to others, and at the same time it constitutes the limit of the freedom of others when relating to the individual. In other words, it constitutes the legal duty to consider the other as a *possible subject of juridical relations*. This sole right of humanity, by virtue of which everyone, by its very nature, advances the right to act in the exercise of his

[34] Cf 6:239. Sadun Bordoni has recently highlighted one of the fundamental aspects of the Kantian distinction between *homo noumenon* and *homo phaenomenon* (one already emerging in the *Naturrecht Feyerabend*): it aims at affirming that freedom alone, and not reason as such, is the element which characterizes *homo noumenon*. See G.S. Bordoni, 'Leggi della natura e leggi della libertà. Kant e il giusnaturalismo', in T. Gregory ed, *Nomos, Lex* (Firenze: Olschki, 2017), 261-270.

[35] Cf 6:427-432. The idea of the human being as an end in itself represents the very content of the categorical imperative, and thus it is the idea – source of every obligation – to which the human must conform (cf 6:404). It is an ideal (cf 6:405), and a duty (cf 6:386) which expresses a responsibility for humankind as a whole.

[36] *Reflexionen* no 6801, 19:165-166.

innate freedom – that is to say, independently from the coercive arbitrariness of others – is the basis of the prohibition of the instrumentality of punishment.

To be primarily considered as an innate personality implies the right not to be judged a criminal without having first performed a (legally recognized) act that damages the external relations of individuals. In the exercise of external freedom, on the basis of human dignity (innate personality) even before civil dignity, the relationship between subjects must be thought of first and foremost as a relationship of innate equality: it is about the

> innate *equality,* that is, independence from being bound by others to more than one can in turn bind them; hence a human being's quality of being *his own master* (*sui iuris*), as well as being a human being *beyond reproach* (*iusti*), since before he performs any act affecting rights he has done no wrong to anyone.[37]

This original and innate meaning of equality between human beings is also applicable to the relationship between the representative of a State power (legislative, judicial, or executive) and the citizen: the 'Relation of the Subject Imposing Obligation to the Subject Put under Obligation' is a 'relation in terms of rights of human beings toward beings that have rights as well as duties' as a 'relation of human beings to human beings'.[38] If the presupposition for the free association of human beings is that imputability depends on the innate personality of an individual, then the relationship between judge and criminal will be first of all a human one, an equal relationship between two persons, going from *homo noumenon* to *homo noumenon* (the criminal *wanted* to carry out the crime, through the full exercise of his freedom, and by choosing so he or she would have acted voluntarily as a free and imputable subject, unlike any animal behaviour or thing). As a human being, therefore, the criminal will never be considered to be an object but, respecting the humanity in its person, always a *subject*. Indeed, he or she is a *co-subject* of a free causality that no punishment can (or better should) take away, and endowed with that innate and fundamental right that is freedom.

3. The Meaning of the Law of Punishment as Categorical Imperative

The Kantian presentation of criminal law as a categorical imperative should be read as part of his broader anti-utilitarian concerns. Due to a certain ambiguity, the following passage has given rise to divergent interpretations within the literature. Object of dispute is the following: the sense in which Kant here uses the term 'categorical imperative', and the identification of the addressee of this command, either a criminal or a judge:

[37] 6:237-238.
[38] 6:241.

> The law of punishment is a categorical imperative, and woe to him who crawls through the windings of eudaimonism in order to discover something that releases the criminal from punishment or even reduces its amount by the advantage it promises, in accordance with the Pharisaical saying, 'It is better for one man to die than for an entire people to perish'. For if justice goes, there is no longer any value in human being's living on the earth.[39]

Through this provocative association of criminal law with the categorical imperative, Kant is taking a stance against any type of punishment motivated by well-being, happiness, or other benefits that either the criminal or society at large may derive at any given time – or even a punishment simply prompted by an arbitrary decision of the judge. An exclusively utilitarian conception of punishment must be countered with an (*absolute*) idea of justice, which cannot be derived either empirically or pragmatically.

That of 'categorical imperative' is a very precise concept in Kantian moral philosophy: categorical imperative refers to, in an ethical context, the internal obligation imposed by the moral law, requiring the simple respect of the law itself. This free and voluntary adherence to duty itself overrides any empirical determination or conditioning. It 'would be that which represented an action as objectively necessary of itself, without reference to another end'.[40] Now, by defining criminal law as a categorical imperative (in the juridical context), Kant wants to say that criminal law must be considered as a '(morally practical) *law*',[41] which determines with apodictic validity and compelling strength[42] the sentencing of a criminal, seen as a necessary consequence of the objective commission of a crime, without reference to any other purpose.

With this juxtaposition of criminal law and categorical imperative, Kant certainly does not want to identify moral law with positive criminal law: rather, he is once again placing the emphasis, on the one hand, on the *necessary* and *unconditioned* nature of the allocation of the punishment under certain given conditions and, on the other, on the limits that the judicial body encounters in the exercise of its punitive power. Although a reading of the categorical imperative as a command addressed to the citizens may also seem plausible, it seems more convincing to argue that Kant is addressing the command to the public officials in charge of the administration of justice.[43]

[39] 6:332.

[40] 4:414.

[41] 'A (morally practical) law is a proposition that contains a categorical imperative (a command)' (6:227).

[42] Cf 6:222.

[43] The text is not clear on this point. On the one hand, the general meaning of the passage evidently concerns the authority that punishes; on the other hand, the textual meaning of the sentence concerns the liberation (even partial) from the punishment, therefore the guilty party. In this sense, this categorical imperative can be associated with the categorical imperative of public law (*Rechtslehre*, §2; §42): if you are unconditionally obliged to enter into the juridical state, you will

When speaking of a categorical imperative, Kant's implicit warning to those exercising judicial power is clear: do not confuse the form of absolute obligation (proper to a categorical imperative) to punish a proven crime with a form of relative obligation (proper to a hypothetical imperative) that would have another type of validity. If the sentence only had a relative validity, the absolute and priceless idea of justice that guarantees the value of the life of all human beings on Earth would disappear. The example given here by Kant is that of someone condemned to capital punishment who would propose to undergo dangerous medical experiments to promote the progress of medicine, in exchange for his life being spared. A court should reject in outrage such a proposal because, Kant argues, 'justice ceases to be justice if it can be bought for any price whatsoever'.[44]

The application of a punishment, therefore, is categorical in the sense that it refers to pure and rigorous justice: the power to sentence to a punishment, proper to the judicial body, must categorically submit to the simple *form* of criminal law and not yield to utilitarian calculations of any kind, for that would lead to the loss of its very character of justice.

Criminal law, therefore, established at the legislative level, *must* be *strictly* applied according to the formal principle of not leaving unpunished a proven crime. As the object of an unconditional and necessary duty, criminal law in its formal character – that is, independently of any empirical or material element that may be relevant to each case – categorically imposes on the judiciary the application of a punishment as a consequence of a crime. This will obviously be the case only in the case of a *proven* crime, committed by a *sane* criminal, and in full respect of the innate freedom of legal subjectivity *tout court*.

With criminal law as a categorical imperative Kant simply wants to warn against the lures of a utilitarian doctrine. Although highly provocative – and object of divergent interpretations – this definition is meant to limit the arbitrariness of the judges' decisions, and to reaffirm the division of powers in the specific field of criminal law.[45] Kant answers the question of the 'if' of the punishment with the idea of absolute justice, the simple and unconditional application of a formal law. If we are to think of a punishment's purpose, the only viable answer, for Kant, is the idea of justice itself.

IV. The Question of the 'how' of Punishment.

1. Retaliation (*Wiedervergeltung*) and Its Restriction

also be unconditionally obliged to act in it in accordance with public laws, including criminal laws. For an overview of the different interpretations of the criminal law as a categorical imperative, cf M.A. Cattaneo, *Dignità umana e pena nella filosofia di Kant* (Milano: Giuffrè, 1981), 225-317.

[44] 6:332.

[45] This is neither a direct derivation of the punishment from the moral imperative nor from the morality of the judge.

A much-discussed problem in the debate contemporary to Kant on punitive measures is that of defining how far one can go when punishing a criminal. From the Kantian perspective, if the investigation was left to empirical observation, and punitive measures were legitimized exclusively on the basis of their preventive effectiveness, there would be no limits to the monstrosity of punishments.[46] Once the guilt of the offender has been established, Kant then raises the question of the criterion for determining the adequate type of the punishment: 'But what kind and what amount of punishment is it that public justice makes its principle and measure?'.[47]

Ignoring the pragmatic value of the punishment, the only valid principle for determining its type and degree is 'none other than the principle of equality (*Prinzip der Gleichheit*) (in the position of the needle on the scale of justice), to incline no more to one side than to the other',[48] because only:

> the law of retribution (*Wiedervergeltungsrecht*) (*ius talionis*) – it being understood, of course, that this is applied by a court (not by your private judgment) – can specify definitely the quality and the quantity of punishment; all other principles are fluctuating and unsuited for a sentence of pure and strict justice because extraneous considerations are mixed into them.[49]

In order to properly establish the adequate amount of punishment, independently from any empirical considerations, human reason can only mobilize the principle of the *ius talionis* 'by its form' ie

> always the principle for the right to punish since it alone is the principle determining this idea *a priori* (not derived from experience of which measures would be most effective for eradicating crime).[50]

The *ius talionis* therefore ensures the proportionality of guilt and punishment, avoiding a punishment disproportionate to the crime committed (on the basis of its preventive effectiveness) and thus guaranteeing equality (as an *a priori* principle of right) between the magnitude of the crime and the measure of retribution, as well as the equality of everybody before criminal law.

Therefore, the law of retribution (retaliation) is not a justification for punishment, but it is the formal criterion employed to establish the proper amount of punishment. That is, it defines a purely formal criterion of equality between the transgression of public law and punitive action. The Kantian employment of the *ius talionis* does not concern *Vergeltung*, ie necessary

[46] Hence Kant places himself on the side of other Enlightenment thinkers, and of Beccaria himself, who have denounced the inhuman tortures of the old penal system.

[47] 6:332.

[48] ibid

[49] ibid

[50] 6:362-363.

retribution under specifics conditions, but rather *Wiedervergeltung*, ie *retaliation* in the sense of delivering to the offender the same type of suffering that he has caused. The latter indicates a just and legitimate need: that of a proportion between crime and punishment, another fundamental theme of Enlightenment debates.[51]

From a formal point of view, therefore, the *lex talionis* will always be applied, if not according to its letter, at least according to its spirit. Among specific forms of application of the principle of equality, Kant includes, as is well known, the death penalty:

> If, however, he has committed murder he must *die*. Here there is no substitute that will satisfy justice. There is no *similarity* between life, however wretched it may be, and death, hence no likeness between the crime and the retribution unless death is judicially carried out upon the wrongdoer.[52]

A closely related topic is that of the equal degree of punishment. If the *lex talionis* is admitted as the only principle of justice, then the equality of everybody before the criminal law – regardless of class or the different sensitivity of individuals to one type of punishment or another – will also be established. Kant's reference here is the Scottish rebellion and the freedom, to be established in court, for any offender to *choose* his or her penalty between *death* and *forced labour*. The criterion of choice here will be subjective, replies Kant, and so the choice would become a matter of honor: 'The man of honor would choose death, and the scoundrel convict labor'.[53] From the point of view of the objective principle of the *lex talionis*, the death penalty would be the only wholly proportionate and just punishment, for both offenders. In these specific cases, therefore, the principle of the equality of punishment – in this case the death penalty – intervenes to vanquish any doubt

> every murderer – anyone who commits murder, orders it, or is an accomplice in it – must suffer death; this is what justice, as the idea of judicial authority, wills in accordance with universal laws that are grounded *a priori*.[54]

This will also be true in the case of several criminals being judged together: 'When sentence is pronounced on a number of criminals united in a plot, the best equalizer before justice is *death*'.[55] Kant writes again:

[51] On the difference between *Vergeltung* and *Wiedervergeltung* see O. Höffe, 'Vom Straf- und Begnadungsrecht' n 1 above, 214-215.

[52] 6:333.

[53] 6:333-334.

[54] 6:333.

[55] 6:334.

> This fitting of punishment to the crime, which can occur only by a judge imposing the death sentence in accordance with the strict law of retribution (*Wiedervergeltungsrechte*), is shown by the fact that only by this is a sentence of death pronounced on every criminal in proportion to his *inner wickedness* (*Innere Bösartigkeit*) (even when the crime is not murder but another crime against the state that can be paid for only by death).[56]

However, a number of factors may limit the application of *ius talionis*. Kant mentions two exceptions. The first concerns the case of an island, where all citizens decide to disperse and thus dissolve the common legal body they compose. Before doing so, all the guilty inmates must first be executed, in order not to leave that common body – as long as it is legally so – unpunished. In this example, death would be the only penalty *corresponding*, not to the crime committed, but *to the time allowed* for serving the sentence *within* the State. In view of the imminent dissolution of the State itself, such a time is contracted to a single instant, that of the execution, the only timeframe corresponding to the disappearance of the State itself.

The second exception is that of a state plot orchestrated by all citizens: if he was to sentence everyone to death, the sovereign – the only innocent member of the community – would perform a 'carnage spectacle of a slaughterhouse'.[57] This would lead to the dissolution of the state itself, and to a return to the state by nature. Therefore, the sovereign would in this case have the exceptional right to assume the role of judge, and to issue a sentence condemning his subjects to a punishment other than death, such as deportation 'which still preserves the population'.[58] This would be a case of necessity (*casus necessitatis*), in which a sentence is issued by an executive decree and not as a public law, as 'an act of the right of majesty which, as clemency, can always be exercised only in individual cases'.[59]

2. Retaliation (*Wiedervergeltung*) and Humanity

Before analysing the Kantian refutation of Beccaria's theses against the death penalty, which completes the argument in these pages of his *Doctrine of Right*, it is worth considering the function performed here by the idea of humanity. The *ius talionis* is subject to a fundamental limitation: the penalty 'must still be freed from any mistreatment (*Mißhandlung*) that could make the humanity in the person suffering it into something abominable'.[60] That is to say, a

[56] 6:333. The *Innere Bösartigkeit* seems to suggest the necessity of considering the intention of the criminal. This, however, would imply an analysis of the intentions that does not pertian to the external freeom of the agent, but to his or her internal relationship with the maxims of action.

[57] 6:334.

[58] ibid

[59] ibid

[60] 6:333. Cf T.E. Hill, 'Treating Criminals as Ends in Themselves' *Jahrbuch für Recht und*

punishment must not involve torture, or a torment that would harm the humanity of the culprit. This raises the question of the respect for humanity, represented, so to speak, by the noumenical side (*homo noumenon*) of the natural person who will physically suffer the punishment (*homo phaenomenon*). This idea of humanity imposes that, despite the principle of equality, punishment must never be so severe as to harm – or even stand contrary to – humanity itself.

The Kantian argument now focuses on the limit posed to the principle of equality by the idea of humanity as an end in itself. If, for what pertains to punishability (the question of the 'if' of the punishment), the idea of humanity established the conditions for the punitive action in general – by prohibiting an abuse of the punitive instrument – when it comes to the *magnitude* of punishment (the question of the 'how'), the very same concept of humanity limits the validity of the *a priori* principle, by demanding an exact proportion between guilt and punishment. It thus establishes the prohibition of imposing, for crimes of enormous entity, penalties so great as to be morally unlawful.

A personal sphere is thus outlined, which cannot be violated by others and must also be respected by criminal justice: the sphere of human dignity (*Menschenwürde*).[61] Such dignity – connected by Kant in his moral works with the *autonomy of will*, in reference to the subject's self-relationship[62] – is associated in the *Metaphysics of Morals* with the human being's status as an *end in itself*, to be respected by others.[63] In this sense 'Humanity itself is a dignity'[64] and as such it is an 'absolute inner worth'[65] not comparable with others[66] and 'ineliminable').[67]

The primordial right to humanity, therefore, imposes to act (and to punish) within the limits of one's personal sphere of dignity, meant as the *status* of a human being as an end in itself. Respecting the inviolable sphere of human dignity thus implies, with regard to the amount of punishment, that the suffering inflicted to a criminal cannot exceed a certain extent, beyond which

Ethik/Annual Review of Law and Ethics, 17-36 (2003).

[61] Cf 6:236. On the meaning of humanity as an end in itself and human dignity see S. Bacin, 'Kant's Idea of Human Dignity: Between Tradition and Originality' *Kant-Studien*, 97-106 (2015); L. Caranti, *Kant's Political Legacy. Human Rights, Peace, Progress* (Wales: University of Wales Press, 2017); J. Glasgow, 'Kant's Conception of Humanity' *Journal of the History of Ideas*, 291-308 (2007); T.E. Hill, 'Humanity as an End in Itself' *Ethics*, 84-99 (1980); H.F. Klemme, 'Die vernünftige Natur existiert als Zweck an sich selbst' *Kant-Studien*, 88-96 (2015); G. Löhrer, *Menschliche Würde. Wissenschaftliche Geltung und metaphorische Grenze der praktischen Philosophie Kants* (Freiburg: K. Alber, 1995); O. Sensen, *Kant on Human Dignity* (Berlin-New York: de Gruyter, 2011).

[62] 'The dignity of humanity consists just in this capacity to give universal law, though with the condition of also being itself subject to this very lawgiving' (4:440).

[63] 6:434.

[64] 6:462.

[65] 6:435. *Person* and *personality* are other designations of human dignity (cf 6:462).

[66] V-NR- Feyerabend, 27.2,2:1319.

[67] 6:436. See also *Reflexionen* no 6801, 19:165-166.

his or her humanity would be demeaned.

In the *Further Discussion of the Concept of the Right to Punish* Kant, returning to the topic of the proper amount of punishment, explicitly raises the problem of respecting the criminal's humanity (albeit in a marginal paragraph):

> The only question is whether it is a matter of indifference to the legislator what kinds of punishment are adopted, as long as they are effective measures for eradicating crime (which violates the security a state gives each in his possession of what is his), or whether the legislator must also take into account respect for the humanity in the person of the wrongdoer (ie, respect for the species) simply on grounds of right.[68]

Referring to the prohibition of an instrumental punishment, Kant once again opposes the idea of punishment as a mere deterrent, this time in the context of a reflection on the kind of punitive treatment to be inflicted on the offender.[69]

Since the two levels (what would be required by the *ius talionis*, and the limit imposed by the idea of humanity) can come into conflict, Kant argues that the question must always be asked while *remaining within the framework* of the foundations of right, and never exceeding the limits of what can be established *a priori*. That is to say, the debate should never veer towards a pragmatic and empirical discussion on the effectiveness of punitive measures. Rather, the question here concerns the relationship, on an *a priori* level, between two ideas of reason: *humanity*, on the one hand, and *equality*, on the other. Since the idea of humanity as an end in itself is the only idea on the basis of which a juridical state *tout court* is possible, it follows that everything that is legitimate in criminal law, according to its *a priori* principle (the *ius talionis* as idea of equality), presupposes this idea of humanity and indeed owes it, so to speak, its status as a principle. The principle of equality (the *ius talionis*) is thus subordinated to the idea of humanity.

When following the principle of equality (*ius talionis*) as the sole *a priori* principle of right, the idea of humanity functions as a positive limit for the extent and the severity of punishments. The limits imposed to *ius talionis* by the idea of humanity will be particularly binding in extreme cases of crimes against humanity. Kant wonders:

> (b)ut what is to be done in the case of crimes that cannot be punished by a return for them because this would be either impossible or itself a punishable crime against humanity as such, for example, rape as well as

[68] 6:362-363.

[69] In this passage it is also clear that Kant is acknowleding the preventive-specific function of punishment (which I will explore below) as long as the punishment is equal and not more severe than the crime.

pederasty or bestiality?[70]

Kant offers the examples of rape, of pederasty, or of bestiality, called unnatural crimes because they are exercised against humanity itself. These cannot be punished with equal forms of suffering: '(t)o inflict *whatever* punishments *one chooses* for these crimes would be literally contrary to the concept of *punitive justice*'.[71] To sentence a punishment for crimes against humanity would go against the letter of criminal justice. Following its spirit, however, it is still possible to hold the criminal responsible and, exceptionally, to sentence him or her to an *arbitrary* punishment: such a punishment would have its intended effect, and the criminal, Kant explains, 'cannot complain'.[72] Through criminal law, then, the legislator establishes a punishment proportional to the crime while not demeaning the human dignity of the criminal. Such punishment will not violate the human dignity of the offender, even though it will determine his or her loss of dignity as a citizen. A *punishment against humanity*, on the other hand, would damage the very free subjectivity that founded and recognized the juridical state: paradoxically, according to Kant, one would be faced with a punishment against oneself as an Institution *of* Humanity *against* Humanity.

When confronted with such unnatural crimes, which demean the humanity of the other, the principle of equality – as a criterion for establishing a punishment – would demand a physically possible but morally impossible punishment. It is interesting, in this regard, that Kant wrote in his 1764 manuscript *Remarks on the Observations on the Feeling of the Beautiful and Sublime*:

> *Die Größe der Strafe ist entweder practisch zu schätzen nemlich daß sie groß gnug sey die Handlung zu verhindern u. denn ist keine größere Strafe erlaubt aber nicht immer ist eine so große Strafe als physisch nothig ist moralisch moglich.*[73]

Therefore, if a punishment proportionate to the crime is physically necessary (a crime cannot go unpunished), it must also be morally lawful (the *ius talionis* as the only principle for determining the severity of the punishment will be applicable only with respect to the inviolable sphere of *human dignity*).

3. The Kantian Refutation of Beccaria's Thesis

In the last part of Section E of *On the Right to Punish and to Grant Clemence* Kant explicitly – and with a somewhat ironic tone – aims to refute Beccaria's

[70] 6:363; cf 6:463.
[71] 6:363.
[72] ibid
[73] 'The magnitude of punishment is either to be evaluated practically, namely, that it be great enough to prevent the action, and then no greater punishment is allowed; but a punishment as great as is physically necessary is not always morally possible' (20:111).

argument on the illegitimacy of the death penalty. 'Moved by overly compassionate feelings of an affected humanity (*compassibilitas*)', Beccaria's theses are, for Kant, 'all sophistry and juristic trickery'.[74] Let us briefly explore Beccaria's ideas, in order to better understand this *juristic trickery* that Kant alludes to.

Beccaria, motivated by a strong utilitarian ethos – the idea of extending economic rationality to the criminal sphere and applying the instruments of logic and calculus to questions of social justice – claims, in *On Crimes and Punishments*, that the purpose of punishment

> is nothing other than to prevent the offender from doing fresh harm to his fellows and to deter others from doing likewise.[75]

He further defines justice as resulting from the '*the greatest happiness shared among the greater number*'[76] and offers an empirical grounding of the main arguments against the death penalty, promoting the greater deterring effect of alternative punitive measures. As we have seen, this kind of analysis is precluded by Kant's anti-utilitarian point of view.

On the other hand, Beccaria was also keen to determine the grounds and the limits of the right to punish, thus being concerned with punishments on the level of State legitimacy and of political obligation (and not, just like Kant, on the moral or the religious level), Beccaria bases his arguments against the death penalty on a specific conception of the social contract, according to which the limits to the sovereign authority's right to punish are imposed by the social contract itself. It is precisely such a theory, aimed at legitimizing the right to punish on the basis of the social contract – and not on *a priori* principles of law – that Kant is interested in refuting. The Kantian rejection of Beccaria's theses, far from being a defence of capital punishment, is a refutation of the theoretical principles of criminal law that are inconsistent with social contract theory. The sovereignty of the state, according to Beccaria, results from the sum total of all the freedoms that individuals renounced in exchange for security:

> (w)earied by living in an unending state of war and by a freedom rendered useless by the uncertainty of retaining it, they sacrifice a part of that freedom in order to enjoy what remains in security and calm.[77]

Against the Hobbesian thesis that individuals surrender *all* their freedoms to

[74] 6:335. Regarding Kant's critique to Beccaria see also AA 27:1391.

[75] C. Beccaria, *On Crimes and Punishments and Other Writings*, edited by Richard Bellamy (Cambridge: Cambridge University Press, 1995), 31. On Beccaria's Theses, cf B.E. Harcourt, 'Beccaria's 'On Crimes and Punishments': A Mirror on the History of the Foundations of Modern Criminal Law', in M. Dubber ed, *Foundational Texts in Modern Criminal Law* (Oxford: Oxford University Press, 2014), 39-60.

[76] ibid 7.

[77] ibid 9.

the Leviathan, Beccaria argues that citizens need giving up only the minimum necessary to achieve security, and more precisely 'the smallest possible portion consistent with persuading others to defend him'.[78] Punishments, therefore, are the main instrument used for enforcing the social contract, keeping individuals from attempting to regain possession of that small part of the freedom they relinquished. However, being a small part, the right of the sovereign authority to punish is equally minimal, that is to say, the minimum necessary to ensure safety:

> (t)he sum of these smallest possible portions constitutes the right to punish; everything more than that is no longer justice, but an abuse; it is a matter of fact not of right.[79]

According to this argument, the contract cannot legitimize the death penalty because it affects a good – life itself – which is a logical presupposition of freedom *tout court*, of which only a small part has been renounced.

From the Kantian perspective, the idea of the contract does not at all imply that the laws, and therefore also the criminal law, are the *object* of the contract; the idea of the contract merely represents the act through which the people turn themselves into a State.[80] The universal principle of right exists, for Kant, regardless of the 'consent' of the parties, and it cannot concern the 'content' of the law but only its form, as an emancipation from a provisional rule of law: criminal law 'could not be contained in the original civil contract'.[81] In Kant's construal of the original contract, therefore, there is no sacrifice of freedom – no matter how small:

> Everyone (*omnes et singuli*) within a *people* gives up his external freedom in order to take it up again immediately as a member of a commonwealth, that is, of a people considered as a state (*universi*).[82]

In this new condition, the human being

> has relinquished entirely his wild, lawless freedom in order to find his freedom as such undiminished, in a dependence upon laws, that is, in a rightful condition, since this dependence arises from his own lawgiving will.[83]

The transition from the state of nature to the juridical state amounts to the passage from an unsecured and provisional form to a peremptory juridical form of the *same content of the law*.

[78] ibid 11.
[79] ibid
[80] '*Status naturalis* is just an idea of reason' (MS/Vigil., 27.2,1:589).
[81] 6:335.
[82] 6:315.
[83] 6:316.

Hence the two main arguments of the Kantian refutation of Beccaria: on the one hand, the impossibility of considering a punishment as an object of the criminal's *will* and, on the other, the impossibility of *identifying* criminal public authority and criminal agent.

Following the first argument, Kant targets the intentionality of the offender, considered in Beccaria's theory. Punishment is, by its very definition, something that is attributed *against the will* of the offender, according to the close connection of punishment with suffering: from this perspective

> no one suffers punishment because he has willed *it* but because he has willed a *punishable action* (*strafbare Handlung*); for it is no punishment if what is done to someone is what he wills, and it is impossible *to will* to be punished.[84]

Beccaria proposes a paradoxical and therefore impossible ground for the illegitimacy of the death penalty, namely a 'promise to let oneself be punished and so to dispose of oneself and one's life'.[85] If the foundation of the right to punish was the criminal's promise to be punished, he or she should also be responsible for finding himself or herself guilty, thus becoming his or her own judge. If the co-legislator was also the criminal (albeit only potentially so, as a person capable of punishable acts) – that is, if the individual as punishable actor was the one to express a desire to be punished as per the social contract – then even the judge, who applies criminal law in court, would be considered a punishable individual. From a Kantian perspective, Beccaria's mistake is precisely that of presuming that the criminal is at the same time the co-legislator of criminal law.

Indeed, the second argument of the Kantian refutation outlines precisely the central distinction between, on the one hand, the subject as co-legislator (*Mitgesetzgeber*) who dictates the criminal law and, on the other, the subject as potential criminal – subject to that law and not entitled to challenge it:

> (w)hen I draw up a penal law against myself as a criminal, it is pure reason in me (*homo noumenon*), legislating with regard to rights, which subjects me, as someone capable of crime and so as another person (*homo phaenomenon*), to the penal law, together with all others in a civil union.[86]

Without this distinction, the impartiality of the judge, the sacredness of public law and the very distinction between the three powers of the State, a central element of Kant's doctrine of law, would be lost.

The person who issues criminal law cannot be the very same person who is

[84] 6:335.
[85] ibid
[86] 6:335.

punished as a subject under that law, because as a criminal it is not possible to participate in the act of legislation. When the subject creates a criminal law against himself as a criminal, it is his *homo noumenon* (endowed with pure reason) who submits him or herself to a criminal law as *homo phenomenon* (capable of crime).[87]

V. Prudence and the Preventive-Specific Function of Punishment

As it clearly emerged, the principle of retribution (*Vergeltung*) – what today would be called the principle of legality – plays a central role in the text of the *Doctrine of Right*. But as we have noted, with the prohibition of an instrumental punishment, on the one hand, and with the principle of equality on the other, Kant has established a relationship of priority of the element of retribution over the preventive element of the punishment: of its vindictive character turned towards the past (*quia peccatum est*) over its exemplary character of utility for the future (*ne peccetur*).

To a certain extent, Kant considers inevitable that human beings in their social life are (and should be) used as a means. With the imposition of a punishment, which by its very definition indicates a type of coercion, an obligation directed against the will of a person, the individual is inevitably *also* treated as a means. The important thing – and this is the central point of Kant's doctrine – is that the accused should *not only* be treated as a means, but *always also* as an end, and that this second aspect should be given priority over the first. Kant therefore does not exclude that the penalty could *also* play a deterrent role, and therefore be aimed at the prevention of a crime, looking towards the future. However, this type of pragmatic assessment must be systematically and clearly distinguished from the *a priori* investigation of metaphysical first principles of the doctrine of right. This emerges clearly from an important footnote from the already mentioned *Appendix*:

> Punitive justice (*iustitia punitiva*) must be distinguished from punitive prudence, since the argument for the former is moral, in terms of being punishable (*quia peccatum est*) while that for the latter is merely pragmatic (*ne peccetur*) and based on experience of what is most effective in eradicating crime; and punitive justice has an entirely different place in the topic of

[87] In his *Reflexionen zur Rechtsphilosophie* Kant writes: '*Die Strafe muß in dem Gesetze selbst bestimmt werden und zwar nicht um der Verbrecher sondern des Publici und ihrer Freyheit willen in Ansehung der Willkühr des Richters. Sonst dem Verbrecher kann nicht Unrecht geschehen*' ('A punishment must be determined within the law itself and not by the will of the criminal, but rather by the will of the public and its freedom in view of the deliberation of the judge. Otherwise no harm can be done to the criminal') (*Reflexionen* no 7995, 19:576).

concepts of right, *locus iusti*.[88]

With this distinction, Kant provides the fundamental *systematic collocation* of two functions of the penalty: the retributive or vindictive function, addressed solely to the fact committed and belonging to the field of the first principles of right (*a priori*), and the specific deterrent function, which instead is the fruit of a calculation on the preventive effectiveness, in the future, of the penalty – a pragmatic and empirical consideration (*a posteriori*).

The concern about the preventive-specific function of the penalty is therefore, for Kant, wholly legitimate, and it introduces the problem of legal reform – although only from a pragmatic point of view, one pertaining neither to the question of the foundation of the right to punish, nor to the questions of the 'if' and the 'how' of punishment.

In a passage of the *Reflexionen zur Rechtsphilosophie* Kant is very clear on this point, and sets out the three basic aims that criminal justice is called to assess from a pragmatic point of view:

> *Die iustitia punitiva hat zur Absicht: 1. den Unterthan aus einem schlimmen in einen besseren Bürger umzuwandeln; 2. durch warnende Beyspiele andere abzuhalten; 3. unbesserliche aus dem Gemeinen Wesen, es sey durch deportation, exilium, oder Tod wegzuschaffen (ob durch Gefängnis). Aber alles dieses ist nur Klugheit der politick. – Das Wesentliche ist die Ausübung der Gerechtigkeit selbst alsdenn noch, wenn die Verfassung aufgehoben würde. (g Ob auch Experimente mit Missethätern der medicin halber gemacht werden dürfen.*[89]

The political art (*Staatskunst*) of criminal prudence (*Strafklugheit*) amounts to the ability (*Geschicklichkeit*) to choose a punishment based on the calculation of its future utility: this is *simply* pragmatic and it is based on the experience of what is most effective, in certain circumstances, to dissuade the potential offender from performing illegal actions.

The fact that a punishment may also have a preventive utility, therefore, is not problematic for Kant: it is only problematic when the preventive aspect becomes exclusive and primary. This is also confirmed by a preparatory manupscript to the *Doctrine of Right*, which focuses precisely on the vindictive, but also *educational*, character of punishment:

[88] 6:363-364.

[89] 'The punitive iustitia has as its purpose: 1. to transform the subject from an evil citizen to a better citizen; 2. to dissuade others through warning examples (*durch warnende Beyspiele*); 3. to export from the common body those who cannot improve, whether by deportation, exile, or death (or by imprisonment). *But all this is only prudence of politics*. The essential thing is always the exercise of justice itself, even if the constitution were to be abrogated' (*Reflexionen* no 8035, 19:587-588).

> *Die Strafe ist ein actus der öffentlichen Gerechtigkeit also des Oberen im Staat gegen den Untergebenen ihm ein Übel zuzufügen was der Läsion gemäs ist die er an einem Anderen (Bürger, passiv oder activ) begangen hat. Sie ist an sich jederzeit rächend kann aber auch mit der Absicht den Verbrecher zu bessern verbunden seyn.*[90]

This passage is crucial to highlight how Kant borrows the *ius talionis* and the vindictive character of the criminal institution from medieval law – even though he distances himself considerably from it by insisting that the right to punish belongs exclusively to the ruler and to the organs of the state, even with all the limitations it encounters. At the same time, however, he also declares that a punishment can be associated with the intention of improving the offender, thus making a significant contribution to the historical transition towards modern law.[91]

This aspect of punishment *qua* warning emerges several times in the context of his *Reflexionen zur Moralphilosophie*:

> *Alle Strafen sind entweder warnende Strafen oder rächende; poenae exemplares, wenn sie nicht diesen gemäs seyn, sind politisch.*[92]

In another reflection, Kant writes:

> *Warnende oder rächende Strafen. (s deterrentes vel vindicativae) … poena est vel exemplaris vel animadversio vel vindicativa.*[93]

Or again, Kant distinguishes pragmatic and moral punishments as follows:

> *Pragmatische Strafen sind warnend und gehen auf das äußere der handlung, moralische auf böse Gesinnung.*[94]

The preventive and specific function of punishment emerges clearly in the pragmatic field. In addition to a general-preventive theory, implicit in the role of punishment *within* the juridical state, and to the retributive function, closely linked to the question of punishability of 'public crimes' and the criterion of punishment, Kant also considers the question of the *usefulness of the*

[90] 'Punishment is an act of public justice therefore of the superior (power) in the state against the subordinate in order to inflict upon him an evil in accordance with the injury he has committed to another (citizen, passive or active). In itself (the punishment) is always vindictive, but it can also be connected to the intention to improve the criminal' (*Vorarbeiten*, 23:343).

[91] Cf also *Vorarbeiten*, 23:347.

[92] 'All punishments are either warning-meaning or vindictive punishments; poenae exemplares, when they are not proportionate, are political' (*Reflexionen* no 6526, 19:56).

[93] 'Deterrent or vindictive punishments (*deterrentes vel vindicativae*) (...) *poena est vel exemplaris vel animadversio vel vindicativa*' (*Reflexionen* no 6527, 19:56).

[94] 'Pragmatic punishments are warning (*warnende*) and are directed towards the exteriority of action, moral punishments towards bad intention (*Gesinnung*)' (*Reflexionen* no 6681, 19:132).

punishment. He therefore considers punishments in their deterrent, preventive-specific, function, but only under the condition that such a pragmatic point of view should be properly distinguished from that of the first principles of the doctrine of right.

Short Symposium on the Punishment

The Necessity for Punishment in Hegel as a Right of Freedom

Sabina Tortorella*

Abstract

The article presents the theory of punishment in the *Elements of Philosophy of Right* focusing on Abstract Right and Administration of Justice. The first part of the essay underlines how punishment allows restoration of the universality of right and plays a role of education to the universal, directed against the natural and immediate will. Through reference to Eumenides' tragedy, the second part points out the limits of Abstract Right in order to then focus on Civil Society, in which, thanks to a court and a trial, punishment is the real conciliation. The conception of punishment reflects the status of the different moments of Objective Spirit, since, while in Abstract Right Hegel sets the problem of the rational foundation of punishment, in Civil Society he questions himself about its purpose and its applicability. This leads not only to rediscussing the opposition between the retributivist interpretation on the one hand and the utilitarian one on the other, but also to highlighting that punishment constitutes the key to access the issue of the validity of right as well as to identify its contradictions.

I. Introduction

The Hegelian conception of punishment has deeply influenced German criminal law up to the most recent times. Although in alternating phases, the legacy of the German philosopher played an important role in Germany and inspired a good number of jurists throughout the 19th century and early 20th century. Even when they disagreed or totally rejected the Hegelian theses, they still showed deep concern for them. Without Hegelianism and its reception by jurists, though to different degree, it would not be possible to understand either the present, or the history of penal legal science.[1] In the field of philosophical literature the topic of punishment is a litmus test of all those prejudices that long characterised Hegel's figure as a statist thinker who reduces the complexity of the existing to a logical and metaphysical structure. According to this interpretation, Hegel would also apply the triadic model of dialectics to the

* PhD, Ca' Foscari University of Venice and Paris 1 Panthéon-Sorbonne.

[1] See M. Kubiciel et al, *Hegels Erben*? (Tübingen: Mohr Siebeck, 2017). About the reception of the theory of punishment in Hegel see also A. von Hirsch et al, *Strafe – Warum? Gegenwärtige Strafbegründungen im Lichte von Hegels Straftheorie* (Baden-Baden: Nomos, 2011) and W. Schild, 'Verbrechen und Strafe in der Rechtsphilosophie Hegels und seiner "Schule" im 19. Jahrhundert' *Zeitschrift für Rechtsphilosophie*, 30-42 (2003).

conception of criminal law, conceiving the State as a superior entity that sacrifices the freedom of individuals. A famous essay of 1968 thus wishes for a final farewell to Kant and Hegel, who should have nothing more to say on the subject of criminal law: not only would the Hegelian argument be 'a pure logical error' or an 'empty statement', but it would coincide with a 'pure metaphysical fantasy', filled with 'irrational lyric-philosophical excesses'.[2]

The value given to the reflection on punishment pinpoints the state of Hegelian studies in general: starting from that movement aimed at a *Rehabilitierung der praktischen Philosophie* already in the 1970s and more recently in the context of the *Hegel Renaissance* which characterised Hegelian philosophy as a whole, even the Hegelian penal conception received renewed interest. Hegel's new depiction as a thinker of freedom and a theorist of recognition and no longer a conservative and totalitarian philosopher had the effect of questioning the long-standing interpretation among interpreters, according to which the Hegelian conception of punishment was a retributivist theory, justifying punishment by basing it on the principle *quia peccatum est*, or as a response to the evil that was done.[3] Numerous studies have been carried out in recent decades alongside this interpretation aiming at pointing out the utilitarian aspects that are present in the Hegelian conception: Hegel's theory of punishment is thus interpreted as a combination of special prevention and rehabilitation,[4] as well as minimal specific deterrence aimed at the resocialisation and reforming of the criminal[5] or again as a general deterrence, in which punishment aims at the restoration of a legal community.[6] Depending on the text passages we take into consideration and following the publication of the lectures in philosophy of right, the critique highlighted how Hegel's theory of punishment is not only – and not really – a retaliatory theory, as it takes into consideration the aspect of dangerousness to society in general, aims at reintegration of the criminal and has a corrective role. For this reason, Hegelian penal theory is thus presented also as a unified theory,[7] which also includes a

[2] U. Klug, 'Abschied von Kant und Hegel', in J. Baumann ed, *Programm für ein neues strafgesetzbuch* (Frankfurt A.M.: Fischer Bucherei, 1968).

[3] For a review of the interpretations of punishment in Hegel see J.-C. Merle, 'Was ist Hegels Straftheorie?' *Jahrbuch für Recht und Ethik*, 11, 145-176 (2003).

[4] G. Mohr, 'Unrecht und Strafe', in L. Siep ed, *Klassiker Auslegen. G.W.F. Hegel, Grundlinien der Philosophie des Rechts* (Berlin: Oldenbourg Akademieverlag, 1997), 95-124.

[5] J.-C. Merle, 'La complexité de la théorie non rétributiviste du droit pénal de Hegel', in J.-F. Kervégan and G. Marmasse eds, *Hegel penseur du droit* (Paris: CNRS Éditions, 2004), 81-96. See also J.J. Kominkiewicz et al, *German Idealism and the concept of Punishment* (Cambridge: UP, 2009), 107-146.

[6] K. Seelmann, *Anerkennungsverlust und Selbstsubsumtion. Hegels Straftheorien* (Freiburg-Munchen: Karl Aber, 1995).

[7] W. Schild, 'The Contemporary Relevance of Hegel's Concept of Punishment', in R.B. Pippin and O. Höffe eds, *Hegel on Ethics and Politics* (Cambridge: Cambridge University Press, 2004), 150-179.

deterrent or rehabilitative function.[8]

In any case, punishment is undoubtedly a recurring topic in Hegelian thought, as it is present in all his practical-political writings, from the juvenile fragments to the years of Jena up to the *Elements of the Philosophy of Right*. The richness of the Hegelian reflection in this context is evident even if we only focus on the text of 1820, since in order to delineate a theory of punishment, it is necessary to go through the whole work: if in Abstract Right the subject of punishment is addressed after the wrong, it is in Morality that Hegel defines the criteria of imputation and in Ethical Life that he describes the functioning of the penal trial and the role of the court. The reasons for such an ambivalent judgment by scholars towards the Hegelian penal conception have to be linked to the fact that in *Elements,* in order to define what punishment is, Hegel uses the lexicon of his philosophy, in particular he uses the most dense and meaningful terms of speculative thought. For example, punishment is defined as *Aufhebung des Verbrechens*, 'cancellation of the infringement',[9] or *Versöhnung des Rechts mit sich selbst*, 'reconciliation of right with itself',[10] and again as *Verletzung der Verletzung*, 'infringement of an infringement',[11] *Vernichtung jener Verletzung*, 'nullification of the infringement'[12] and finally *Negation der Negation*, 'negation of the negation'.[13]

The importance given to the concept of punishment is all the more surprising when one considers Hegel's definition of right, which is presented as 'the realm of actualized freedom' and coincides with 'the existence of the free will', with 'freedom as Idea'.[14] Therefore, it is possible to identify an apparent paradox according to which the person who identifies right and freedom is the same who not only admits the possibility of punishment, but also makes it a necessary aspect of right. How is it possible to state that right does not constitute a limitation on freedom, but on the contrary its realisation, and at the same time justify a form of constraint that may appear as a denial of freedom? In other words, if right is not a restriction of freedom and freedom does not coincide with the arbitrary will, then what is the foundation on the basis of which Hegel introduces the possibility of repression? If for the authors who consider right as

[8] T. Brooks, 'Is Hegel a retributivist?' *Hegel Bulletin*, 25, 113-126 (2004); Id, 'Hegel and the Unified Theory of Punishment', in T. Brooks ed, *Hegel's Philosophy of Right* (London: Blackwell, 2012). For a non-retributivist interpretation of punishment in Hegel see also S. Moccia, 'Contributo ad uno studio sulla teoria penale di G.W.F. Hegel' *Rivista italiana di diritto e procedura penale*, 27, 131-174 (1984); S. Fuselli, 'La struttura logica della pena in Hegel' *Verifiche*, 28, 27-106 (1999); P. Becchi, 'Il doppio volto della pena in Hegel' *Verifiche*, 28, 191-209 (1999).

[9] G.W.F. Hegel, *Elements of the Philosophy of Right*, edited by A.W Wood, translated by HB Nisbet (Cambridge: Cambridge University Press 1991), § 98, 124. From now on abbreviated as *Rph*.

[10] ibid § 220, 252.

[11] ibid § 101, 127.

[12] ibid § 97, 123.

[13] ibid § 97, 123.

[14] ibid §§ 29, 4 and 58, 35.

regulation of external behaviour or organisation of force, introducing constraint is the way to ensure the effectiveness of right, then for Hegel who makes the right of freedom the cornerstone of his legal theory, it seems apparently contradictory to conceive a duty of coercion.

The paradox seems even more evident if we think that in Abstract Right punishment is explicitly related to the topic of justice, since Hegel explicitly states that the different modalities of infringing right at the same time raise the question of the 'objective consideration of justice'.[15] Far from being presented as a pathological and marginal phenomenon, the problem of wrong is addressed by Hegel, focusing on the relationship between punishment and justice and between right and freedom: precisely as a transgression of right, a crime represents an experience of injustice and punishment as the annihilation of the non-right constitutes the denial of injustice. Instead of starting from a definition of justice and therefore from a substantive and essentialist conception of it, Hegel reverses the reasoning: if denying right means doing violence against what has value and the meaning of fair, punishment acquires the role of restoration of what is *Richtig* and it is therefore an act of justice.[16] Wrong, *Unrecht*, is not only the reverse of *Recht*, that is the denial of right and therefore its opposite, but an unavoidable moment of the juridical as it allows its realisation. Consequently, punishment is not only a tool aimed at denying the transgression and restoring the validity of legal principles, but it is also what allows us to highlight its finitude and its limits and it is therefore intrinsically linked to the broader topic concerning the definition of right as freedom.[17] From this point of view, crime and punishment are two sides of the same coin: if right cannot fail to admit its maximum denial, or crime, the moment of coercion and punishment is not a necessary evil, but a constitutive aspect of the juridical, which makes it possible to examine the statute of right, its foundation and legitimacy as well as its applicability and effectiveness. As was pointed out, Hegel rejects any attempt to moralise punishment, since the field of penal treatment is properly juridical and constitutes the key to access the issue of the validity of right, to question its immediacy as well as to highlight its contradictions, above all in the gap between universality and application.

So far, only the aspects that are related to the necessity of punishment and are defined as subjective have mainly been highlighted, since attention was focused on the abrogation of the crime starting from the perspective of the particular will of the offender, while it is necessary to take into consideration the juridical necessity for punishment from an objective point of view, since it coincides with restoration of the juridical universal. Punishment must be

[15] G.W.F. Hegel, n 9 above, § 99, 125.

[16] See J.-F. Kervégan, 'La théorie hégélienne de la justice', in P. David and B. Mabille eds, *Une pensée singulière* (Paris: Harmattan 2003), 101-113.

[17] See on this point M. Foessel, 'Penser la peine' *Revue de Métaphysique et de Morale*, 4, 529-542 (2003), in particular 530.

justified precisely because the very possibility of admitting right is at stake. If for Ricoeur Hegel's choice to place the discussion of punishment within Abstract Right, that is outside of Ethical Life, would seem to demonstrate Hegel's intention to contribute to the 'deconstruction of the myth of punishment', this choice was actually born from the intention to question the rationality of the latter and therefore to lay the foundations outside its concrete application.[18] In spite of the *Elements'* scan, the topic of punishment makes it possible to identify a common thread inside the work, which directly relates to two sections that are apparently very distant from each other, but actually characterised by continuous references: Abstract Right and Administration of Justice in Civil Society. If in Abstract Right Hegel deals with the rational foundation of punishment, he shows at the same time its limits, stating the need to insert the same penal law within the proper institutional context of Ethical Life. The age-old problem that opposes a retributivist Hegel on the one hand and Hegel as an advocate of the corrective function of punishment on the other can hence be solved through a reading that takes simultaneously into account the paragraphs of Abstract Right and the ones of Civil Society as textual sources to fully reconstruct, in a unitarian way, the penal theory in Hegel.[19] According to the perspective of a 'philosophical science of right' which Hegel's work wishes to be, indeed the discussion of punishment has to be inscribed within the developmental stage of the 'Idea of right', which implies, as is well-known, 'the concept of right and its actualization',[20] and therefore it has to be read in light of the different moments of the Objective Spirit: according to the formal and abstract point of view, which is the concept of right belonging to *abstraktes Recht*, and according to the point of view belonging to the historical-social reality that characterises *Rechtspflege*.

To this aim, a first part of this essay will focus on the concluding paragraphs of Abstract Right in order to then show the limits of this perspective and focus on its discussion within Civil Society. If Ricoeur stresses how the logic of punishment is 'a logic without myth',[21] it is possible, on the other hand, to highlight how, according to Hegel's use of the Greek tragedy, as in the case of Antigone and Creon in *The Phenomenology of Spirit*, it is precisely the use of the myth of the Eumenides that allows him to directly link Abstract Right to Administration of Justice. Therefore, this essay aims to show first that, depending on the textual passages that are examined, it is possible to identify a double justification of

[18] P. Ricoeur, 'Interprétation du mythe de la peine', *Le Conflit des interprétations* (Paris: Seuil, 1969), 354.

[19] The relevance of the paras about Civil Society with respect to the theory of punishment was highlighted by: D. Klesczewski, *Die Rolle der Strafe in Hegels. Theorie der bürgerlichen Gesellschaft. Eine systematische analyse des Verbrechens – und des Strafbegriffs in Hegels Grundlinien der Philosophie des Rechts* (Berlin: Duncker & Humblot, 1991) and K. Seelmann, *Le filosofie della pena di Hegel*, in P. Becchi ed (Milano: Guerini e Associati, 2002).

[20] G.W.F. Hegel, n 9 above, § 1, 25.

[21] P. Ricoeur, n 18 above, 360.

punishment by Hegel in order to then underline as a conclusion, how both arguments converge in a single objective: to legitimise the duty of punishing as a right of freedom and to show how both in Abstract Right and in Civil Society punishment accomplishes the task of overcoming the opposition between universal will and particular will.

II. Coercion as Retribution in Abstract Right

Abstract Right constitutes, as is well known, the first part of *Elements of the Philosophy of Right* and it corresponds to the first stage in the process of objectification of freedom since 'the will which is free in and for itself' is 'in the determinate condition of *immediacy*'.[22] The notion of the person is the protagonist, which indicates 'the universality of this will which is free for itself' and 'the simple reference to itself in its individuality'.[23] Defined by Hegel as capacity for right in general, the person is 'the inherently individual will of the *subject*', 'a consciousness of itself as a completely abstract "I" '.[24] The fundamental rule of Abstract Right can be exemplified by the imperative that proclaims 'be a person and respect others as persons!' and translates on the juridical side into a 'permission or warrant'.[25] At this level, right is formal because it just recognises the universality of the person, ie of the subject of right, without taking into consideration a particular interest, or the intention, or motive of the action. It does not establish what can be done, nor what should be done, but what should not be done, in such a way that the only obligatory constraint that is envisaged is to refrain from violating the person.[26]

Abstract Right presents the characters that refer to an original condition, outside of any social relationship and in the absence of political power. However, this is not an ideal moment placed at the origin of history, as is the state of nature for contractualist thinkers, but rather it presents the rational principles of private right. Abstract Right, on the one hand, is the result of the historical process, but, on the other, it presents its categories as if they were independent of history, in the same way as a rational normative order. Only the evolution of the spirit recognised the universality of individual freedom and of the juridical capacity, but in the modern world they have become a foundation that has been removed from political bargaining, thus assuming features that not only place them outside the historical dimension, but that make them an

[22] G.W.F. Hegel, n 9 above, § 34, 67.
[23] ibid § 35, 68.
[24] ibid §§ 34-35, 67-69.
[25] ibid §§ 36-37, 69.
[26] In Abstract Right the person is not yet a moral subject and therefore the aspects related to the proposal and responsibility of the action are not taken into consideration, but they will be discussed within morality, in particular in paragraph 132, where Hegel introduces the notion of imputation.

element of legitimacy of the established order. Abstract Right is therefore simultaneously the result of universal history and the formal and universal prerequisite of modern ethical life.

After presenting the institutions of private right such as property and contract, the third section of Abstract Right, which is entitled 'Coercion and Crime', corresponds to 'coercive right'[27] and coincides with 'the sphere of *penal law*'.[28] Wrong arises because of an opposition between the particular will of one of the contracting parties and common will, as a result of the agreement of the particular wills of the owners that was reached in the contract. In Abstract Right, particular will complies with universal will, therefore it complies with right, only accidentally and consequently 'right *in itself* is present as something *posited*', whose universality presents itself *als Gemeinsames*.[29] The immediacy of the person is initially established as identical to the universal and therefore corresponds to common will, but, by a kind of retaliation, the same person then discovers himself as particular, contradicts common will and breaks the pact. Wrong determines a detachment between 'right *in itself* or the will as universal *in itself*', and 'right in its *existence*, which is simply the *particularity* of will'.[30] Wrong is the negation of that contractual relationship which aspired to mediate, albeit in a still apparent form, particular will and universal will and represents the radical opposition between the particularity of the person and the universality of right:[31] the particular will takes the place of universality and therefore the criminal denies the common will, behaving as if he himself was right, or as if his particular will coincided with the universal. The limit of Abstract Right is therefore on the one hand the lack of a coercive force that is capable of ensuring respect for the contract and on the other hand, the way in which subjective will relates to universal.

Hegel presents three types of wrong, even if the most serious is the crime, *Verbrechen*, which constitutes infringement of right as such, that is the denial of the right in its very formality. Unlike unintentional wrong and deception,

[27] ibid § 94, 121.

[28] ibid § 95, 122. On this aspect G. Mohr insists, n 4 above. On the contrary, G. Marmasse, 'Hegel et l'injustice' *Les Études philosophiques*, 3, 331-340 (2004), for whom the third part of Abstract Right is the conflict between the fair and the unfair and throughout the notion of fair distribution presents some similarities with the Aristotelian conception.

[29] G.W.F. Hegel, n 9 above, § 82, 115.

[30] ibid § 81, 113.

[31] Hegel states that the right is *Erscheinung*, appearance, as an immediate agreement between essence and existence, between right and particular will, as shows with respect to the contract. With wrong, right is the semblance, *Schein*, as the identity and the agreement between right and particular will is lost and the same right receives *die Form eines Scheines*. By referring to the categories of the logic of essence, Hegel underlines the inadequate relationship of mediation between the terms involved: if, with respect to the exposition of the *Science of Logic,* the order is reversed, it is because Hegel wants to stress how in common will right still presents itself as immediacy and wrong coincides with a moment of involution and retreat with respect to the identity gained in the contract.

crime does not only concern the value of property or of an asset, the object of a dispute between one individual and another, but damages right as a right, because it is a violence against the person and his freedom. In this case punishment is necessary,[32] indeed, it not only has the aim of intervening with respect to the specific episode, for example by establishing a compensation or redress, but therefore has a universal vocation that is completely absent from the civil law dimension. The violence of the crime is consummated 'directed against the existence of my freedom in an external thing' and therefore it 'infringes the existence of freedom in its concrete sense', that is 'right as right':[33] a crime is an action in which 'not only the particular – ie the subsumption of a thing under my will' is negated, as happened in the case of civil wrong, but 'also the universal (...) my capacity for rights' and it does not happen without 'the mediation of my opinion', that is, without my knowledge, as happened with fraud.[34] Thus, crime denies both the subjective and the objective side: the particular right of the person, but at the same time more generally the universal sphere, ie the juridical capacity of the person and therefore the presupposed recognition of the right, since his freedom is not recognised, which is an implicit condition in every legal relationship.[35] It coincides with a form of violence against another will, aimed at denying the possibility of the manifestation and externalisation of subjective freedom: with crime, says Hegel, 'the principle of will is attacked', as committing a crime against someone means not admitting or denying that someone has a right.[36] The criminal action is presented as contradictory, in that it denies the very possibility of giving a legal relationship. By being itself a substitute for universal will, the criminal claims to be a substitute for the right, but thus ends up denying de facto the juridical bond: a crime is not simply a refusal to be submitted to legal constraints, it is instead the criminal's claim to deny the juridical relationship as such, so that at first he refuses the right to then direct his own action to the same universality of the right he had previously denied.

[32] ibid § 97, 123. Concerning the studies on Hegel's penal conception, we refer, among many works to O.K. Flechtheim, *Hegel's Strafrechtstheorie* (Berlin: Duncker & Humblot, 1975); R. Hohmann, *Personalität und strafrechtliche Zurechnung* (Frankfurt A.M.-Berlin: Peter Lang, 1993); P. Stillman, 'Hegel's Idea of Punishment' *Journal of History of Philosophy*, 14, 169-182 (1976); A. du Bois-Pedain, 'Hegel and the Justification of Real-world Penal Sanctions' *Canadian Journal of Law & Jurisprudence*, 29, 37-70 (2016).

[33] G.W.F. Hegel, n 9 above, §§ 94-95, 121.

[34] ibid § 95, 121-122.

[35] Since this is not the right place to go in depth in the relationship between recognition and Abstract Right, it is sufficient to quote Hegel himself, who states that 'Contract presupposes that the contracting parties recognise each other as persons and owners of property; and since it is a relationship of Objective Spirit, the moment of recognition is already contained and presupposed within it', G.W.F. Hegel, n 9 above, § 71, 103.

[36] G.W.F. Hegel, *Lectures on Natural Right and Political Science. The First Philosophy of Right*, transcribed by P. Wannenmann, translated by J.M. Stewart and P.C. Hodgson (Oxford: Oxford University Press, 1995), § 45, 96.

Crime is an infringement of right as it coincides with the unilateral affirmation of particular will and testifies to the latter's failure to adhere to the right. If the crime corresponds to force and it is defined as *Zwang or Gewalt*, as *Gewalt überhaupt*[37] or even in the *Encyclopaedia*, as *unmittelbaren Zwang*,[38] punishment is defined first in § 93 as *zweiter Zwang*: while then the first coercion is simple force against my freedom and as such it is *unrechtlich*, that is to say contrary to right, the second coercion tends to delete the previous force through a constraint, in such a way that force represents in this second moment a properly juridical element.[39] On the one hand with the wrong my will 'may either experience force in general' or 'it may be forced to sacrifice or do something',[40] on the other hand punishment enacts 'a force which supersedes the original one'.[41] In this perspective, extra-juridical force as a power to materially prevent freedom is followed by properly legal force, a tool which the same right uses in order to reaffirm itself. If force is the anti-right, since committing violence means placing oneself outside the juridical horizon, the punishment as juridical coercion is *Zwang* and not only *Gewalt*: for sure this is a form of force, because it involves coercion, but it does not exhaust itself in the simple production of evil or suffering. On the one hand then, we can say that violence is not only what distinguishes the pre-juridical reality, as it receives full legitimacy within the juridical field, but on the other hand it appears in right as a means and not an end, as the use of coercion is the *modus operandi* of the right as such. Punishment is force used and directed against another force, in such a way that it is the right to coerce, *zwingen*, through the use of *Gewalt*.[42] *Zwang* is the form of force belonging to the legal field and an instrument of reaffirmation of right.[43]

Instead of being justified on the basis of utilitarian or instrumental reasons, as is the case of the Enlightenment theories, punishment is founded, for Hegel, on the idea that 'punishment in and for itself is just'.[44] This means that it must

[37] G.W.F. Hegel, n 9 above, § 90, 119.

[38] G.W.F. Hegel, *The Encyclopedia of the Philosophical Sciences (1830)*, translated by W. Wallace (Oxford: Clarendon Press, 1971).

[39] G.W.F. Hegel, *Elements* n 9 above, §§ 92-94, 120-121.

[40] ibid § 90, 119. The German term *Gewalt* is translated into English mainly with the word 'force' and sometimes with 'violence' while the term *Zwang* is translated with 'coercion'.

[41] ibid § 94, 121.

[42] ibid § 90, 119.

[43] See L. Marino, 'Violenza e diritto in Hegel' *Rivista di filosofia*, 205-233 (1977). The author refers to the definitions of *Zwang* and *Gewalt* that can be found in J.H. Campe, *Wörterbuch der deutschen Sprache*: *Zwang* means 'der Zustand, da die freien Handlungen eines Wesens durch Gewalt eingeschränkt werden, es möge diese Gewalt eine körperliche oder sittliche sein', while *Gewalt* 'gehet hier vor Recht'. As Marino stated, the specific scope of the force is thus explicitly identified with the extra-juridical or pre-juridical one.

[44] G.W.F. Hegel, n 9 above, § 99, 125. Hegel criticises the Enlightenment criminal theories such as those of Klein, Beccaria and Feuerbach: if the former confuses injustice with evil, thus giving a moral value to the juridical categories, the limit of the second lies in the use of natural law

be justified not only from the point of view of universal will, ie from the point of view of universal represented by right, but also starting from the perspective of the criminal, as a result of his own will. In paragraph 100 Hegel states that 'the injury which is inflicted on the criminal' is 'a right posited in his existent will', it is 'his right', since through punishment he 'is honoured as a rational being'.[45] The Hegelian strategy consists of justifying punishment as a sort of right of the criminal, who has the duty to be submitted to the punishment as a necessary consequence of the behaviour itself carried out by him, since his action implicitly has a universal value and therefore implies that the same action is subsumed under the same law. This is a Kantian topic that follows the principle of the universalizability of the maxim:[46] if at the moment in which he formulates a norm that is useful to act, he is obliged to be subject to it, it can therefore be said that, in the case he steals, he should himself be liable to theft of property. By stealing he claimed a universal rule, which consists of rejection of private property and which becomes a source of obligation even for him. The criminal disregards the individual as a subject of right, he violates the principle that characterises cohabitation and the objective dimension, placing himself above the juridical and social bond, since with his action he states the principle by which 'it is allowed to harm the will'.[47] On the one hand, the crime denies 'the recognition of the right to the universal and deciding factor' and therefore rejects the intersubjective relationship of reciprocity starting from which the juridical dimension is generated, since it claims to manifest its freedom as a natural and particular will. On the other hand, punishment is what makes the offender a victim and restores the recognition relationship.[48]

arguments, since he grounds the lawfulness of the sentence on the basis of an implicit consent of the offender which the latter would have granted on the basis of the social contract. Finally Feuerbach gives a dissuasive psychological value to the sentence, transforming the penal sanction that follows the crime into an anticipated threat aimed at intimidating and removing the drive to violate the law, thus ending up treating the human like a dog.

[45] ibid § 100, 126.

[46] In this regard, Seelmann identifies in Abstract Right two distinct justification strategies of punishment: the first set out in para 97 is what he calls the argument from recognition, the second, in para 100 is what he defines the argument from the law, (n 6 above). K. Seelmann, 'Does Punishment Honor the Offender?', in A.P. Simester et al eds, *Liberal Criminal Theory: Essays for Andreas von Hirsch* (Oxford: Hart, 2014), 113-114. According to Foessel, the definition of punishment as a criminal's right inaugurates the field of morality, since the criminal becomes aware of himself as subjectivity (M. Foessel, n 17 above, 538).

[47] G.W.F. Hegel, *Vorlesungen über Rechtsphilosophie nach der Vorlesungsnachschrift von H.G. Hotho 1822/1823*, in K.H. Ilting ed, *Vorlesungen über Rechtsphilosophie 1818/1831* (Stuttgart-Bad Cannstatt: Fromman-Holzboog, 1973), III, § 100, 316.

[48] G.W.F. Hegel, *Elements* n 9 above, §§ 84-85, 117. Regarding the relationship between recognition, crime and punishment K. Seelmann, 'Wechselseitige Anerkennung und Unrecht' *Archiv fiir Rechts und Sozialphilosophie*, 79, 228-236 (1993) and L. Siep, 'Anerkennung, Strafe, Versöhnung. Zum philosophischen Rahmen von Hegels Strafrechtslehre', in M. Kubiciel ed, n 1 above, 7-28; L. Di Carlo, 'Il riconoscimento nella *Filosofia del diritto* di Hegel' *Teoria politica*, 145-154 (2003). For Fossell crime shows the insufficiency of the recognition as it is realised in Abstract Right (n 17 above, 533).

As we read in a margin note, the coercion that characterizes punishment is a force against a natural being.[49] In this perspective, declaring that the force of right is a power that is enacted as a reaction with respect to wrong means simultaneously stating that the force of right is directed against the affirmation of a will that is still natural, immediate and abstract, in such a way that this violence constitutes an element by which the will itself is formed to the universal. Thus, legal force allows us to go beyond the natural horizon, so that it can be understood as education to obedience, formation to the universality incarnated by right: that same *Gewalt*, which could be considered as oppression, abuse of power and injustice, acquires in the context of Abstract Right the value of a culture, which allows us to access spiritual universality. Punishment is juridical coercion and it is presented and justified first of all as a means by which it is possible to restore right, as it allows one to manifest 'its necessity which mediates itself with itself through the cancellation (*Aufhebung*) of its infringement',[50] ie to give right 'the determination of something fixed and valid', since from being immediate it can become 'actual as it returns out of its negation'.[51] In the same way, from the point of view of the person, the right as *Zwang* is simultaneously both a foreign power, which imposes itself as violence, and a process of liberation by which consensus and obedience to the right are developed.

The infringement of right is something negative because it exists only as an activity that denies, only in relation to something else, that is right, and brings with it the same criminal since 'the positive existence of injury consists solely of the particular will of the criminal'.[52] Therefore, if the only thing that has a positive value in the context of wrong is particular will, it can be said that the injustice corresponding to the denial of right is precisely the particular will against which right must carry out its work of integration towards the universal in opposition to the particular and natural drive. If right remains 'an obligation', an ought, and therefore it is abstract, as long as 'the will is not yet present as a will which has freed itself from the immediacy of interest in such a way that, as particular will, it has the universal will as its end',[53] the penal right implements a coercion in order to overcome the naturalness of the will and therefore uses punishment as *Bildung* against the particularity that resulted in the violation of the right.

Hegel then explains that while a crime is a *Verletzung des Rechts* and therefore it is an existence within itself null, which exists only as far as it denies right, punishment is the 'manifestation of its nullity' and therefore the *Vernichtung* of the *Verletzung* itself.[54] As well as crime having a claim to

[49] G.W.F. Hegel, *Grundlinien der Philosophie des Rechts*, in *Werke in zwänzig Bänden*, hrsg. von E. Moldenauer und K.M. Michel (Frankfurt a.M.: Suhrkamp, 1969), § 92 Randbemerkung, 179.

[50] G.W.F. Hegel, *Elements* n 9 above, § 97, 123.

[51] ibid § 82Z, 116.

[52] ibid § 99, 124.

[53] ibid § 86, 117.

[54] ibid § 97, 123.

universality, also punishment has the same vocation as it restores universal will. Since crime is not something positive, but essentially negative, then punishment is the negation of negation: thanks to punishment, right reaffirms its validity and is confirmed as necessary, but at the same time it appears as the result of mediation. If a positive value were to be attributed to a criminal action, it could be argued that its merit lies in bringing out the inadequacy and insufficiency of Abstract Right. It can then be said that *die Äußerung* of crime coincides with *die Verwirklichung* of right, since, thanks to punishment, it is restored and confirmed, after overcoming the initial immediacy. Punishment and before it wrong must not then be considered as an unforeseen inconvenience, but as the manifestation itself of right. Force shows itself to be necessary as long as it allows the overcoming of abstraction and immediacy: thanks to punishment, right develops a process of actualisation and it does not remain a mere 'must be', a postulate, a requirement – *Forderung*[55] – at the mercy of subjective will, but comes to terms with its maximum negation. Precisely as a lesion of the lesion, punishment compensates and indemnifies, determining a positive condition that is the result of a negative one. It performs a double task: the first function, which can be considered the strictly juridical objective, consists of the fact that, thanks to coercion, right overcomes its abstract character, gaining real application and effectiveness, while the second shows how the penalty is the means by which it is the person who overcomes arbitrariness and naturalness. Punishment represents the moment in which the same right is shown as an instrument aimed at the universality of freedom, since it works in order to favour the integration between the individual and society through the attempt to neutralise the conflict and to stabilise the recognition relationship.

At this level of the *Elements* punishment is presented as retribution, ie as a consideration for the wrong that was committed in order to reintegrate the violation. The purpose of punishment is *wiederherstellen*, that is to restore the established order after the lesion, but in this way also *wiedervergelten*, retribution, the lemma that in German contains the term *gelten,* ie to make valid and to give value as it operates so that to be valid, *gelten*, is not the *daseienden Wille* of the criminal, but the right in itself.[56] The repressive function of punishment therefore justifies itself as it is the restoration of the universality of the right – regardless of the intentions or expectations of the actors that are involved – and it is the universalisation of the subjective will. Punishment is something that at the same time is due to the individual as he is free, it restores the relationship of recognition between the contracting parties and gives the right effectiveness and existence, since the possibility of forcing cannot but be up to the right, because it is a matter of its very existence.

[55] ibid § 89, 119.
[56] ibid §§ 99 and 101, 124 and 127.

III. From Revenge to Punishment

When Hegel seems to have demonstrated the legitimacy of punishment, he introduces a topic that apparently questions the argument he has discussed so far. He states that 'in this sphere of the immediate, the cancellation (*Aufheben*) of crime is primarily *revenge*'.[57] In Abstract Right punishment seems to contradict itself, appearing as revenge rather than justice. In order to avoid contradicting ourselves, and therefore as Hegel always states, it is possible to fully speak of 'punitive justice' and not of 'an avenging justice',[58] it is necessary to highlight the limits of the presentation we carried out so far and of punishment when it is exclusively meant as retribution. We can say that these limits may be identified in two distinct elements, one characterised by a formal character, the other by a content nature.

Regarding the latter, the problem concerns the 'determined qualitative and quantitative magnitude' of the punishment and the relationship of proportionality and correspondence between the wrong and the punishment. Hegel is aware that in criminal science we are dealing with 'the realm of finite things', which excludes any 'absolute determination' and only allows 'an approximate fulfilment'.[59] The error of this conception of punishment lies in the fact that it conceives the payment between infraction and coercion as 'an equality in the specific character', when it should concern that of the 'character in itself' and it should be established starting from 'inner equality of things which, in their existence, are specifically quite different'.[60] The problem of the proportionality of the punishment led to a misunderstanding of the notion of equality, leading to a representation of retribution as 'robbery for robbery' and therefore leading to apply the logic of 'an eye for an eye and a tooth for a tooth'. In order to establish a just and fair punishment, one must proceed to the *Vergleichbarkeit* of the value and an 'inner identity' between the two terms must be considered, while not being confined to applying the law of retaliation.[61]

From the point of view of form, instead, punishment as it is presented in *Abstract Right* remains 'the action of a subjective will', which 'exists for the other party only as a particular will', so that which should be a legal constraint appears to be 'a new infringement': justice then presents itself as 'in altogether contingent' and then turns out to be revenge, generating 'an infinite progression'.[62]

[57] ibid § 102, 130.
[58] ibid § 103, 131.
[59] ibid § 101, 128.
[60] ibid § 101, 128. Hegel highlights how it is important not only to distinguish a crime against the person from the one towards property, but it is fundamental to respond to the wrong with fair punishment: to punish a robbery with death means confounding a wrong against a property relationship with the punishment given in the case of a lesion against the person, just as it cannot be permissible to punish a crime with a mere indemnity.
[61] ibid § 101, 128.
[62] ibid § 102, 130.

Although punishment allows retribution of the lesion, it does not go beyond the logic of the settling of accounts because punishment is put in place by another single will, which does exactly what the former has done. In the moment in which it is another single individual who presides over the attribution of the punishment, from the formal point of view the two wills do not realise anything but two reciprocal lesions. Each one is for the other a particular individuality that acts against the other, in such a way that neither of the two actions is placed on a different horizon from that of resentment, retaliation or payback: hence this produces the effect that the restoration of right is left to the contingency of the will that punishes. Set in this way, the issue of punishment does not come out of the vicious circle of a continuous reiteration, of the bad infinity, since it will not be possible to interrupt the chain by which every wrong deserves a subsequent sanction. In order for the sanction to stop being considered the same as and analogous to the wrong, it must be taken away from the merely individual intervention: the overcoming of revenge for the benefit of justice therefore occurs when punishment is no longer imposed by a subjective will, but is established by the will as universal. In this way Hegel emphasises the insufficiency of the private sphere, since the relationship between two persons not only does not exhaust the juridical sphere, but cannot even found it as it reiterates the same act in an action-reaction dynamic which reduces right to personal revenge.

In order to explain this passage, Hegel uses the example of the Greek tragedy and cites Aeschylus's *Eumenides*, claiming that 'among the ancients, revenge and punishment are not yet distinct: *Dike* is revenge and punishment, the Eumenides are goddesses of revenge and punishment'.[63] This reference, which was already present in the essay on natural law written in Jena, leads to emphasise the political value of the work of the Greek dramatist in order to explain through the interpretation of the tragedy the conclusive passage of Abstract Right. Indeed, the *Eumenides* display an incurable conflict because each one of its protagonists is at the same time innocent and guilty, a murderer but without guilt, who killed just in order to correct a previous crime. This raises the issue of the 'right right', that is the right that must win, whether the one of Apollo, which absolves Orestes, or the one of the Erinyes, who want to do justice, making Orestes suffer the shame of his crime. If every case of revenge seems to be legitimate, at the same time, when it proposes private justice, it cannot be considered the adequate form of resolution of a conflict, because it redresses evil with another evil and produces a spiral of violence. The tragedy represents precisely the situation depicted in paragraph 102, in which Hegel states the impossibility of restoring right with a private sanction, since in this case revenge 'is inherited indefinitely from generation to generation'.[64] In the same way the turning point told by the tragedy will coincide with the solution claimed by

[63] G.W.F. Hegel, *Lectures* n 36 above, § 48.
[64] G.W.F. Hegel, *Elements* n 9 above § 102, 130.

Hegel. The *Eumenides* stage both the conflict and the composition of the conflict, which presents two opposing instances, the one of Apollo and the one of the Erinyes: the court established by Athena, the Areopagus, issues the sentence about the innocence of the accused Orestes. In this way it overcomes the two principles personified by the modern and ancient divinities for the benefit of consolidating the ethical dimension, exemplified by the new function that the Erinyes will perform as Eumenides. In the *Oresteia,* Aeschylus represents the transition from summary and primitive punishment to the dominant one in the *polis*, from taking justice into one's own hands up to being members of a community that intervenes – in the form of the court – in order to settle disputes, from revenge, as the only way of retribution, to justice, as a power standing above private interests. Athena's action leads to the establishment of a political dimension as a place of justice, in which right is not the prerogative of the individual, but it is the instrument that helps to restore the violated freedom. What is decisive is not only the outcome of the judgment, that is the absolution of Orestes, but the fact that the justice of the *polis* interrupted the chain of revenge and defeated the law of retaliation.

A reading of the *Eumenides* therefore makes it possible to enlighten the Hegelian passage from revenge to justice, precisely because Hegel understands the urgency of removing the universal juridical dimension from particular feelings, in which subjectivity is wrapped, and becomes aware of the risk, that the sanction undergoes, when it is the discretionary prerogative of the victim of the wrong. Only in so far as the punishment becomes justice does it lose its revenging character, it overcomes the accidentality and the contingency in which it arises and acquires a universal value for the entire collective dimension, because, when a right is violated, restoring justice is the general and common interest. As in the *Eumenides*, so Hegel concludes the discussion of wrong by postulating the occurrence of an institute that, as a public authority, presides over the application of justice. This institution, however, can only arise within Ethical Life, where individual interest is reconciled with the general one: institutions and the laws express a stable content, that is independent of the subjective opinion, objectively realise freedom and manifest the rationality that distinguishes the course of universal history. In this perspective the Hegelian strategy consists of exposing, through Abstract Right, the fundamental categories of private right in order to present their strength and limits: Abstract Right is a presupposition of Ethical Life, but at the same time it refers to the latter in order to be able to find an application, as a non-self-sufficient horizon due to the lack of conciliation between the universal and the particular. The penal right constitutes the *trait d'union* between Abstract Right and Administration of Justice: precisely because indeed, a right in itself, as a punitive right, requires the overcoming of revenge as a form of retribution, it leads to admitting the need for a public right.

IV. Punishment as Institutionalisation of Force

In Civil Society, Hegel presents the 'objective actuality of right',[65] since right itself assumes the form of positive right, which is universally valid and applies to particular cases. Those rational principles set forth in Abstract Right are included in the field of *Dasein*, as 'right in itself is posited in its objective existence', so that right shows its validity as *Allgemeines* as is recognised, known and willed even by consciousness.[66] In Administration of Justice, *abstraktes Recht* on the one hand assumes the characteristics of historically existing right, on the other hand it is no longer subordinated to subjective will nor is it deprived of coercive power but acquires a necessary and autonomous value from arbitrariness of the individual will. The categories that are exposed in Abstract Right, such as property, contract and punishment, acquire reality because they are defined as laws in a system, so that the same right becomes effective in actual Civil Society. Likewise, while in Abstract Right person coincided with legal capacity, the person who is the protagonist of Civil Society is now a 'concrete person',[67] that is an individual moved by specific interests and needs: since the context around him is historically, socially and economically developed, the concrete person acts as a member of a family, a worker placed in a productive system or a citizen of a state. In some paragraphs of Administration of Justice Hegel explicitly returns to the considerations set out about coercion. If Abstract Right indeed ends with the reversal of punishment into revenge, bowed to a particular interest and to the arbitrariness of subjective will, the administration offers a different picture of it:

> since property and personality have legal recognition and validity in civil society, *crime* is no longer an injury (*Verletzung*) merely to a *subjective infinite*, but to the *universal* cause (*Sache*) whose existence (*Existenz*) is inherently (*in sich*) stable and strong.[68]

In this perspective it is necessary to revise the penal conception which identifies itself with the principle of retribution, since the crime – and consequently its punishment – now acquires new meaning. In Administration of Justice, a crime violates a positive law and therefore of course the victims, but, in an overview, it violates the same system that codifies the behaviours and the exchanges between citizens, ie the true universal will, that represents the cornerstone of Civil Society as such. While in fact in Abstract Right the criminal action was an act directed against another subjective will, against any other person, so much so that Hegel claims that 'there is still no mention of

[65] G.W.F. Hegel, *Elements* n 9 above § 201, 240.
[66] ibid §§ 209-212, 240.
[67] ibid § 182, 220.
[68] ibid § 218, 250.

punishment in the form of punishment', now it appears as an act against the entire system that sanctions subjective rights. If in fact right in the 'the form of revenge' is merely 'right in itself', or 'not just (gerecht) in its existence (existenz]', in Civil Society it is universal to be harmed, thus generating a redefinition of the same concept of punishment, which 'ceases to be merely subjective and contingent retribution of revenge' to become 'the genuine reconciliation of right with itself'. It objectively corresponds to the 'conciliation' of the law 'which restores and thereby actualises itself as valid', and subjectively 'it applies to the criminal in that *his law* which is known by him and is valid for him and for his protection'.[69] Indeed, already in Abstract Right, Hegel identified the anti-juridical nature of the crime in the lesion of the universality of right in itself, but punishment turned into revenge, since it was itself an action that was committed by a subjective will, since the conditions were not identified nor was a person responsible for imposing it. On the contrary, in Administration of Justice, the fact that a crime is a lesion of the universal means that the universal has the task of punishing the infringement of the right through the institution of the trial and thanks to the court.

It is in Ethical Life that punishment fully realises itself as *Versöhnung* of the will of the offender and of universal will: the penal sanction is then founded and legitimised in Civil Society, since they are the judges, officials of the state and not private persons, to determine the modality and the extent of the punishment in the trial, according to what is stated by positive law in light of the gravity of the crime. In Abstract Right, wrong involved the loss of right in itself, whereas in Civil Society the latter has the tools to bring the particular back to the universal. In a penal trial the universal constitutes at the same time both the injured party, since the crime is itself a lesion of the universality of the right, in this case of the law, and what intervenes to re-establish the right as a *super partes* instance, which has the task of remedying the opposition. In this way a crime generates a break within the universal, which splits and duplicates itself as a part and as a whole, so that the conciliation takes place on a double track, the will of the offender and the law, on the one hand, and the right with itself on the other hand. The law also appears in multiple ways, since it is both what has been transgressed, and what allows the recomposition of the violation, establishing the ways in which to serve the sentence.[70] The offended person then participates in the trial, but in a mediated way, since his protection is delegated to a State's official, so that, says Hegel,

> the right is the Eumenides, the well-intentioned, which is equally protection of the criminal and it only realises what is in the necessity of the

[69] ibid § 220, 252.

[70] About this passage and, more in general, about the structure and articulation of the process, that are only mentioned, see a more detailed discussion by S. Fuselli, *Processo, pena e mediazione nella filosofia del diritto di Hegel* (Padova: CEDAM, 2001).

thing, in order to ensure that with respect to the offender there is no 'valid individual arbitrariness.[71]

This conciliation is also possible because in Ethical Life right rests on an inner adherence by the individual will and presupposes a trust in the universal that is rooted in the very structure of the will. If in Civil Society right is universally recognised, it is because the opposition between the abstract universality of right and the particular will of the individual is overcome. Obviously, this does not mean that right cannot be subjected to transgressions or that the individual spontaneously adheres to every imposed obligation, but simply that, beyond individual infractions, the individual has developed a subjective disposition which is at the basis of the relationship of obligation. In this way the wrong must be brought back to the arbitrary will, but this does not undermine the fact that the individual became capable of abstracting from his own particularity, recognising himself as equal to the others and being aware of being able to find his own wellbeing within the State. As long as subjectivity has undergone a process of formation, in Civil Society it can recognise the same right as universal: the possibility for the right of being objectively valid and independent from the particular individual will must be understood as the outcome of the awareness of his universality gained by the individual. Also, in Civil Society, the criminal right therefore retains the function of *Bildung*, since the different phases of the juridical process create the conditions for the members of Civil Society to be 'spiritually present, with their own knowledge' and to consider the law as 'the most proper', 'the substantial and rational'. The publication of the laws and that of the sentence, as well as the participation to the trial of courts of jurors made up by every educated person shows how the right in the *Rechtspflege* requires knowledge and the will of the individual and requires that this knowledge contributes to the application of right.[72] The general principle of *Öffentlichkeit* therefore ensures a dual function in Administration of Justice. On the one hand, it guarantees the correctness of the trial as a protection from any eventual abuse by the judge who is called to express his opinion on the legal aspect of the dispute regardless of the subjective opinions; on the other hand, it forms public opinion and carries out a role of education to the universal towards the same individuals, strengthening the trust they have in the law.[73]

[71] G.W.F. Hegel, *Vorlesungen* n 47 above, § 220, 670-671.

[72] G.W.F. Hegel, *Elements* n 9 above, §§ 222-227, 253-256.

[73] In particular in the Lectures on natural right of 1817 Hegel stresses the role played by jury courts 'to foster a trust and awareness' of right and to avoid that right appears to individuals as a 'an alien power'. That is the reason why Hegel affirms that 'the judicial system is nearly as important as the law itself and among civilised peoples should be as fully developed as possible' see G.W.F. Hegel, *Lectures* n 36 above, §§ 115-116, 200-207. See also G.W.F. Hegel, *Elements* n 9 above § 319, 355-358.

If in positive right accidental and contextual elements come into play, that are linked to historical conditions and to the character of Civil Society,[74] this also has an effect on the definition of penal law, since the punishment is not determined 'in terms of its concept', but 'in terms of its outward existence'.[75] What is defined as a crime, its gravity and the punishment that it entails are established in a contingent manner with respect to a series of variable historical and social factors and in relation to the way in which that same act is considered by the citizens. Because of this, Hegel specifies, 'this gives rise to the viewpoint that an action may be a danger to society',[76] since the extent of the punishment depends on specific assessments that let the same crime be punished differently depending on the historical epochs and in particular on the level of culture and 'education of the people'.[77] Therefore, if in a particular age an action can be considered as a serious crime, while the same act in a successive period might theoretically not even be subject to sanction, this depends on the fact that 'a penal code is therefore primarily a product of its time and the current condition of civil society'.[78] From a general point of view, it is possible to see how the strengthening of civil society in modern times has meant that the punishment of crimes has become progressively milder, as is shown by the opposition to torture and abolition of the death penalty.[79] Hegel claims that 'with the progress of education, however, attitudes toward crime become more lenient, and punishments today are not nearly so harsh as they were a hundred years ago. It is not the crimes or punishments themselves which change, but the relation between the two'.[80] The more a society is stable and strong, the milder the penalties will be, the more a crime puts the political system in crisis, the more it will be severely punished.

The justice of punishment is not determined on the basis of universal and rational principles, but on the contrary it is closely linked to particular circumstances. Therefore, a punishment that is right in itself does not exist, as a strictly retributive principle would lead to admit: since punishment is defined in relation to 'the conditions of their time' and cannot be valid 'for every age',[81] it may not be unfair to punish criminals differently for the same crime. This means that aspects that are completely absent in Abstract Right are now taken into consideration, as in the first place the consequences that a given wrong can have with respect to the stability of society. Just as in Abstract Right crime had a universal value, the same happens, or even more, in Civil Society: in Civil

[74] ibid § 212, 243.
[75] ibid § 218, 250.
[76] ibid § 218, 250.
[77] G.W.F. Hegel, *Vorlesungen* n 47 above, § 101, 322.
[78] G.W.F. Hegel, *Elements* n 9 above § 218, 250.
[79] ibid.
[80] ibid § 96Z, 123.
[81] ibid § 218Z, 251.

Society a crime 'is an injury to all others' and not only the victim, ie to the individual, so that crime is never a private matter between two individuals but calls into question the very foundations of civil coexistence, ie the 'basis and the ground'[82] of Civil Society. This implies that the same action produces external effects that go beyond the criminal's own intentions and that his action has implications that go beyond the action itself. The case of theft or robbery is paradigmatic: it certainly violates the principle of private property, but it also has an additional effect on Civil Society, because it affects the feeling of security of the members of the society. This fact is taken into consideration by the judge in the sentence and with respect to the determination of the punishment.

However, by identifying the factors that contribute to establishing the entity of the punishment, Hegel also makes a shift with respect to the foundation of the justification of the punishment itself. To the extent that criminal law depends on the degree of stability and cohesion of the society or on the actual threat that the crime produces, it follows that 'the attitudes and the consciousness of civil society' come into play. From this point of view, then, punishment must be considered not simply as a 'lesion of the lesion' and therefore as retribution, as if the same offender was subjected to the same law that he set, but in relation to the social consequences that the wrong involves, both with respect to the danger to civil society and to the representation that individuals make of a particular crime.[83] In his lectures, Hegel is even more explicit in that he states on the one hand that a criminal act 'embodies a bad example', thus assuming a dissuasive value up to the extent that punishment appears to be a deterrent to the repetition of the crime by other individuals.[84] Indeed, Hegel states that 'under the conditions of civil society the aim and purpose of improvement can enter the question of punishment. It is important that it does so, and is even necessary'.[85] Precisely for this reason it can be admitted that the punishments that are established for repeat offenders are harsher than those established for the ones who commit the same crime for the first time. This shows that the purpose of punishment consists of educating the criminal and therefore in his reintegration into society in order to ensure that 'a person can be reintegrated by society' and to avoid that the realisation of crimes 'becomes a habit'.[86] Since it is established in relation to the social consequences that the crime implies and to the power it has to condition the behaviour of the entire community, punishment now holds a social function aimed at influencing the future action of other citizens, as well as re-educating and reintegrating the offender into

[82] G.W.F. Hegel, *Lectures* n 36 above, § 114, 198.

[83] G.W.F. Hegel, *Elements* n 9 above, § 218, 250.

[84] G.W.F. Hegel, *Vorlesungen über Rechtsphilosophie nach der Vorlesungsnachschrift von K.G. Griesheims 1824/25*, in K.H. Ilting ed, *Vorlesungen über Rechtsphilosophie 1818/1831* (Stuttgart-Bad Cannstatt: Fromman-Holzboog 1973), IV, § 218, 549.

[85] Ibid § 218, 553.

[86] G.W.F. Hegel, *Lectures* n 36 above, § 113, 197 and § 114, 200.

society, so much that here the Hegelian conception of punishment assumes the traits of a utilitarian theory: since punishment is no longer distinguished as an element of private law, but by its public function and value, punishment certainly retains its 'quality of justice', but it can also have other ends like the ones of reformation and deterrence.[87]

V. Conclusion

Hegel's penal theory shows to be two-sided: on the one hand, as it was presented in Abstract Right, it can be considered among the retributive conceptions; on the other hand, as was stated in Administration of Justice, it seems to approach utilitarian positions. As in the case of the different moments of Objective Spirit, while in the case of Abstract Right, Hegel sets the problem of the rational foundation of punishment, and in Civil Society he questions himself about its purpose and its applicability. However, these are not two opposing positions or two antithetical theories, but two different points of view: on the one hand that of Abstract Right, rational and formal, on the other hand, that of positive, historical and contingent right as it is characterised in Administration of Justice. In the first case, Hegel identifies the conditions by which it is possible to justify coercion, in the second case the same principle, by finding application on the terrain of the society, can only change by taking into consideration aspects that were previously irrelevant and concern the improvement of the offender or prevention in general. This passage is the result of a depersonalisation of the determination and application of the punishment, which must make abstraction precisely from the subjectivity of the victim, the guilty and the one who imposes the punishment, because the latter is not imposed because of the accidentality of a subjective will: it is not only a right of the criminal, but a right of the right.

Despite such a double justification of punishment, both in Abstract Right and in Civil Society, punishment plays an analogous function as far as it presides over the universalisation of consciousness and the effectiveness of right. From the point of view of right, indeed, it is directed not only to re-establish legality, but also to overcome its abstract character, showing the necessity for it to gain the concrete level of action, whereas from the point of view of the individual, coercion conceived as violence accomplishes a task of *Bildung* against the natural and immediate impulse. Therefore, punishment guarantees the role of

[87] G.W.F. Hegel, *Vorlesungen* n 84 above, § 218, 554. This aspect of the punishment is more developed in the lectures of philosophy of right than in the published text, in particular in the Manuscript by Wannenmann (§ 114) and in that by Griesheim (§ 218, 548), where Hegel states that the criminal 'injures society as such' and the 'crime receives then the determination of the danger', so that a single case presents 'the character of universality'.

mediating between the universal and the particular will on a double level: it allows one to heal, on one hand, the separation between right and reality, ensuring the effectiveness of justice and, on the other hand, the opposition between right and subjective will, contributing to the development of that individual disposition that Hegel calls rectitude in Ethical Life. It is also thanks to punishment that right is not only an ought: punishment thus represents the unification of the abstract rationality of right and the particularity of the will, having the function – perhaps the ambition – of guaranteeing the realisation of right as injustice is the violence committed both against someone and against the right tout court. Thus, far from being what opposes itself to freedom, legal coercion is necessary precisely as a right of freedom, because on the one hand – and this is what is evident mainly in Abstract Right – it is presented by Hegel as a result of the dialectic of the will of the criminal: not only a duty, but at the same time a right of the individual thanks to which the universality of the will is acknowledged, as far as punishment is a limitation of freedom only if freedom is identified with arbitrariness. On the other hand, punishment is also something by which right in itself as the realm of realised freedom ensures its own existence, it is the restoration of the universal like the idea of freedom, thanks to which the latter is embodied in social and political structures.

Through a subsequent reading of Abstract Right and Administration of Justice, it is possible to state that punishment highlights the necessity for Hegel that private right be open to public law, as an axiologically, logically and chronologically prior horizon, as a presupposition that orders, organises and regulates intersubjective relations. Consequently, right can only be accompanied by a strictly political moment, as a place of decision aimed at establishing an order, in such a way that right represents an institution that is functional to guaranteeing stability through its own authorities that are capable of resolving disputes and conflicts. From this point of view, Hegelian penal theory seems to be possible to interpret as the justification of the thesis that the State is the only subject that has the monopoly of legitimate force, as opposed to those positions that support the possibility of the private use of force:[88] the constraint is *Zwang* only because it is framed within a political-institutional dimension, without which it would be nothing but mere *Gewalt*. Precisely for this reason, Hegel states the necessity of an institutionalisation of punishment that through the authorities of the State ensures respect of the law and that, taking charge of punishment, demonstrates that crime is not only a violation of another individual, but of that universal that is represented by the juridical order in which he places himself.

The ambitious project of combining the legal constraint with the concept of justice shows at the same time how Hegel is far from a naively natural law approach, as it is the frame of the positive right that outlines the conditions not only for the application of right, but also for the identification of the parameters

[88] J.-F. Kervégan, n 16 above.

of justice. Hence justice does not depend on a natural order, nor does it correspond to universal, transcendent and absolute principles, but we could say that it is defined in the context of Ethical Life: it depends on what is established by the law and is ensured by respect of the procedures of the State of right – therefore from respect of legality – but precisely for this reason, being always contingent and historically determined, it is exposed to the risk of being unfair. Thus, we can conclude with the umpteenth paradox, which is even more relevant for those who, like Hegel, have been portrayed for decades as the thinkers of logic and absolute reason and who in this case appear to be well aware of the incurable contradictions that are present in Objective Spirit. The use of violence does not appear as an episodic or isolated event, but rather as a trait which is proper to the juridical, witnessing the fact that in the same Ethical Life, the quintessential dimension of conciliation and pacification, the conflict and the division are not deleted at all. It is precisely through punishment that order is maintained in the State and any form of subversion is expelled: right is the instrument through which the existing system is renewed and continuously restored, as well as the device by which any opposition is contained and any centrifugal force that is dangerous for the state itself is cancelled.

Punishment is at the same time a clear example of the tension that is typical of the Hegelian approach between two opposing instances, the one of the universality of the principle, which imposes internal equality between crime and punishment, and the one of adherence and rightness with respect to the concrete case. The central aspect is indeed the dialectic that is established between rationality and generality of right, on the one hand, and the particularity and specificity of the individual case, on the other hand. In this sense, the logical normative principle underlying the crime-punishment relationship, by which the crime must be punished on the basis of a retributivist logic, can only be placed in historically determined contexts, as is highlighted by the fact that the same types of crimes will be punished, in different societies and in different historical epochs, in different ways depending on the contingent elements that inevitably exceed the abstract dimension. Right then inevitably lies in the contradiction by which, when it comes to establishing, for example, the right measure of punishment, this cannot be entirely rationally determined. The act of determining the quantitative of the punishment and thus its entity finally belongs to the judge's decision, which represents, in the field of the imponderable, the only way through which it is possible to solve the issue. Hegel then specifies that it is not possible to establish by reason 'whether the just penalty for an offence is corporal punishment of forty lashes or thirty-nine', although 'even one lash too many' is 'an injustice', *eine Ungerektigkeit*.[89] If, as a result of a decision, the punishment is then always questionable, accidental, contingent and therefore involves an element of arbitrariness that is impossible to remove,

[89] G.W.F. Hegel, *Elements* n 9 above § 214, 245.

once again we can say that it constitutes an access key to question the statute of the right, which is constitutively open to injustice: the latter is indeed a possibility that is present within the same juridical field, which is dependent on a residue of irrationality that can never be completely removed, since it derives from the implicit contradiction between universality of the norm and adherence to the particular case.

Nietzsche, La Mettrie, and the Question of the Legitimacy of Punishment: A Hidden Source?

Marco Piazza*

Abstract

Friedrich Nietzsche (1844-1900), starting from the years of *Human, All Too Human* (*Menschliches Allzumenschliches*: 1876-1878) elaborates a conception of punishment based on an organic reflection on the origin of morality, the function of custom, the critique of remorse and the origin of justice, a reflection that then finds a definitive reworking at the time of On the Genealogy of Morality (*Zur Genealogie der Moral*: 1887). About one hundred and thirty years earlier, in his Discourse on Happiness (*Discours sur le bonheur*: 1748-1751), Julien Offray de La Mettrie (1709-1751) had elaborated a conception of punishment with several analogous lines. Starting from this theoretical coincidence, in this article we ask ourselves: did Nietzsche know the theories of La Mettrie? Had he read his works? Do the two philosophers really support the same theories? To try to give an answer to these questions, we will first present the doctrine of La Mettrie and then that of Nietzsche, before presenting a final balance sheet of the survey.

I. La Mettrie and the Legitimacy of Punishment

The ethics of Julien Offray de La Mettrie (1709-1751) is consistent with the materialistic framework of his thought.[1] In its most mature form it is contained in his Discourse on Happiness, published in three different versions between 1748 and 1751.[2] One of the fundamental mainstays on which it is based is the idea

* Associate Professor of History of Philosophy, Roma Tre University.

[1] On the La Mettrie's thought see: R. Boissier, *La Mettrie. Médécin, pamphlétaire et philosophe* (Paris: Les Belles Lettres, 1931); P. Lemée, Julien *Offray de La Mettrie, St-Malo, 1709, Berlin, 1751, médecin, philosophe, polémiste, sa vie, son oeuvre* (Mortain: Éditions du "Mortainais", 1954); K. Wellman, *La Mettrie. Medecine, Philosophy and Enlightenment* (Durham/London: Duke University Press, 1992); C. Morilhat, *La Mettrie. Un matérialisme radical* (Paris: Presses Universitaires de France, 1997); A. Punzi, *I diritti dell'uomo-macchina: studio su La Mettrie* (Torino: Giappichelli, 1999); U.P. Jauch, *Jenseits der Maschine: Philosophie, Ironie und Ästhetik bei Julien Offray de la Mettrie, 1709-1751* (München/Wien: C. Hanser Verl., 1998); M.Á. Cordero del Campo, *Materialismo y voluptuosidad en la filosofía de Julien O. de la Mettrie* (León: Universidad de León, 2003); A. Paschoud and F. Pépin eds, *La Mettrie, philosophie, science et art d'écrire* (Paris: Éditions matériologiques, 2017).

[2] J.O. de La Mettrie, *Discours sur le bonheur*, critical edition by J. Falvey, *Studies on Voltaire and the Eighteenth Century*, ed. by T. Besterman, vol. CXXXIV (Banbury: The Voltaire Foundation/Thorpe Mandeville House, 1975); partial Engl. transl. *Anti-Seneca or the Sovereign Good*, in Id, *Machine Man and Other Writings*, transl. and ed. by A. Thomson (Cambridge: Cambridge University Press, 1996), 117-144 (texts respectively cited hereafter with the abbreviations

that free will is a chimera.[3] An idea that is at one with the principle of moral irresponsibility, anchored in the absolute determinism that governs our actions, which we erroneously consider free, unwilling to accept to consider ourselves slaves of necessity:

> The will is necessarily determined to desire and seek what is to the immediate advantage of the soul and the body. (…) And yet I think that I have chosen and congratulate myself on my liberty. All our freest actions are like that one. An absolutely necessary determination carries us away, and we will not admit that we are slaves! How insane we are, and all the more unhappily insane for permanently reproaching ourselves with not having done what it was not at all in our power to do (!) (DB 160-161; AS 141).

The determinism professed by La Mettrie finds its roots in a sort of medical philosophy or philosophical medicine, according to which our conduct would find its causes in the individual temperament,[4] which is therefore configured as a permanent and substantially non-modifiable element, once the process of its development is completed:

> When I do good or evil, when I am virtuous in the morning and wicked in the evening, it is the fault of my blood, which makes me cheerful, serious, lively, playful, amusing, mocking or mad, and which makes me will and determines me in everything (DB 160; AS 141).

The medical philosophy promoted by La Mettrie is therefore in open contrast with the traditional theological metaphysics, and has no fear of assuming unpopular and such radical positions as to recall the anger of conformist people:

> What a point have we reached, cry the theologians, if there are no inherent vices or virtues, no moral good or evil, no justice or injustice? (…) We shall leave them to make speeches and start calmly along this new path, where we are led by the best philosophy, that of physicians (DB 150; AS 135).

The radical materialism that guides medical medicine allows us to establish a second key principle of the ethics of La Mettrie, that of the arbitrariness of virtue, which, so to speak, assumes an unequivocal validity when the field is cleared of metaphysical prejudices: 'Stripping away little by little of his prejudices, he (the

DB and AS, followed by the Arabic number of the page or pages cited). Henceforth, the text of the *Discours sur le bonheur* will be quoted in the mentioned English translation, where available (in fact it stops at the end of Section III, on page 117 of the manuscript or on page 165 of the critical edition edited by John Falvey), while always maintaining the reference to the pagination of the Falvey edition. The translation of the passages not included in Ann Thomson's English translation is ours.

[3] C. Morilhat, n 1 above, 93.

[4] See K. Wellman, n 1 above, 188.

philosopher) will esteem, the virtue for that which it is, arbitrary' (DB 215).

The criticism of metaphysics then leads the physician-philosopher to establish another fundamental principle of his materialistic ethics, namely the indifference of good and evil with respect to happiness:

> As the pleasure of the soul is the true source of happiness, it is therefore very obvious that in relation to felicity, good and evil are totally indifferent in themselves, and that he who has greater satisfaction in doing evil will be happier than whosoever has less satisfaction in doing good. Which explains why so many rogues are happy in this world, and which shows us the existence of particular individual happiness without virtue and even in crime (DB 161-162; AS 141-142).

In other words, according to La Mettrie, as we have already noted, there is no good and no evil in itself, but this does not mean rejecting the existence or usefulness of a relative good or evil. However, these are moral norms whose sole foundation is social, as they are established solely and exclusively 'to make life in society possible'.[5] It follows therefore not only that these norms produce socially useful behaviors, but also a certain kind of happiness, which is therefore induced and in some way stands out and in some ways even contrasts with natural happiness, indifferent to the social good. Despite being advantageous for society as a whole, La Mettrie, consistent with this materialistic framework, does not consider this a 'purer' happiness:

> One source of happiness, which I do not believe to be any purer for being nobler and finer to the minds of almost everybody, is that which derives from the order of society. The more man's natural determination has seemed wicked, and as it were monstrous, in relation to society, the more it has been thought necessary to counteract it in different ways. Hence the ideas of generosity greatness and humanity have been linked to actions which are important for men's intercourse. Esteem and consideration have been accorded to the man who would never harm anyone however much good it might bring him; respect, honours and glory have been given to the man who would serve his country, friendship, love or even humanity to his own cost; and by means of these noble incitements, how many animals with human faces have become immortal heroes! (DB 162; AS 142).

The man of La Mettrie is essentially a natural being before a social being, and therefore the history of morality must be understood as a history of contrasting the natural unsociability of individuals, who are molded thanks to education, which corrects their natural egoism: in other words the social virtues, as far as 'socially

[5] C. Morilhat, n 1 above, 96. See also: K. Wellman, n 1 above, 219.

useful', 'are unnatural'.[6] In this way, however, the pedagogical action applies a mask over the nature of man that can always split apart, making the underlying physical being re-emerge. And just as the feelings of good and evil turn out to be social constructions lacking a natural foundation, even remorse turns out to be another construct, or an acquired habit:

> Remorse is (…) only an unpleasant remembrance, a former habit of thought, which returns in force (…) an old prejudice (DB 150-151; AS 135).[7]

La Mettrie adheres to the Cartesian philosophy of habit, based on the physiological doctrine of animal spirits, which we find operating in numerous authors between the late sixteenth and eighteenth centuries, substantially from Francis Bacon to David Hume.[8] This doctrine seeks to provide a scientifically founded explanation of a phenomenon of which we are aware by analyzing our life experience, that is, the possibility of replacing an acquired habit with another habit, even making it stable. A phenomenon that can be guided by reason, that is it can be oriented through a voluntary effort aimed at replacing a prejudice with a rational counter-habit.[9]

> Luckily this cruel enemy is not always the victor. Any longer-standing or stronger habit must necessarily defeat it. The most beaten track fades away, as a path is closed or a precipice filled. Another kind of education (*habitude*) brings another route for the spirits, other dominant traces and other feelings, which can enter our soul only on the ruins of the earlier ones, which are abolished by a new mechanism (DB 151; AS 135-136).

Therefore remorse, also because it is 'not is an innate sentiment', is not invincible, however rooted and 'engraved on the brain at a very early age':[10] by relying on reason it is possible to escape the power of habit, which La Mettrie identifies with an unhealthy education, which torments man by preventing him from following his own nature, his natural propensity to happiness, thus imposing an excessive weight on him which unjustly distances him from a disposition to pleasure in itself honest and innocent:

> Was man – whom nature has tried to attach to life by so many allurements, destroyed by depraved artifice – in particular the honest man,

[6] ibid 227.

[7] La Mettrie had already talked about the remorse in *L'Homme machine* (1747), showing it to be not at all connected to a presumed natural right of way: see K. Wellman, n 1 above, 197-198.

[8] See M. Piazza, *Creature dell'abitudine. Abito, costume, seconda natura da Aristotele alle scienze cognitive* (Bologna: il Mulino, 2018), 100-101.

[9] See ibid 105-110.

[10] K. Wellman, n 1 above, 220. Wellman shows the Lockean root of epistemology professed by La Mettrie about the difficulty of replacing primitive impressions with later ones when the former are simple and strong (ibid 220-221).

> created in order to be delivered up to tormenters? No, let him use the power of his reason to provide him with what is provided for so many rogues by the force of habit (*habitude*). For one villain who stops being unhappy and returns to peace and tranquillity, which he did both deserve in his relations with other men, how many wise and virtuous individuals, undeservedly tormented amidst a charming and innocently delicious life, would finally throw off the yoke of an oppressive education, enjoy clear cloudless days and replace the cruel worry which devours them with sweet pleasures? (DB 153; AS 137).

La Mettrie clearly contrasts a moralistic socialization with an eudemonistic socialization, in which man is led back to his own nature through an appropriate education that frees him from the chains of prejudice and remorse, thus demonstrating that his physiological determinism does not coincide with a rigid metaphysical determinism.[11] And to give a sort of demonstration of the validity of his own perspective, he uses a utilitarian argument: what is the purpose of remorse if it is useless to avoid the action considered immoral?

> Who has ever abstained from doing what gave him pleasure or what could make his reputation or fortune, simply through fear of feeling remorse? It (…) is therefore useless before a crime. But while one is committing it and is carried away by one's passion, one thinks of nothing less than that feeling by which one is going to be racked. And when the crime has been committed and remorse rises up as if to avenge society, only those who do not need it can profit by it. The suffering of the others, whose wickedness is innate and organic, rarely (if ever) prevents them from reoffending. Thus, remorse is in itself, philosophically speaking, as useless after as during and before a crime (DB 154; AS 137).

If remorse is useless, as La Mettrie tries to prove, the same cannot be said of punishment, the necessity of which is exclusively political, or arbitrary, and conversely void philosophically, since the determinism to which our actions are subjected relieves us of responsibility, to which the legitimacy of the punishment would connect. Therefore, the punishment is in itself unjust, but useful and even necessary from a social and political point of view.[12] Obviously the awareness of this makes its use more conscious, but it is only inspired by the principle of social utility, to which it looks – with a mixture of lucid cynicism and 'tolerance'[13] – the 'philosophical prince', split between the recognition of the injustice of punishment and the conviction of its inevitable necessity for the purposes of social governance:

[11] See K. Wellman, n 1 above, 223.
[12] See ibid 222.
[13] ibid.

> If someone who is guilty in relation to society is not free in his actions, it no doubt follows clearly that he was not free not be guilty and that he is guilty as if he were not guilty; he is guilty in one sense – in the sense of arbitrary, wisely established relations – but is not at all in another, not intrinsically, in the absolute sense or philosophically speaking. To put it bluntly, he is clearly not guilty at all and only deserves compassion. Even when a philosophical prince punishes him, he groans at being forced to come to this sad extremity; he knows that legal punishment is as absolutely unjust as it is relatively necessary and that consequently the political reasons which are the basis of law of retaliation do not prove that the man we hang is hanged with justice or equity (DB 164; AS 143).

The result is a double truth: the philosophical one, which however has a purely individual and intrapsychic range of action, and the political one, which has no philosophical foundation, but governs human relations. The philosopher is he who manages the painful contradiction between these two heuristic levels, and if on the one hand he represents the subject most capable of emancipating himself from prejudice, on the other he cannot escape the social law, the custom, embodied by the *nomos*:

> You, who we usually call unhappy and who are such in the face of society, can therefore feel comfortable before yourself! You just have to stifle remorse with reflection (…) or with contrary habits, which are much more powerful. If you had been raised with other principles, or without the ideas that underlie yours, you would not have had to fight these enemies at all. (…) Thief, parricidal, incestuous, thief, wicked, infamous and legitimate object of the execration of honest people, you will still be happy! In fact, what unhappiness or pain can cause actions that, no matter how black and horrible they may be, would not leave (according to the hypothesis) any trace of crime in the criminal's soul? But if you want to live, be careful, politics is not as lenient as my philosophy. Justice is his daughter; the executioners and the gallows are at his command: fearful more than your conscience and the gods (DB 195).

The reasoning is therefore taken to extremes by La Mettrie: a parricide or an incestuous person who came for hypotheses made completely devoid of any trace of moral education and were not influenced by the moral judgments of those around them, would be happy, since their acts are actually effects of a necessary determinism that has nothing to do with their freedom of action. But it is clear that this is a philosophical fiction, since the hypothesis disregards the political dimension from which we cannot escape and which, so to speak, obliges us to deal with the historical contingency. A contingency made up of real politics and the execution of laws inspired by principles of social utility that

reintroduce morality expunged in theory. Therefore, if we can be free from the conditionings of morality in the space of our conscience, we cannot act in conformity with this amoralism when we act as members of a political-social collective governed by principles, legislators and executors of punishments.

II. Nietzsche and the Legitimacy of Punishment in the Epoch of *Human, All Too Human*

Starting from the composition of Human, All Too Human, between 1876 and 1878, Nietzsche elaborates a conception of punishment that is part of a project of deconstruction of morality and metaphysics fed among other things by the assiduous and repeated meditation of Neo-Kantian and Darwinian theses by Friedrich Albert Lange. Of this author he had begun to read the monumental History of Materialism (first published in 1866) very early, one of the readings, together with that of Schopenhauer, which influenced him to move from philology to philosophy.[14] We have traces of his reading this volume in the letters of Nietzsche from August 1866 and then already in the fragments of autumn 1867 – spring 1868, about the ancient materialism of Democritus, Epicurus and Lucretius.[15] Unfortunately we do not have the copy of the first edition of Lange's text that belonged to Nietzsche, because he would have given it away before 1875.[16] Nietzsche again mentions the book in his correspondence and fragments starting from 1884 and in 1887 he purchases a copy of a reprint of the fourth edition of the volume (first published in 1882), still held in the Nietzsche Library, with numerous traces of reading.[17] Compared to the first

[14] F.A. Lange, *Geschichte des Materialismus und Kritik seiner Bedeutung in der Gegenwart* (Iserlohn: J. Baedeker Verlag, 1866). The second edition, with major revisions, was published by the same publishing house in 2 Vols: Vol. I, 1873; Vol. II, 1875. It is probable that Nietzsche also had this edition in his hands: see G. J. Stack, *Lange and Nietzsche*, in E. Behler et al eds, *Monographien und Texte zur Nietzsche-Forschung*, Band 10 (Berlin/New York: Walter de Gruyter, 1983), 13 (fn 9), 23. In this article the works of Nietzsche, except when not included therein, will be cited by: F. Nietzsche, *Digitale Kritische Gesamtausgabe Werke und Briefe*, based on the critical text by G. Colli and M. Montinari, edited by P. D'Iorio (Paris: Nietzsche Source, 2009), available at https://tinyurl.com/57hx8wrv (last visited 30 June 2021). References to this edition are indicated by eKGWB followed by the standard abbreviations used for Nietzsche texts, for example GT-8; preceded by www.nietzschesource.org this abbreviation allow readers to directly consult the text at the address https://tinyurl.com/3b9nzyxu (last visited 30 June 2021). I warmly thank my friend Paolo D'Iorio for his invaluable bibliographical information on Nietzsche and Lange as well as on the English editions of Nietzsche's works, without which I could not have completed this article.

[15] F. Nietzsche, *Nachgelassene Aufzeichnungen. Frühjahr 1864 – Herbst 1868*, *KGW I/4*, *Nietzsche Werke. Kritische Gesamtausgabe*, ed. by J. Figl and I. W. Rath (Berlin /New York: Walter de Gruyter, 1999), P I 6, Fr. 57[26], 390.

[16] T. H. Brobjer, *Nietzsche's Philosophical Context. An Intellectual Biography* (Urbana and Chicago: University of Illinois Press, 2018), 16.

[17] F.A. Lange, *Geschichte des Materialismus und Kritik seiner Bedeutung in der Gegenwart*. Wohlfeile Ausgabe. Zweites Tausend. Besorgt und mit biographischem Vorwort versehen von H.

edition, it is increased and revised, and we are at the time of the composition of *On the Genealogy of Morality*.

The influence of Lange's book on Nietzsche's philosophical formation for some years has been the subject of attention by scholars.[18] After all he could not go unnoticed; just think of this statement by Nietzsche himself: 'Kant, Schopenhauer and this book by Lange: I don't need anything else'.[19]

In summary, Lange considers materialism to be a philosophical school and as a kind of prophylaxis against idealism. His reading of materialism is aimed at its re-actualization (see the subtitle of the work, which literally sounds: Criticism of its Importance in the Present) which passes through neo-Kantianism. For him it is a question of developing Kantianism in harmony with the development of the physiology of the sensory organs. But materialism is also treated as an ontological metaphysics that is problematically founded on the assumption that reality is composed of matter and force. George J. Stack, who dedicated an entire volume to Lange's influence on Nietzsche, states that

> In his early notes of the mid-1860s one finds direct references to Lange's History and even in the notes of the late 1880s there are numerous entries that are identifiable Langean themes.[20]

Or that 'In his earliest notations of the 1870s and in his last notes before madness overtook him, traces of Lange's influence can be found'.[21]

According to Stack, the ingredients of the anthropomorphic idea of truth and knowledge presented by Nietzsche in On Truth and Lies in a Nonmoral Sense (1873) are directly influenced by his reading of the first edition of Lange's book, which in several passages, and above all where it deals with La Mettrie, exposes a theory of evolutionary signs in which words are the effect of excitations of specialized brain areas struck by sounds.[22] Furthermore, after noticing some similarities with the moral theory illustrated in Human, All Too Human (as well as analogies between Lange/La Mettrie and a passage from Dawn),[23] Stack writes:

Cohen (Iserlohn und Leipzig: J. Baedeker, 1887, XXX + 852). See the file relating to the copy held by Nietzsche's Personal Library (from now on: BN) in G. Campioni et al eds, *Nietzsches persönliche Bibliothek (BN)* (Berlin/New York: Walter de Gruyter, 2002), 346. There is a poor-quality translation of the second edition, reprinted several times: *The History of Materialism and Criticism of Its Importance*, transl. by E.C. Thomas, 2 vols. (London: Trübner & Company, 1877–1881).

[18] J. Salaquarda, 'Der Standpunkt des Ideals bei Lange und Nietzsche' *Studi Tedeschi*, XXII, 1, 133-160 (1979); G.J. Stack, 'Nietzsche and Lange' *The Modem Schoolman*, LVII, 2, 137-148 (1980); Id, n 14 above.

[19] eKGWB/BVN-1866,526 - Brief AN Hermann Mushacke: November 1866.

[20] G.J. Stack, n 14 above, VII.

[21] ibid 2.

[22] See ibid 138, 59.

[23] See ibid 140.

> Although there are other similarities between the views of Lamettrie and Nietzsche, a consideration of them would carry us too far afield.[24]

The ground on which the similarities to which Stack refers is certainly that of the critical genealogy of morality conducted by Nietzsche from the mid-seventies, within which he places his theory of punishment.

In a first phase Nietzsche elaborates a 'history of moral sentiments' in which morality is the object of scientific investigation: that is, you have to study its evolution, you have to grasp its natural causes. It is a matter of overcoming our resistance to investigate the motives of human actions and therefore we must proceed to a psychological dissection of morality.[25] On the basis of a common reflection with his friend and disciple Paul Rée,[26] in this phase Nietzsche will develop an articulated theory based on some key assumptions: a) the moral irresponsibility of the individual, according to which we erroneously believe that our actions are based on free will and instead are the effect of natural determinism and therefore of necessity (principle shared with Rée and well summarized by the aphorism 39);[27] b) the origin of 'good' (*Gut*) and 'evil' (*Böse*) brought back to a play of forces, to an exchange of power in which those who have the power to reciprocate, recognizing the power of those who are more powerful, are said to be 'good' (*Guten*), while 'bad' (*Schlechten*) is one who is unable to requite, and is part of the mass of impotent;[28] c) the origin of morality from the law, according to which morality is not spontaneous and therefore if we perform defined good actions we do not do them other than because without knowing it we follow the custom, so that altruism results an effect and not the cause of our 'good' action (in this regard see in particular aphorism 96, where Nietzsche crosses the theme, dear to him, of the first and second nature);[29] d)

[24] ibid 140-141. After Stack's work, the critique highlighted how about the theory of language developed by Nietzsche since *On Truth and Lies in a Nonmoral Sense*, it was also influenced by Nietzsche's reading of Gustav Gerber's book *Die sprache als Kunst* (in zwei Bänden, Bromberg: H. Beyfelder, 1871-1872). See: E. Behler, *Selbstkritik der Philosophie in der dekonstruktiven Nietzschelektüre*, in G. Abel and J. Salaquarda eds, *Krisis der Metaphysik. Wolgang Müller-Lauter zum 65. Geburstag* (Berlin/New York: Walter de Gruyter, 1989) 283-306.

[25] See the precise reconstruction in: M. C. Fornari, *La morale evolutiva del gregge. Nietzsche legge Spencer e Mill* (Pisa: ETS, 2006), 17-120.

[26] While Nietzsche was working on *Human, All Too Human*, Rée for his part was writing: P. Rée, *Der Ursprung der moralischen Empfindungen* (Chemnitz: Schmeitzner, 1877) and was already the author of: Id, *Psychologische Beobachtungen* (Berlin: C. *Duncker, 1875)*. The works of Rée can be consulted today in: *Supplementa nietzscheana*, hr. von Th Böning, W. Müller-Lauter, K. Pestalozzi, Band 7: *Paul Rée: Gesammelte Werke 1875-1885*, hr. von H. Treiber, mit einer Einleitung und einem Kommentar (Berlin/New York: de Gruyter 2004). On the theory of moral sentiments of Rée see: J. Salviano, 'O naturalismo moral e o pessimismo em *A Origem dos Sentimentos Morais* de Paul Rée' *Cadernos Nietzsche*, 39.2, 197-204 (2018).

[27] eKGWB/MA-39; Engl. transl. with an Afterword, by G. Handwerk, F. Nietzsche, *Human All Too Human, I* (Stanford, California: Stanford University Press, 1995), 47-49.

[28] eKGWB/MA-44 and eKGWB/MA-45; Engl. transl., 51-52.

[29] eKGWB/MA-96; Engl. transl., 73: 'We call someone "good" who, as if by nature, after long

the political character of morality, for which 'customs', through a process of habituation, make 'mild' what is initially 'hardest' and at the same time transform what is 'useful',[30] from the moment that the 'custom' itself, through the obedience with which it imposes itself, generates in us a kind of 'instinct' to take pleasure in behaving according to morality,[31] an instinct that can be silenced and replaced with another in the moment in which the critical history of moral sentiments allows us to realize that we can derive 'higher degrees' of well-being through 'other customs';[32] e) the instinct of conservation guides our actions: thus the actions considered 'evil' derive from 'the individual's striving for pleasure and avoidance of pain' and therefore not only 'are they not evil',[33] but they also do not derive from the will to cause pain in itself, to hurt the other in itself (since they are not arbitrary), even if there is pleasure derived from the 'feeling of superiority' on the other, which is proved when the other suffers because of us, but this pleasure is due to a sense of fulfillment in the exercise of one's power (therefore it is neither good nor bad, but rather useful or useless).[34]

On the level that closely concerns us here, that of the legitimacy of punishment and of penal responsibility, in this phase Nietzsche reaches the following conclusions:

a) Justice must be understood as a game of equal forces, or as an exchange or compensation: since the struggle between equals would lead to annihilation, one agrees by negotiating one's reciprocal claims;[35] in this perspective justice is a balance promised by the powerful (as an alternative to the marauder who does not do the same): the weak either unite to have equal weight or submit to the powerful, but since they fear annihilation, they choose the second option, generating the aforementioned balance.[36]

b) The origin of punishment lies in justice as revenge: when the balance of forces is broken, the disgrace that falls on those who undermine the balance in view of their advantage against others, as a social disadvantage, restores the troubled balance; the punishment is imposed as a castigation for those who oppose dominance, aspiring to something to which they are not right; the punishment, therefore, recalls the 'harshness of the state of nature'.[37] Thus judicial

inheritance, hence easily and readily, does what is customary'. On the doctrine of the first and second nature in Nietzsche we refer to: M. Piazza, 'Nietzsche e a dialética aporética entre primeira e segunda natureza' *Cadernos Nietzsche,* 39[3], 121-139 (2018).

[30] eKGWB/MA-97; Engl. transl., n 27 above, 74.

[31] eKGWB/MA-99; Engl. transl., 76.

[32] eKGWB/MA-97; Engl. transl., 74.

[33] eKGWB/MA-99; Engl. transl., 75.

[34] eKGWB/MA-103; Engl. transl., 79.

[35] eKGWB/MA-92; Engl. transl., 70-71.

[36] eKGWB/WS-22; Engl. transl. with an Afterword by G. Handwerk, F. Nietzsche, *Human, All Too Human II and Unpublished Fragments from the Period of Human All Too Human II (Spring 1878-Fall 1879)* (Stanford, California: Stanford University Press, 2013), 164-166.

[37] eKGWB/WS-22; Engl. transl., 166.

punishment restores the honor of both private and society. We turn to the court because we want private revenge with respect to the damage suffered and 'public revenge' of our honor publicly trampled and at the same time revenge of the honor of society itself.[38]

c) Punishment is detached from individual responsibility: the person who is punished in fact 'does not deserve the punishment' just as the one who is rewarded 'does not deserve this reward', as both act deterministically.[39] If, by hypothesis, we try to grant the existence of free will, we note that this would invalidate the concept of punishment: punishment is imposed because it is presumed that the offender at the moment in which he committed the crime knew what is good and what is bad and was free to choose between one and the other. But then he would arbitrarily choose evil, that is, without reason. Consequently, he should not be punished because he did not deny his reason voluntarily (in reality, for Nietzsche, as we know, the offender acts in a certain way for reasons he believes to be good, driven by necessary circumstances).[40]

d) Punishment does not punish a fault: justice does not punish guilt because if it really did, it should punish the circumstances that led an individual to commit the crime, that is, it should punish the educators, the parents, and even the judges themselves, who are members of the community to which the offender belongs.[41]

e) Remorse has no reason to be: 'Pangs of conscience are as stupid as the pangs of a dog biting a stone'.[42] If you have understood that you have done wrong, it is sufficient to act well. If an individual is punished for his actions, he bears the punishment, considering himself as 'humanity's benefactor', since we are punished for others not to behave like us (deterrent value of punishment).[43]

f) Punishment is arbitrary: habitual offenders, who should be punished with greater leniency – because they are more conditioned by their nature – are instead punished more harshly. While occasional offenders – who therefore have a less rooted inclination to the crime and could therefore resist more the push to the criminal action – are punished less harshly. This shows that the criterion of punishment is calibrated on society and not on the individual.[44]

g) The social utility of the punishment:

> If punishment and reward were to disappear, the strongest motives that impel us away from certain actions and toward certain actions would

[38] eKGWB/WS-33; Engl. transl., 175.
[39] eKGWB/MA-105; Engl. transl., n 27 above, 81-82.
[40] eKGWB/WS-23; Engl. transl., n 36 above, 166-167.
[41] eKGWB/WS-28; Engl. transl., 170.
[42] eKGWB/WS-38; Engl. transl., 177.
[43] eKGWB/WS-323; Engl. transl., 286.
[44] eKGWB/WS-28; Engl. transl., 170.

also disappear; the utility of human beings requires their perpetuation.[45]

Therefore, for the Nietzsche of the Human All Too Human's era punishment has only social and not moral or metaphysical value, in perfect harmony with what La Mettrie had affirmed in his time.

III. Nietzsche and the Legitimacy of Punishment in the Age of *On the Genealogy of Morality*

In a second phase, partly anticipated in The Gay Science and substantially corresponding to On the Genealogy of Morality, after reading Spencer, Nietzsche returns to the origin and purpose of punishment, completely distancing himself from Rée and embracing a position more attentive to the physiological implications of morality (gregarious structures, herd instinct etc.). Morality is now understood as the fulfillment of every function proper to the human species. The program of a reversal of values is clearly visible, bringing up again the will to live, animality, health, even the wickedness of knowledge against the submissiveness, passivity, anti-naturalness of traditional morality. While Rée and English philosophers hold the equivalence between good action and selfless action, exchanging the effect with the cause, Nietzsche intends to question the very value of morality, reading the power of the custom in filigree: 'Thus no one until now has examined the value of that most famous of all medicines called morality; and for that, one must begin by questioning it for once'.[46] If good is what is spiritually noble, it is a matter of reviewing the meaning of what is noble, it is a question of overthrowing the process that has made the non-egoistic something beautiful and pleasant.[47]

Nietzsche overturns Rée's explanation of justice as the effect of punishment understood as retribution or retaliation (this would presuppose that the offender could act differently).[48]

> "The actions that are necessary cannot be repaid" p. 49. Of course they can! He believes that they shouldn't be, that it would be unfair! That is, he is also subjected to the conditions of morality.[49]

The feeling of justice based on a relationship of forces is the cause of punishment. The historian of morality must study the real utility of defined

[45] eKGWB/MA-105; Engl. transl., n 27 above, 81.

[46] eKGWB/FW-345; Engl. transl. by J. Nauckhoff, F. Nietzsche, *The Gay Science with a Prelude in German Rhymes and an Appendix of Songs*, ed. by B. Williams (Cambridge: Cambridge University Press, 2001), 203.

[47] See M.C. Fornari, n 25 above, 83.

[48] ibid 116.

[49] eKGWB/NF-1883,16[15]. The quotation contained in the fragment's excerpt is taken from: P. Rée, n 26 above, 49.

good and bad actions, not only the origin of moral judgments. Thus, he will discover that punishment does not originate in an alleged purpose (a critique of the utilitarian point of view), but is merely functional, has taken on many meanings over time (and therefore has been used for different purposes) and its origin is not easily identifiable.[50]

Specifically, on the question of punishment, its purpose and its origin, Nietzsche reaches the following conclusions:

a) Justice does not come from resentment (*ressentiment*). Active affections have priority over reactive ones. While resentment is reaction, the lust for domination, the desire for possession have priority, they are the action that generates the reaction. This against Dühring who considers 'the seat of justice is found in the territory of reactive sentiment'. The active man is therefore closer to the justice of the one who reacts![51]

b) The law represents the fight against feelings of reaction. That is to say, it consists of a stop to the release of the reactive feeling and also to the constraint of an agreement. And this in various ways:

I. tearing the object of resentment from the hands of revenge;
II. putting in place of revenge the fight against the enemies of peace and order;
III. devising and imposing agreements;
IV. raising to laws certain forms of compensation for damage;
V. above all, through the establishment of the law (what is permitted and legitimate vs. what is prohibited and illegitimate);
VI. treating any illegitimate action as an infringement of the law and therefore departing from the victim's perspective, in favor of an impersonal evaluation of the action.[52]

c) Punishment is not born to cause guilt, 'bad conscience' or remorse. Erroneously it is thought that the value of the punishment lies in its arousing in the offender a feeling of guilt, that is to say it is the instrument to provoke that psychic reaction called '"bad conscience" or "pang of conscience"'. In reality, genuine remorse is rare among criminals: penitentiaries are the least suitable places to give birth to this feeling. So 'the evolution of feeling of guilt was most strongly impeded through punishment' due to judicial and punitive procedures, acts that closely resemble those being punished ('spying, duping, bribing, setting traps', 'robbery, violence, slander, imprisonment, torture and murder'). Bad conscience 'did not grow in this soil', because for centuries judges have never thought of having to deal with a guilty person, but with 'someone who had caused harm, an irresponsible piece of fate'.[53] A bad conscience and

[50] See M. C. Fornari, n 25 above, 113.

[51] eKGWB/GM-II-11; Engl. transl. by C. Diethe, F. Nietzsche, *On the Genealogy of the Morality*, ed. by K. Ansell-Pearson (Cambridge: Cambridge University Press, 2nd ed, 2006), 48.

[52] See ibid; Engl. transl., 49-50.

[53] eKGWB/GM-II-14; Engl. transl., 54-55.

remorse are born therefore by effect of traditional morality, not of justice.

d) The incommensurability of punishment and guilt. The 'instinct of freedom' is 'forced back, repressed' by the rules, so much so that it does not find in individuals another object on which to be discharged if not on themselves: this is the origin of a 'bad conscience'. Then it becomes a 'debt towards God', so that our 'animal instincts' are a 'guilt before God'. Thus 'the punishment' will never be 'equivalent to the level of guilt'! Therefore, bad conscience is a 'sickness'.[54]

e) Punishment has only a functional role. In punishment it is necessary to distinguish between the 'permanence' of a succession of procedures ('custom': *den Brauch*) and the 'fluidity', relative to their execution ('purpose': *den Zweck*). Genealogists, Nietzsche says, were wrong to believe that the procedure had been devised for the purpose of punishment. It is, however, impossible today to say why you arrive at the punishment. It can only be argued that the punishment is a synthesis of meanings that forms a unit that is difficult to analyze and completely impossible to define or reduce to a single meaning.[55]

f) Punishment serves only to lower the other.

> The sense of punishment is not that of being a deterrent, but rather that of putting someone lower down in the social order: he is no longer one of our peers.[56]

It turns out that the question of the intentionality of doing damage does not matter, but the mere fact of having been damaged and how much. From here follows the punishment, which humiliates a peer who has broken a pact of peace and loyalty founded on the fictitious presupposition of an equality of feelings and actually founded on a relationship of forces, which is subverted by the offender, with respect to whom the power of the one who commands (and exercises justice) must be valid again.[57]

g) In its evolution penal law is being mitigated, up to an ideal self-suppression of justice. This happens with the growth of power and self-awareness of a community. It is like a relationship between a creditor and a debtor: the richer the former becomes, the more compliant he is. The maximum of 'power' (*Macht*) would be the one in which the community leaves its offenders unpunished, as 'parasites' of which it does not fear the effect. This is a 'self-suppression (*Selbstaufhebung*) of justice', known as 'mercy', which goes beyond the law itself.[58] Note in this regard the right-moral homology: in *The Dawn* (1881) Nietzsche speaks of 'the self-suppression (*Selbstaufhebung*) of morality'.[59]

[54] eKGWB/GM-II-17 and eKGWB/GM-II-22; Engl. transl., 57-58, 63-64.
[55] eKGWB/GM-II-13; Engl. transl., 52-54.
[56] eKGWB/NF-1883,16[29].
[57] See M.C. Fornari, n 25 above, 116.
[58] eKGWB/GM-II-10; Engl. transl., n 44 above, 47-48 (translation modified by us).
[59] eKGWB/M-Vorrede-4; Engl. transl. by B. Smith, F. Nietzsche, *Dawn. Thoughts on the Presumptions of Morality* (Stanford: Stanford University Press, 2011), 6 (translation modified by us).

Therefore, it could be concluded that for Nietzsche right in itself does not exist and the social necessity for penal law is based only on contingent political considerations, in a further radicalization of the thought already expressed at the time of Human, All Too Human. But with results that do not deviate dramatically from those at the time reached by the materialism of La Mettrie.

IV. Nietzsche Reader of La Mettrie: An Open Question

After having illustrated both the theory on the legitimacy of punishment of La Mettrie and that of Nietzsche, it is appropriate, before verifying whether it is possible to prove a direct reading of La Mettrie by Nietzsche, to attempt a direct comparison between their doctrines to evaluate in a more analytical way, similarities and differences, beyond the generic feeling of familiarity that the reader feels by combining certain passages of the two authors.

From a comparison between the doctrines of La Mettrie and Nietzsche different and significant common doctrinal elements emerge, which we summarize here in a synthetic way:

i. the determinism applied to human action, from which moral irresponsibility follows;

ii. the arbitrariness of good and evil, deprived therefore of metaphysical ground;

iii. the existence of a natural foundation of human actions, rejected on a moral level and concealed for social and political reasons, so as to produce the distortion of individuals in favor of a certain project of domination based on prejudice and fanaticism, whose overall human costs are unsustainable;

iv. the idea that justice and penal law have a political-social and non-philosophical foundation;

v. the uselessness of remorse, which, since harmful and unfounded, must be eliminated;

vi. punishment, unfounded from the philosophical point of view, assumes an exclusively social function: there are 'bad' actions for society as a whole, that is actions that damage it and that must be punished for this (a doctrine that however is valid for Nietzsche of Human, All Too Human, but becomes problematic for that of On the Genealogy of Morality).

However, there are also differences between the two doctrines which we can summarize as follows:

i. while for La Mettrie our actions are oriented by the pursuit of pleasure, Nietzsche, especially from On the Genealogy of Morality onwards, considers this inexact, as they are rather oriented by the will for power;

ii. according to La Mettrie, egoism is harmful to society even if it is not reprehensible: Nietzsche cannot agree because for him egoism is useful and it is necessary to give him back a 'good conscience';

iii. for Nietzsche the philosophical critique of morality can call into question the same value of morality and therefore to envisage his own 'self-suppression' and therefore 'good' and 'bad' are not only unfounded but useless: it is an alien problem to La Mettrie, who tends rather to distinguish the fields, on the one hand, of philosophy and on the other of morals, law and politics, endowed with different claims and aspirations.

From this comparison it is clear that on a purely theoretical level it is not at all unreasonable to suppose that Nietzsche may have thought of the problem of the legitimacy of the penalty by confronting the thought of La Mettrie. Indeed, it seems more than probable. What remains to be defined, however, is whether it was a direct or indirect comparison. So let us take a closer look at the elements in our possession in order to try to answer this question.

First of all, with today's computer systems, it is possible to establish with accuracy any occurrences, in the Nietzschean work, of the name of La Mettrie or direct quotations from his work. From this investigation it emerges that the name of La Mettrie is never mentioned by Nietzsche, including letters and fragments, just as there is no citation from the work of this author in the entire Nietzschean corpus, including letters and fragments.[60] However, it is possible to find in this corpus the expression 'man-machine', which does not appear as such, but disjointed in the form of a 'machine "man" ' (*machine "Mensch"*) and in a context that has no explicit connection with the theories of La Mettrie.[61] Just to make a comparison with another possible Nietzschean source, or the philosophy of Helvétius, the name of this philosopher occurs in six textual places and a copy of his *De l'homme* in German translation is still preserved in the BN, with numerous traces of reading.[62]

[60] The survey was conducted on 10.01.2019 within F. Nietzsche, *Digital critical edition of the complete works and letters*, based on the critical text by G. Colli and M. Montinari (Berlin/New York: de Gruyter 1967-), ed by P. D'Iorio, on the web site http://www.nietzschesource.org/ with the following result: 'The expression "La Mettrie" ' occurs in 0 textual units'. Same result by entering the search engine exclusively "Mettrie".

[61] The only occurrence found by us is eKGWB/NF-1884,25[136]. There are also seven occurrences of the expression 'man-plant' or 'plant "man"' (*Pflanze Mensch* or *Planfze "Mensch"*), which apparently refers to the title of one of the works of La Mettrie (eKGWB/NF-1884,27[40]; eKGWB/NF-1884,27[59]; eKGWB/NF-1885,34[74]; eKGWB/NF-1885,34[146]; eKGWB/NF-1885,34[176]; eKGWB/NF-1885,37[8]; eKGWB/JGB-44), but the context has no explicit connection with the theories of the French philosopher. In fact, the true source of this expression has been recognized in Stendhal's *Rome, Naples et Florence* (Paris: Michel Lévy Frères, 1854), 383. See: F. Nietzsche, *Einführung Siglenverzeichnis Kommentar zu den Band 1-13*, in: Id, *Sämtliche Werke. Kritische Studienausgabe in 15 Bänden*, ed. by G. Colli and M. Montinari (München/Berlin: Deutscher Taschenbuch Verlag/de Gruyter, 1980), *Band 14: Einfürhung – Siglenverzeichnis – Kommentar zu Band 1-13*, 354, 724-725**;** N. Regent, *A 'Wondrous Echo': Burckhardt, Renaissance and Nietzsche's Political Thought*, in H.W. Siemens and V. Roodt eds, *Nietzsche, Power and Politics. Rethinking Nietzsche's Legacy for Political Thought* (Berlin/New York: Walter De Gruyter, 2008) 629-665, particularly 654-657; G. Campioni, *Der französische Nietzsche* (Berlin/New York: Walter De Gruyter, 2009) 182.

[62] The name appears, to be precise, in: eKGWB/NF-1883,7[19]; eKGWB/NF-1883,7[77];

Precisely with regard to the Philosopher's Library it is opportune to recall the existence of an 'ideal library' alongside the real one, that is, the one formed by the volumes belonging to Nietzsche and which have been preserved until today, as was well illustrated by Paolo D'Iorio, where he distinguishes between 'library' and 'readings':

> The first set, the library, includes all the books owned by the author. The second set, the readings, includes all the books that the author read, whether they belonged to him or not. The first set includes volumes that Nietzsche did not read, at least in that preserved specimen, while the second includes books that the philosopher read with passion, from which he drew quotes, which influenced the development of his thought, but which we do not have evidence possessed.[63]

In the wake of these considerations we can suppose that Nietzsche had in his hands the works of La Mettrie – on loan from a friend, consulted or borrowed from a library, or purchased and then resold or given away or lost – although not those being included today in the BN. Besides, there were several volumes that were part of it, lost especially in the first years after the death of the philosopher, in addition to those deliberately destroyed by his sister Elizabeth because she considered them immoral or harmful to her brother's posthumous reputation:

> As for the library, the disappearance does not only concern, as she herself acknowledged, a series of volumes lost in Paraguay, among which the *Tales* of Bret Harte, or *Henry the Green* and *The People of Seldwyla* by Gottfried Keller, but even those that she herself offered as a gift to illustrious visitors and patrons of the Nietzsche-Archive, as well as those who suffered her censorship: she did in fact throw away some works that she considered embarrassing or scandalous (something she never recognized). Montinari believed that, in order to preserve his brother's good reputation, he had made Stirner's *The one and his property* disappear, a highly condemnable reading in his eyes, as well as some novels with licentious content, or at least those which she regarded as such, as Stendhal's *De l'amour*.[64]

If we consider that Stirner's work was the object of a purge, a similar fate could be touched in La Mettrie! And at least two hypotheses can be attempted in this regard. According to the first, Nietzsche, on the wave of Lange's youth reading,

eKGWB/NF-1884,25[366]; eKGWB/NF-1885,34[39]; eKGWB/JGB-228; eKGWB/NF-1888,14[97]. As for the copy of the *Discurs über den Geist des Menschen* by Helvétius still preserved in the BN see: G. Campioni et al eds, n 17 above, 289.

[63] P. D'Iorio, 'Geschichte der Bibliothek Nietzsches und ihrer Verzeichnisse', in G. Campioni et al eds, n 17 above, 33-77, 68.

[64] ibid 33-34.

around 1875-76, reads La Mettrie when the first translation of the *L'homme plus que machine* is published in German, edited by Adolf Ritter.[65] And perhaps he goes to look at the second volume of the works of La Mettrie cited in the note in the second enlarged edition of the book by Lange, in which the latter refers in a note to the *Discourse on Happiness*,[66] edition of *History of Materialism* which, as we have already mentioned, according to some scholars Nietzsche may have had in his hands. According to the second hypothesis, which postpones the direct reading of La Mettrie by Nietzsche, the latter could have read it directly, in French and/or in Ritter's translation, in the wake of his careful re-reading of Lange, carried out on the fourth edition of his book, purchased in 1887.

In fact, if we take the copy of Lange's book that belonged to Nietzsche and is still kept at the BN, we find traces of reading on page 257, in the chapter dedicated to the influence of English materialism in France and Germany. On that page we read:

> Nothing was left but to make the experiment of placing sensation as a property of matter in the smallest particles themselves. This was done by Robinet in his book on 'Nature' (1671), while La Mettrie in "L'Homme Machine" (1748) still kept to the old Lucretian conception.[67]

Other traces appear on page 344, in the chapter on German materialism, where the concept of 'Homme Machine' (*Maschinemann*) in direct relationship with La Mettrie is recalled.[68] It should be noted that on the pages immediately preceding and following the name of La Mettrie is mentioned several times, concerning the German reception of his most famous work, *L'homme machine*, the title of which is repeated five times on four pages.[69]

What can we conclude from this? That Nietzsche could not have not read the chapter expressly dedicated by Lange to La Mettrie (although we have no traces of reading on the pages of that chapter in the edition kept in the BN). That Nietzsche probably also read the two pages that Hermann Hettner, in his *Geschichte der französischen Literatur im achtzenten Jahrhundert*, dedicate to La Mettrie.[70] That Nietzsche was certainly aware of the existence of Ritter's

[65] J.O. de La Mettrie, *Der Mensch eine Maschine. Übersetz, erläutert und mit einer Einleitung versehen von Adolf Ritter* (Berlin: Erich Koschny, 1875).

[66] In the second edition, including notes, Lange refers to several French editions of the works of La Mettrie, including the Berlin edition of 1774, in 8th (in two volumes), and the edition published in Amsterdam in the same year in 12th (in 3 volumes): *Œuvres philosophiques. De Mr. de La Mettrie, corrigée & augmentée* (Amsterdam: s. n., 1774). From this last Lange quotes the *Discours sur le bonheur*, which is contained in volume II, on pages 95-190. See F.A. Lange, n 10 above, vol 2, 79 (n 75), 83 (n 77), 84 (n 79).

[67] ibid vol. 2, 29.

[68] ibid vol. 2, 140.

[69] Two occurrences of the title are counted in ibid vol. 2, 137 and three occurrences of the same in ibid vol. 2, 138.

[70] H. Hettner, *Literaturgeschichte des achtzenten Jahrhunderts, 2. Th.: Geschichte der*

German translation, but that it is not certain that he ever had it in his hands nor did he have the French works of La Mettrie, since he does not insert any direct quotation from these in his writings.

While awaiting further investigations and discoveries, we believe, however, that it is right to claim that Nietzsche, surely stimulated by Lange's reading, retraced the steps taken by La Mettrie, leading to even more extreme consequences of his radical materialism and his conception of punishment.

französischen Literatur im achtzenten Jahrhundert (Braunschweig: Vieweg, 1860), 373-375. La Mettrie is discussed by Hettner in the second of the three volumes of which his work is composed. Nietzsche quotes Hettner's work in a generic way in a letter to his mother dated 2 May 1863: eKGWB/BVN-1863,353. There is, however, a certain trace that Nietzsche possessed the second volume of the work, which contains the pages on La Mettrie; see the relative file in the BN: G. Campioni et al eds, n 17 above, 300.

Short Symposium on the Punishment

The Mith of Re-Education

Patrizio Gonnella*

Abstract

The prison model has won, as it has developed from the sixteenth-century workhouses to the ten million detainees currently imprisoned in the world. Despite these huge numbers, our penal and penitentiary legal framework is all about the myth of re-educational treatment. The treatment model is progressively overflowing, as was inevitable, towards a disciplinary model. Everything in prison is treatment and everything is discipline. Shifting the spotlight from the re-educational utopia to human dignity and the rights deriving from it helps to read the aporia of prison, helps to reestablish the foundation of the prison system in a clearer way, imposing ethical limits that cannot be crossed and making it compatible with the rules of the welfare state and the rule of law.

I. The Legal Functions of a Sentence. Prison Has Won

Penal law jurists, constitutionalists, philosophers of law, law historians, prison officers, ordinary and constitutional judges have been wondering for decades about the legal and formal function of punishment. However, they argue about an abstract punishment, a punishment written in rules and codes but that does not exist in reality.

More in touch with the political, social, criminal and prison reality was a scholar and politician, Arturo Rocco,[1] the father of the Italian penal code currently in force, who wrote as follows:

> It is evident how this character of necessary defence (of the vital interests of the nation) can be found, not only in those crimes that directly attack the existence or security or state, but also in those serious common crimes that, for the atrocious manner in which they are committed, and in the absence of mitigating circumstances, denote in the guilty such perversity as to render all hope of amendment and re-education vain. (…) The severity of sentences cannot be justified, except by a concrete and immediate purpose of more

* Professor of Sociology of Law and Philosophy of Law, Roma Tre University; President of Antigone, an association involved in the penal issues and human rights of prisoners; Expert for the Council of Europe to monitor places of deprivation of liberty.

[1] He was Minister of Justice and for Religious Affairs for seven years between 1925 and 1932. His Penal Code is imbued with fascist ideology. His report to the code is an ideological manifesto of the fascist legal thinking.

vigorous repression.[2]

Alfredo Rocco did not need to lie about the effective function of punishment. He admitted that the punishment system must serve the apparatus of repression. He was the jurist ideologue of fascism, and the fascist regime could afford to claim an idea of punishment as an affliction, without those hypocrisies that are necessarily and inevitably present in democracies.

In the last seventy years, a wide-ranging debate has taken place about the legal functions of penal sanctions. Still today, there is broad and meticulous concern with the doctrine and jurisprudence on the old and noble theme of the function of punishment, trying to disengage the sentence from the needs of social defence and mere repression. Academics, judges and lawyers, at all levels, discuss, investigate, write essays and sentences on what the function of punishment should be, without questioning the material essence of what has been universally – in space and time (for at least two centuries) – considered the only punishment worthy of the name, namely imprisonment. To understand what the function of a penalty is, the material and daily nature of that penalty must be understood.

Any answer to the main questions about what the function of a criminal sanction is or should be cannot ignore knowledge of the system of penalties and imprisonment in its concreteness. No scientist would renounce direct field observation to better understand the good state of his theoretical studies.

First of all, a view of reality helps to clear the way for a first interpretative misunderstanding. Although the whole contemporary legal culture is oriented towards the search for alternatives to detention, prison dominates the planet. Prison is, indeed, the only sanction considered as such by politicians, public opinion and, not rarely, by judges and security actors. All the other criminal sanctions, starting from the pecuniary one, are marginal, envisaged for minor crimes.[3] They have favoured the extension of criminal law towards areas that traditionally belonged to civil or administrative law, without affecting the centrality of prison as the main penalty.

What happened in Italy in 2013 and 2014 is paradigmatic of the ineffectiveness of non-custodial criminal measures in reducing the use of prison as the main sanction. Non-custodial measures are incapable of taking away the centrality of a prison sentence. In the daily work of social services and courts, alternative measures to detention and trial have taken up what would probably have been the space of non-punishability. Yet the declared normative function

[2] Preparatory work for the Penal Code and the Code of Criminal procedure, V, Final project, Report by Alfredo Rocco. Part I, 68-69.

[3] The forecasts made by G. Rusche and O. Kirchheimer, *Punishment and social structure* (London: Routledge, 2003) were wrong. In any case, to them we owe the investigation of the relationship between social, economic and penal relations. The book was translated for the first time in Italy by Dario Melossi and Massimo Pavarini.

was different. Sanctions other than imprisonment have usually broadened the area of punishment without reducing the area of imprisonment.

In January 2013, Italy was condemned by the European Court of Human Rights in the Torreggiani case.[4] It was a pilot sentence because it concerned thousands of applications filed by prisoners forced to live in very limited spaces. According to the Strasbourg Court, every detainee must have at least three square metres, otherwise, the state will incur violation of Art 3 of the Convention, prohibiting torture and any inhuman or degrading treatment. Italy was called upon to take systemic steps to reduce overcrowding, as well as to ensure effective judicial remedies in the case of prisoners' complaints. Among the measures taken by the Italian Parliament was the extension of the institution of probation from the juvenile system to that of adults.[5] Judicial probation consists, at the request of the defendant, in suspension of the criminal proceedings at the first instance decision stage for crimes of low social alarm. In the space of a few years, great recourse has been made to probation in the adult system. The number of measures in progress rose from eight hundred four on 31 January 2015 to seven thousand three hundred forty five in 2016, nine thousand two hundred seven in 2017 and eleven thousand one hundred two in 2018. The increase continued during 2018, with thirteen thousand four hundred eighty-one measures in progress on 30 June and fourteen thousand nine hundred eighty on 30 November 2018. At the same time, however, the total number of prisoners also increased, despite the decline in crime rates. Therefore, probation did not result in an erosion of the prison area. It probably affected all those people who would not have been imprisoned anyway. Prison has continued to be the main, sometimes the only, penalty, the one that no one is willing to give up. It continues to be used as a threat and is applied without hesitation and without taking crime rates into account. All this happens because prison responds to a social, public, and political need for affliction, revenge, and neutralisation that has nothing to do with the legal function of the penalty itself.

Prison is a penalty that has become the only penalty around the world in the last two centuries. Even in our constitutional and penal law, the word penalty is written in its plural form; however, in reality, penalties tend to be just one: prison. A prison sentence is used as a threat, executed, condoned, suspended, amnestied and converted. But the main penalty is always the same,

[4] Eur. Court H.R., *Torreggiani et al* v *Italy*, Judgment of 8 January 2013, available at www.hudoc.echr.coe.it.

[5] It was introduced with Legge 28 April 2014 no 67, which modified respectively: the Criminal Code, with the provision that introduces the new institution in Arts 168-*bis*, 168-*ter* and 168-*quater*; the Code of Criminal Procedure, with the introduction of Arts 464-*bis* and following which regulate the activities of investigation of the proceedings and the trial, as well as Arts 567-*bis* which indicates the modalities of the evaluation of the probationary period; the norms of implementation, coordination and transition of the Code of Criminal Procedure; the Consolidated Text on the subject of the legislative and regulatory provisions on the subject of criminal records.

ie prison. Jurists need to come out of their bubble and look at the facts. Prison is the penalty. Prison has won in the criminal justice system in time and space, imposing itself as the penalty. Whether we look at trials and courts, but even more if we look at social perception of the penalty system and at numbers, the prison model has won. It has exceeded the limits of criminal justice by returning to its origins, which were those of administrative detention. The prison model is widely used in the field of administrative detention of migrants, and the institutionalisation of the elderly or minors. The birth of prisons has its origin in administrative detention. Today, the prison criminal model also incorporates its origins. In the English correctional workhouses and the Dutch workhouses (Rasphuis), in the past (in the XVI century) there were perpetrators of minor offences, vagrants, petty thieves, beggars by administrative or judicial decision.[6] The punishment of imprisonment has no competitors in its being a force of symbolic reassurance, in its capacity to respond to a plurality of functions or in its being without a precise legal aim.

II. The Social and Economic Function of a Sentence

Crime and punishment are not related. For the first time, in 1939, the two scholars from Frankfurt, George Rusche and Otto Kirchheimer, investigated punishment and how its origin was not criminality. They wrote that penalty has changed in modern times because the model of economic production has changed. They used the historiographic method to explain all the changes. Punishment is not a simple consequence of crime, nor the hidden side of it. It is undeniable that a penalty has specific functions but it is equally undeniable that it cannot be explained based only on these. In terms of punishment and social structure, the two scholars trace back the changes in punitive forms to economic-social systems' change. Their work suggests that it is possible to study economy and society through criminal law and punishment in practice.

> The criminal system of every historically determined society is not something isolated, subject only to its specific laws, but is an integral part of the entire social system.[7]

If punishment and crime are not univocally related, the purpose of a penalty requires more complex explanations. On the other hand, this lack of connection between punishment and crime is perfectly proven by statistical data. Detention rates and crime rates are not moving in parallel. Detention

[6] D. Melossi e M. Pavarini, *Carcere e fabbrica. Alle origini del sistema penitenziario* (Bologna: il Mulino, 1st ed, 1977, 2nd ed revised, 2018). See also M. Ignatieff, *Le origini del penitenziario. Sistema carcerario e rivoluzione industriale inglese* (1750-1850) (Milano: Mondadori, 1982).

[7] G. Rushe and O. Kirckhheimer, n 3 above.

rates also rise when (as in recent years in Italy) crime rates fall. Massimo Pavarini started his sociological studies from an awareness of the absence of a link between punishment and crime. However, he is not interested in giving explanations entirely within the capitalist model and goes to the heart of the anthropological needs of punishment present in society. To understand what a penalty is, the interpreters must know daily life in a prison.

Massimo Pavarini[8] defines the essential and factual attributes of a penalty from a sociological point of view. He is not theoretically interested in the legal nature of a penalty because the normative functions legitimise the legal penalty, but are unable to fully explain it. The sanction, according to him, has four attributes: the punitive one (production of deficits against the person punished, ie reduction of his/her rights and/or reduction of satisfaction of his/her needs), the programmatic one (the repressive action must appear explicit and intentional so that the punished person feels it as a censure against him/her), the expressive one (the sanction must distinctly and symbolically express the claim of authority of those who punish) and the strategic one (it must be such as to perform the function of maintaining certain power relationships). The four characteristics of punishment, according to Pavarini, explain the essence of the punishment itself.

A penalty is to punish, to inflict pain. If this were not the case, some details of life in prisons would not be explained. To prohibit a prisoner from meeting his son/daughter whenever (s)he wants has a purely punitive value. This is a rule designed only to create pain. It has nothing to do with the legal, preventive or re-educational function of punishment. The Constitutional Court in 1993, in a very relevant decision, stated:

> Those who are in detention, even if deprived of most of their freedom, always retain a residue, which is all the more valuable because it is the last area in which they can expand their individual personality.[9]

Therefore, it is the same court that recognises that the sanction of deprivation of liberty draws everything into it and can determine the loss of fundamental rights other than personal liberty. For this reason, judges' attention must be maximised.

A punishment, as Pavarini argued in depth, has a programmatic aspect. It makes sense to programmatically exclude a detainee from all sources of information. In this way, it is possible to programme his/her social exclusion over time. Prohibition to use a PC and an Internet connection for people with long sentences has a clear will to reduce the person to an unthinking animal and to exclude him/her from the community of the people already included. The rule responds to a need for socio-economic planning. In this way, in fact,

[8] M. Pavarini, *Governare la penalità* (Bologna: Bononia University Press, 2014).
[9] Corte Costituzionale 28 July 1993 no 349, available at www.cortecostituzionale.it.

prison management prevents any attempt to break the social barriers.

A punishment must be expressive. This is the only reason why prisoners are forced to humiliate themselves for every request they have to make in the prison context. They have to submit to authority, represented by prison guards. And if they leave the prison to work, they must do so for free and clearly visible by the crowds of innocent people. It is never clear in prison what the source of any provision is. An imposition does not always have its own legal legitimacy. Often it is the result of an internal regulation of the prison central administration, or even more frequently, it is the result of an order issued by the director (warden) of the prison. In prison life, it can happen that what is considered a rule is instead a mere practice, completely devoid of formal legitimacy. The deep sense, in the world of prisons, of such disinterest in the legality and formal hierarchy of the sources, is given by the will to distinctly highlight the unlimited power of the guards over the people deprived of their liberty. It is easier to do so if this punitive power is clearly arbitrary and not legally confined. In this way, the imbalance with the prisoners will appear even stronger.

Finally, a punishment must be strategic. The more migrants there are in prison, the more the poor are imprisoned, the more the balance of power between the different social groups is left unaltered. The entire system of procedural guarantees is designed for a type of person (indigenous, wealthy, well-educated on average) who rarely passes through the criminal system. All the others – that is, those already excluded from society and welfare –will find it difficult to avail themselves of those guarantees. Thus, once they have ended up in prison, the selection, which is not so much of class today, but of ethnicity, passport, wealth and status, is perpetuated.

According to Justice offical data reported in June 2019, by adding detained foreigners and prisoners from the four most populous regions of southern Italy, we obtain seventy-seven percent of the total number of people detained in Italy. By adding the prisoners from Sardinia, Basilicata, Abruzzo and Molise the percentage exceeds 80percent. The rest of the country, which tends to be richer, produces only one-fifth of the detained population, even though it makes up about two-thirds of the free Italian population. Recent data also tells us that over a thousand prisoners, three hundred fifty of whom are Italians, are illiterate. In free Italy, illiterates add up to zero point eight percent. In prison, the percentage is more than one point five percent. Moreover, six thousand five hundred prisoners, more than ten percent of the total, only have an elementary school certificate. University graduates amount to just over one percent (six hundred ninety-eight), while in free society they reach eighteen point seven percent. This is the strategic framework to imprison poverty. Investing in education and welfare would constitute an extraordinary form of crime prevention that in the long run would produce security. However, the penalty of imprisonment has no connection with security, but with symbolic reassurance

and social and political consensus.

III. Global Mass Incarceration

The prison model has therefore met with undeniable success, as it has developed from the sixteenth-century workhouses to the ten million detainees currently imprisoned in the world. An impressive number that cannot but have its own close link with the prevailing social and economic model. Given the exponential growth in prison numbers over time, and considering that no political system renounces prison as a form of social control, mass incarceration can be explained by the recourse to several competing causes. It could be linked to the planning of mass discipline[10] or to the complex demonstration of the use and abuse of prison by the capitalist system[11] or also to the gradual withdrawal of the welfare system.[12] The numbers are so high that it is not possible to be satisfied with simple explanations. The legal function of a sentence is challenged by mass incarceration and by the centrality of the custodial model even outside the penal system. In the United States, there are about 2.2 million prisoners. In light of the total number of people living in the US, the detention rate is incredibly high: for every one hundred thousand inhabitants, about six hundred fifty-five are in prison (data from the Bureau of Justice Statistics). After New York, Los Angeles, Chicago and Houston, the American prison population would be the fifth American city in terms of the number of people with a total cost of eighty billion dollars. The federal system has produced a multiplication of the detention level. More than fifty percent of prisoners are held in state prisons. There are as many as three thousand local prisons. A bargain for the security multinationals that manage about eight percent of the total number of prisoners, fuelling the vicious circle of judicial corruption and the financial market without scruples. In recent times, millions of US taxpayers' dollars are being invested in private prison operators involved in the detention of migrants around the United States. At the national level, at least twenty pension funds and plans have invested in the two largest private operator groups (Geo Group and CoreCivic).[13] Though African-Americans comprise only about twelve percent of the total US population, they add up to thirty-three percent of the federal and state prison population. Meanwhile, whites, who constitute sixty-four percent of American adults, amount to only thirty percent

[10] Michael Foucault is always worth reading, as his book is the most organic explanation of the birth and excellent fitness of the prison system. M. Foucault, *Surveiller et punir: Naissance de la prison* (Paris: Gallimard, 1975).

[11] See fn 6.

[12] But also L. Wacquant, Iperincarcerazione. Neoliberismo e criminalizzazione della povertà negli Stati Uniti d'America (Verona: Ombre Corte, 2013).

[13] See the journalistic enquiry of The Guardian available at https://tinyurl.com/57d68bhh (last visited 30 June 2021).

of those behind bars.[14] Given that detainees in almost all American states lose their voting rights (and do not regain them immediately after serving their sentences), this means that a part of the population, the most marginal, is deprived of the opportunity to participate in political choices. Prisoners and former prisoners are out of community life. Therefore, mass incarceration is a source of wild enrichment, as well as functional to neoliberal policies of social and political exclusion of non-productive sectors. The political and economic system uses the custodial model as an easy source of enrichment, to fuel its own gift of consent and to exclude those most reclutant to inclusion. The prison-centred penal system is, therefore, a selective system based on wealth, skin colour and nationality and is functional to leaving power blocks unaltered.

All this is happening not only in the United States but in a similar way and with similar aims, also in Europe and in developing countries. In Italy, the detention rate is one hundred prisoners for every one hundred thousand inhabitants. Much lower than in the United States but the trend (and the social composition of the prison population) is the same and leads to overlapping considerations. If these are the numbers, as David Garland suggests,[15] then criminal institutions should be analysed from the outside, in an attempt to understand their role in a broader perspective, as an expressive form of social change. Having seen the numbers, the costs, the gains and the social composition of the prison population, the re-educational function of punishment is fiction and the social afflictive function of punishment is evident.

IV. Observing Prison and the Myth of Re-Education

It is necessary to see a prison to fully understand its social function and measure its difference compared with the justifying legal functions attributed to punishment. Not having seen a prison is not a fault to be ascribed to the jurist or philosopher: but we cannot ignore the voices of those who have seen one. At the beginning of the last century, it was the powerful voice of Filippo Turati who opened the gates of prison to public opinion and the political world. Prisons 'are the greatest shame of our country', he denounced in a memorable speech to the parliament:

> They represent the expression of social revenge in the worst form that has ever happened: we believe that we have abolished torture, and our prisons are themselves a system of torture, the most refined; we boast of having cancelled the death penalty from the penal code, and the death

[14] Data from Pew Research analysis of Bureau of Justice Statistics. See also CNN available at https://tinyurl.com/ye28d8ph (last visited 30 June 2021).

[15] See D. Garland, *The Culture of Control: Crime and Social Order in Contemporary Society* (Oxford: Oxford University Press, 2001).

> penalty that our prisons administer drop by drop is less compassionate than that which was given by the hands of the executioner; we swell our cheeks to speak of a re-education of the guilty, and our prisons are factories of criminals or schools of improvement of criminals.[16]

With strong words, Filippo Turati expressed his warning to the bourgeois of the *belle époque* of the Italian Giovanni Giolitti's time. Therefore, nothing that recalled a noble perspective of re-education, of correction of bad souls, but only a living in misery and violence. A prison is a factory of recidivism and potential future repeating offenders.

What is Turati saying at the beginning of the XIX century? The abolition of the death penalty does not weaken, on the contrary, it strengthens a hard, degrading, inhuman and violent prison, as the right response, precisely because it is painful to crime. The spread of the prison system has not led to a reduction of the risk of dying in the hands of the state. Of the prison, Turati knew the law void and the concentration of power; the obtuse bureaucracy and despotic hierarchy; the ineffective rules and the painful practices, the unfulfilled aims and the loss of self-determination: how it is much easier to imprison a condemned person, to frighten him, to brutalise him, than to educate him and make him a new man; how ferocity requires neither intelligence, nor effort, nor financial means, while education demands all these things; how of all prison regulations, the brutal ones are widely applied, those in which the spirit of social vengeance against the unfortunate detainee survives. Instead, Turati explained, all those regulations which reflect the duty of the state to provide for the rehabilitation (he used the word redemption) of the guilty person, while at the same time providing public security against recidivism, are left completely aside and have remained a dead letter.

Direct observation makes it possible to reason beyond the functions that the law attributes to punishment and not to limit one's experience to the scholarly debate on the doctrines of general or special prevention, retribution or re-education. It allows us to examine those essential and factual attributes of punishment, identified by the penetrating look of Massimo Pavarini: afflictive, programmatic, expressive and strategic.

Prison also needs an ethnographic investigation to be properly understood. Anthropologists and ethnographers should make their investigations and studies available to expose the real function of punishment.[17] We have understood a great deal about prison thanks to the stories and the analyses of those who have told them from behind bars.

It is necessary to have seen a prison to understand its afflictive nature,

[16] Filippo Turati gave the speech in the (Italian) Chamber of Deputies on 18 March 1904. It was then published in a brochure with the title 'The cemetery of the living'.

[17] A recent ethnographic survey is that of G. Torrente, *Le regole della galera. Pratiche penitenziarie, educatori e processi di criminalizzazione* (Torino: L'Harmattan Italia, 2018).

made up of denied rights far beyond what is explicitly indicated in the law; its programmatic essence, based on oppressive messages and submission of the prisoner to the unwritten rules of the prison; its expressive force, consisting of the symbolism of the asymmetrical relationship between the prisoner and the guard; its strategic character, aimed at not changing the relationships of domination between social classes. Anyone who has visited a prison knows the useful and right indicators to understand the actual social function of a prison sentence: the downward or upward glances of the prisoners, the smell of coffee or rottenness in the cells, the silence or noise, the life or absence of life that there is in the sections, the provision or not of underground spaces for solitary confinement, the use of military uniforms or the absence of uniform for the prison staff. Indicators that go beyond any normative codification and which only direct observation can examine and recognise.

After the end of the Second World War and the fall of fascism, Piero Calamandrei published a monographic issue of the magazine '*Il Ponte*' dedicated to torture, prisons and the need for a parliamentarian reform. The volume contains extraordinary essays written by intellectuals and politicians who were imprisoned during the fascist regime.[18] Calamandrei asked for testimonies about prison from the most important anti-fascist people of his time. So, re-education is completely demythologised. Altiero Spinelli undertook the task of demystifying the praise of prison as a pedagogical tool:

> It is a very strange way of re-educating the one that consists of detaching (a man/woman) from the whole network of social relations, and putting him/her into a set of new rules, to respect in which (s)he no longer needs any sense of responsibility. The prisoner gets up, washes, sweeps, eats, works, rests, speaks, is silent, reads, writes and goes to sleep at the sound of a bell. They ask him/her to be a machine and nothing more.[19]

The myth of re-education had already been spiced for a century and a half with bombastic rhetoric about the amending virtues of work, education and religion: 'those who think that prison, however modified, can be an instrument of moral and social redemption are victims not of an illusion, but of hypocrisy', Spinelli wrote.

Vittorio Foa was of the same opinion:

> we lock them inside four walls, we entrust them to specialists in repression, not to see them, (...) to live in peace. And hypocritically we add

[18] The issue was recently published by P. Gonnella and D. Ippolito eds, *Bisogna aver visto. Il carcere nella riflessione degli antifascisti* (Roma: Edizioni dell'Asino, 2019). The book includes essays by Altiero Spinelli, Vittorio Foa, Giancarlo Pajetta, Ernesto Rossi and many others.

[19] A. Spinelli, 'Esperienza di prigionia', in P. Gonnella and D. Ippolito eds, *Bisogna aver visto. Il carcere nella riflessione degli antifascisti* n 18 above.

that we want them to improve.

One can also disagree from this point of view. But we must be aware that the utopia of re-education, in the reality of prison, always risks plunging into the dystopia of correctionalism.

The autobiographical method[20] can be a suitable instrument of knowledge of prison reality and its ontological irreconcilability with the re-educational function of punishment. The tales of those who have suffered the misadventure of spending a period in prison are one of the ways of verifying what the effective function of a prison is. The stories and reflections of Spinelli and Foa well explain, in the light of their ability to analyse, how much prison is mainly inflicting suffering, triggering the worst feelings in those who manage them. Indeed, guards have infinite power over the bodies of the prisoners. And, as an anthropologist or a psychotherapist could explain better than a jurist, guards can make arbitrary use of this power unless the wall of opacity is torn down. A prison is a closed place, out of reach of external looks. This makes it not very permeable to the demands of social help. It makes it a place potentially at risk of abuse. Eligio Resta is responsible for the great intuition of what has been called the illusion of criminal justice, that thought of progressing by transforming a trial from a private fact to public history and punishment from a public spectacle to a matter taken away from the eyes of the curious and the enthusiasts. The problem is that a penalty taken away from the outside world has become arbitrary, illegal, vengeful and irremediably painful.[21]

Fifty years after Foa and Spinelli, a person serving a life sentence who managed to emancipate himself from crime and took two degrees maintaining the lucidity of the raconteur gives his testimony:

> (the guards) were all dressed in the same way, they moved in the same way and said the usual things. They looked like priests, but they didn't work for God or the devil... For them, if you came to prison, you had done something and if you hadn't done anything it means that you are dumb to be in there. I have learned only now to forgive the men and women who work in prison... In prison, it is difficult not to bend and not to resign, only the bad guys manage not to do so, so I thought of trying to become even more-evil.[22]

There have been dozens and dozens of testimonies from prisoners starting

[20] See D. Demetrio, *Raccontarsi. L'autobiografia come cura di sé* (Milano: Raffaello Cortina editore, 2006). The author theorised the relevance of the autobiographical method both for self-knowledge (as well as self-satisfaction), and for knowledge of the outside world through the story of one's own life.

[21] See E. Resta, *Il Diritto vivente* (Bari: Laterza, 2008).

[22] C. Musumeci is a 'lifer' who wrote essays and books. Together with A. Pugiotto, *Gli ergastolani senza scampo. Fenomenologia e criticità costituzionali* (Napoli: Editoriale Scientifica, 2016).

from Filippo Turati to the present, in Italy and other countries, about the cruelty of punishment, its ambiguities and the attempt to infantilise the detainee.[23]

In 2013, after the sentence against Italy in the Torreggiani case from the European Court of Human Rights, a broad and interesting public and deep debate were opened, involving hundreds of experts convened by the Ministry of Justice within the framework of the General States on Penal Execution.[24] Among the proposals that emerged there was that of overcoming prison infantilisation, ensuring the opening of cells, sections, the self-management of the day according to the will to make detainees responsible for managing their relationship with the time of deprivation of liberty. All this took the name of dynamic surveillance. In a short time, the project was abandoned and the Italian penitentiary system returned to a model of segregated punishment. According to the majority of prison guards (and their unions), a detainee must stay in a cell, always ask permission to do anything during the day, must not express personal opinions, must always be accompanied in any even brief movement within the prison. So, the prisoner should be treated as a child or as a dog (kept on a leash). All of this confirms that punishment is thought of as suffering, based on distrust and not on the prospect of social reintegration.

There is a common thread that links the stories from prison in time and space. Both in the reflections of the anti-fascists during the 1930s and in those of Pierre Clementi, a famous French actor imprisoned in the 1970s in the Regina Coeli prison in Rome:[25] there is no trace of and no possibility for the re-educative function of punishment. According to their stories, the only practiced alternative to a harsh regime of life and suffering is a prison model based on authoritarian paternalism. Here we could turn to the pedagogical culture to understand how it is not through paternalistic management that a path of responsibility and social reintegration is built. Paternalism transforms rights into concessions, legality into discretion.

The same interpretative result can be obtained by reading the reports of the European Committee for the Prevention of Torture, which is authorised to visit, sometimes by surprise, all the places of deprivation of liberty within the framework of the Council of Europe.[26] The monitoring of prisons by international or national agencies responds to the objective of removing prison from the punitive arbitrariness of its managers. It is thus highlighted, through direct observation, what the true afflictive nature of the prison is. All this in an attempt to set limits in the name of human dignity and fundamental rights.

[23] About the process of infantilisation, see also A. Sofri, *Le prigioni degli altri* (Palermo: Sellerio, 1993).

[24] There is still a trace of it here https://tinyurl.com/dmajreu2 (last visited 30 June 2021)

[25] See P. Clémenti, *Pensieri dal carcere* (Roma: Il Sirente, 2007).

[26] See www.cpt.coe.int.

V. Human Dignity as the Only Limit That Works[27]

Prison as punishment is an invention of modernity connected with great questions that transcend it: from the model of economic production to the ideology of work, from the more general objectives of justice to the more specific theme of the rite of a criminal trial. Prison as punishment has to do with social and fiscal systems, with urban and architectural choices, with human rights and the residue of their executability, with the dignity of the body and the salvation of souls, with ethics and religion. Prison as punishment is within the system of law but is historically not inclined to be caged by law. It is the result of a judgment that turns into prejudice. Prisons are not to be reduced to a historical description; they should also be read through an epistemological investigation that uses the classical categories of space and time. As seen above, prison as punishment imposes a reflection on its function and its methods of execution.

A prison sentence in a democratic society has unsurpassed limits, imposed by the legal system and by ethical sense. Limits that can be traced back to protection of human dignity understood in its Kantian meaning of humanity and the impossibility of treating men as mere means to achieve an end. Art 27 of the Italian Constitution, in its third paragraph, envisages that

> punishments may not consist of treatments contrary to the sense of humanity and must aim at the re-education of the person sentenced,

and suggests not putting a re-educative function and respect for dignity in competition. The scholars of punishment, starting from jurists, but not exclusively them, over time have chosen the function of punishment as their main area of interest, which is the second of the constitutional objectives. Around it, reforms have been built and cancelled, and opposing theses have been endorsed. Some have erected, not only metaphorically, monuments to redemption[28], those who have sought to eliminate the non-educable and those who have developed a model of prison open to the territory and aimed at the social recovery of convicts. In all these cases the same constitutional expression was evoked and used. Re-educational rhetoric, unrelated to human dignity, has for decades hindered the emergence and consolidation of a conceptual, normative and jurisprudential reflection on the first of the constitutional objectives, ie a penalty according to humanity. Mass incarceration and penal populism have produced a macroscopically illegal prison.

This is difficult to tolerate for a liberal or constitutional democracy. Therefore, human dignity works as a limit to the excessive and arbitrary power

[27] This paragraph follows the first chapter of the book of P. Gonnella, *Carceri. I Confini della dignità* (Milano: Jaca Book, 2014). Regarding human dignity in prison see also M. Ruotolo, *Dignità e carcere* (Napoli: Editoriale Scientifica, 2014).

[28] This happened in front of the Pisa prison in the nineties.

to punish, but only if you give up the rhetoric of re-education. If re-education, even good and non-invasive re-education, was reduced to a myth and public attention inevitably shifted around humanity, there is a concrete possibility of more effectively guaranteeing prisoners' human rights. Since 2010, observers, judges and academics have placed human dignity and human rights in the spotlight. The absence of a personal minimum space evoked images reminiscent of the tragedy of the Holocaust or the great tragedies of the last century.[29] The limit had been exceeded.

The aim of re-education works worse than human dignity as a limit to oppose an illegal and violent punishment; this has been noticed by the Supreme Courts in the United States as well as in Europe[30] in Germany[31] and the courts in Italy.

What took place in the United Staes of America is paradigmatic. In 2011, the Governor of California, following an order of the United District Court of California of 8 April 2009, which required him to prepare a suitable plan within forty-five days to reduce the number of detainees by at least forty-six thousand people in two years, ordered a real emptying of Californian prisons. The judges of the United District Court of California found a violation of the eighth amendment of the US Constitution, which prohibits the use of cruel and inhuman punishment. The decision was then confirmed by a ruling[32] of the Supreme Court of the United States, which was also called upon to decide on the violation of rights in the overcrowded Californian prisons.

The re-educational paradigm works worse because of correctionalism – the idea according to which through the prison sentence the detainee should be 'corrected' of his deviant nature – is not conceptually and logically an antithesis to treatments contrary to the sense of humanity. It is so in its democratic version, it is so in the intentions of many scholars and social and legal workers,

[29] It is no coincidence that the UN minimum standard prison rules are named after Nelson Mandela.

[30] See the case-law about art 3 of the Convention from the Eur. Court H.R., *Sulejmanovic* v *Italy*, Judgment of 16 July 2009 to Eur. Court H.R., *Mursic* v *Croatia,* Judgment of 20 October 2016, available atwww.hudoc.echr.coe.it**.**

[31] Bundesverfassungsgericht 22 February 2011, 1 BvR 409/09 available at www.bundesverfassungsgericht.de. The German Constitutional Court, following the appeal of a detainee who complained of particularly harsh conditions of imprisonment (twentythree hours a day closed in a cell of eight square metres to be shared with another detainee who smokes), stated that the state must ensure full respect for human dignity even by renouncing application of the penalty. In Germany, the court has a stronger juridic instrument, as Art 1 states that: 'Human dignity is inviolable. To respect and protect it is the duty of all state authority. The German people, therefore, acknowledge inviolable and inalienable human rights as the foundation of every human community, of peace, and justice in the world'.

[32] *Brown* v *Plata* 131 S. Ct. 1910 US Cal. (2011). See G. Salvi, 'La Costituzione non permette questo torto: la Corte Suprema degli Stati Uniti e il sovraffollamento carcerario' *Questione Giustizia*, 205 (2011) and M. Lombardi Stocchetti, 'Il carcere negli USA oggi. Una fotografia', *Diritto Penale Contemporaneo*, 23 December 2014, available at https://tinyurl.com/yf8vhpys (last visited 30 June 2021). For a critical analysis of mass incarceration and the role of the courts see S. Anastasia, *Metamorfosi penitenziarie* (Roma: Ediesse, 2013).

but it is not so everywhere and in any case. The correctional model – even in its most modern, less paternalistic and authoritarian versions – always carries the germ of the instrumentalisation of a human being for another function. The detained to be re-educated becomes a means to achieve his/her change, social tranquillity, the pursuit of a less tense environment in prison. A non-recoverable prisoner can also be condemned to an inhuman punishment without this theoretically undermining the corrective model.

The inhumanity of the prison regime, on the other hand, undermines the human-centred prison model based on dignity. The re-educational emphasis, when it is not linked to the protection of human dignity, is potentially in conflict with it. The attention given to the function of punishment and all that it entails did not help to design a penitentiary system that is clear in its rights and duties, that connects them indissolubly without subordinating them to one another. Blindly relying on the idea of re-education means believing fideistically or hypocritically in impossible investigations of the deepest feelings of the person. For example, the Italian prison system of 1975 subordinates the granting of a wide range of benefits (which reduce the extent and intensity of the prison sentence) to 'participation of the prisoner in rehabilitation work'.[33] Participating or not participating will therefore not be indifferent to a prisoner. His future, even his being free or a prisoner, will depend on his participation in the rehabilitation work. All this introduces elements of interest in the asymmetrical relationship between the prisoner and the guard. Individual destinies are entrusted to a synallagmatic game that has little to do with the sphere of law. Shifting the spotlight from the re-educational utopia to human dignity and the rights deriving from it helps to read the aporia of prison, helps to reestablish the foundation of the prison system in a clearer way, imposing ethical limits that cannot be crossed and making it compatible with the rules of the welfare state and the rule of law.

Once this paradigm shift has been made, then re-education becomes capable of acquiring a high, secularised and de-ideologised social sense. With a fixed gaze on the horizon of human dignity, any intervention aimed at offering opportunities for social reintegration comes out of the game of hypocrisy and becomes an intervention to promote the rights of the person. In recent years it has happened that the Italian Constitutional Court has legally 'threatened' parliament by imposing measures to contain prison overcrowding.[34] The decision of the Constitutional Court states that 'the questions of constitutional legitimacy of Art 147 of the Italian Criminal Code, raised by the Venice and Milan Surveillance Courts, have been declared inadmissible, insofar as that

[33] Art 13 of Italian Prison Law no 354, 1975 never emended. At the end of 2018, a new reform of the prison system was approved, but Art 13, with its correctionalist function, remained unchanged. See P. Gonnella, *La riforma dell'ordinamento penitenziario* (Torino: Giappichelli, 2019), chapter 1.

[34] Corte Costituzionale 22 novembre 2013 no 279, available at www.cortecostituzionale.it.

provision does not include the situation of overcrowding in prisons among the cases of optional postponement of the execution of the sentence'. The court considered that it could not replace the legislator in identifying a judicial remedy to the problem of prison overcrowding but, at the same time, reserved the right, in the event of legislative inaction, to adopt in any subsequent proceedings, the necessary decisions aimed at ending the execution of the sentence in conditions contrary to a sense of humanity. The keyword of this wave of jurisprudence precisely is 'humanity', that is human dignity, in whose name judges are trying to obviate those policies of mass incarceration that have produced prisons in which life is degraded and treatment is inhuman.

It is not conceivable, however, that the mere normative, doctrinal and jurisprudential revival of human dignity can be sufficient to bring the system back to effective legality. A gap between legal proclamations and punitive practice exists and persists. There remains the strident paradox of an illegal punishment inflicted in the name of a broken legality, which must lead policymakers and academics to find solutions not only at the legal level but also at the cultural and operational level.

VI. How to Get Rid of the Need for Imprisonment

Prison must, therefore, be freed from the correctionalist ideology that legitimises a penalty that would otherwise be difficult to justify in contemporary societies. Everything in prison is based on the pedagogical ideology of treatment. 'Treatment' is a word that linguistically, before evoking people, habitually refers to textiles. It is customary to say, about fabrics, that they are treated whenever they are subjected to colouring or other interventions. Treatment has its etymological origin in the Latin *tràctum*, which then derives from the verb *trahere*. Tract indicates a physical space crossed by those who are in motion, but tract is also the supine form of the verb 'to acquire', which means, among other things, 'to take something' even if in a non-material sense. The word treaty also has the same root as treatment. A treaty is a pact, an agreement between two or more parties. Each treaty has its own rules, but it also has its own threats of sanctions. A treatise is also a formal work; it must be complete, self-contained, without gaps or flaws within itself. The entire structure of our prison law, as well as the prison laws of democracies and non-democratic regimes around the world, is based on the ideology of 'treatment'. The treatment of prisoners evokes each of the meanings mentioned above. It evokes what happens to the treatment of textiles, because prisoners, like fabrics, are subjected to a proposal for change aimed at a possible embellishment or some improvement. 'Individualised treatment' also evokes a path, one that goes from deviance to resocialisation. It also evokes deception because that path is based on the ideology of treatment and correctionalism, which has no objective

parameters of verification and that is often based on the same hypothesis as pentitism, or hypocrisy, with personal calculation of costs and benefits of each behaviour. Treatment evokes a treaty, because prison treatment is a sort of informal contract between prisoners and guards, in a synallagmatic game with sanctions and prizes. Equally, concerning the logical and philosophical completeness of a Scientific Treatise, the treatment, both the penitentiary and even more the re-educational one, despite its irremediable imperfection, aims at absoluteness.

Social and prison officers, but also penal law jurists and criminologists, starting from the contents of the 1975 Penitentiary Law and its subsequent evolutions and regressions, have mainly been concerned with classifying every aspect of prison life in the 'treatment' container. Therefore, treatment is not a good word. It presupposes an intervention of an exogenous nature. It makes one think about the need to put one's hands on the person, to want to change him or her to improve him or her. It does not seem too different with other words strongly marked from the ideological or religious point of view, as moral re-education or redemption. It does not bring to mind anything good, or at least anything authentic. It does not suggest a free choice. The treatment of a non-free person will always be based on blackmail, even if not explicitly proposed as such, even if occasionally not perceived as such.

Our penal and penitentiary legal framework is all about the myth of re-educational treatment. The treatment model is progressively overflowing, as was inevitable, towards a disciplinary model. Everything in prison is treatment and everything is discipline. Everything in the same prison law is reduced or elevated, depending on the case, to an element of treatment. It is no coincidence that this is the keyword of prison life. Outdoor exercise, permits awarded, work, education, even religion, human rights are all qualified by the same law as elements of treatment, whose philosophy permeates every part of life within prisons.

Relying on human dignity, on the other hand, means re-qualifying all prison life (and what produces it) in terms of human rights, reducing the power to punish, setting limits on those who believe the objective of punishment to be re-education and treatment, prohibiting torture. It is no coincidence that it is only by resorting to the notion of human dignity (and not discussing re-education) that judges have succeeded in setting limits to life imprisonment.[35]

The notion of human dignity is a healthy bath of realism. Panpenalism has shifted criminal law towards the construction of artificial crimes that live and die during an election round. Faced with forty thousand criminal laws in Italian

[35] See the Eur. Court H.R., *Viola* v *Italy*, Judgement of 13 June 2019, available at www.hudoc.echr.coe.it and Eur. Court H.R., *Vinter* v *the United Kingdom*, Judgment of 9 July 2013, available at www.hudoc.echr.coe.it. Italian Constitutional Court justified life imprisonment as compatible with the re-education function of the punishment.

law, there are no theoretical or practical possibilities for those who believe in the myth of re-education and treatment. Instead, there is always space for human dignity to undermine established powers, in all circumstances.

It is no coincidence that the clearest sociological position on the question of punishment was taken by Pope Francis. He based his observations and criticisms on the notion of human dignity, without relying on the more usual (for religious people) notion of re-education. If even the Pope no longer believes in the salvific power of punishment, it is unreasonable for jurists, philosophrs and lay criminologists to believe in it.[36]

[36] See P. Gonnella and M. Ruotolo eds, *Carceri e giustizia secondo papa Francesco* (Milano: Jaca Book, 2016). The book comments on the 2014 speech of Pope Francis addressed to the international association of scholars of criminal law.