Volume 5
Issue 2
Winter 2012

European Journal of Legal Studies

Young Scholar's Lab

Stefan Mayr
Putting a Leash on the Court of Justice? Preconceptions in National Methodology v *Effet Utile* as a Meta-Rule 3

Patrick R Goold
The Evolution of Normative Legal Scholarship: The Case of Copyright Discourse 22

Filippo Fontanelli
Necessity Killed the GATT - Art XX GATT and the Misleading Rhetoric about 'Weighing and Balancing' 39

General Articles

Alessandra Asteriti
Social dialogue, *Laval*-style 69

Kushtrim Istrefi
The Application of Article 103 of the United Nations Charter in the European Courts: The Quest for Regime Compatibility on Fundamental Rights 100

Davide Strazzari
The Scope and the Legal Limits of the 'Immigration Federalism': Some Comparative Remarks from the American, Belgian and the Italian Experiences 118

F E Guerra-Pujol
Modelling the Coase Theorem 179

www.ejls.eu

Website

www.ejls.eu

Submissions

The European Journal of Legal Studies invites submissions from professors, practitioners, and students. In particular, we welcome articles demonstrating a comparative, contextual and interdisciplinary approach to legal scholarship. Articles should be submitted in electronic format to subejls@gmail.com. If electronic submission is not practicable, please send a hard copy to:

The Editor-in-Chief
European Journal of Legal Studies
c/o EUI Department of Law
Villa Schifanoia, Via Boccaccio 121
50133 Florence – Italy

Footnotes should conform to The Oxford Standard for Citation of Legal Authorities, available for downloading from **www.law.ox.ac.uk/publications/oscola.php**.
The name of the author and contact details should appear only on a separate cover sheet to facilitate objective, anonymous evaluation. The Journal is committed to multilingualism and accepts submissions in any European language within the competence of the Board.

EJLS

European University Institute
DEPARTMENT OF LAW

European Journal of Legal Studies
Vol. 5 No. 2 (Winter 2012)

Table of Contents

Editorial

Tiago Andreotti 1

Young Scholar's Lab

Stefan Mayr
Putting a Leash on the Court of Justice?
Preconceptions in National Methodology v *Effet Utile*
as a Meta-Rule 3

Patrick R Goold
The Evolution of Normative Legal Scholarship:
The Case of Copyright Discourse 22

Filippo Fontanelli
Necessity Killed the GATT - Art XX GATT and the
Misleading Rhetoric about 'Weighing and Balancing' 39

General Articles

Alessandra Asteriti
Social dialogue, Laval-style 69

Kushtrim Istrefi
The Application of Article 103 of the United Nations
Charter in the European Courts: The Quest for Regime
Compatibility on Fundamental Rights 100

Davide Strazzari
The Scope and the Legal Limits of the 'Immigration
Federalism': Some Comparative Remarks from the
American, Belgian and the Italian Experiences 118

F. E. Guerra-Pujol
Modelling the Coase Theorem 179

EDITORIAL

Tiago Andreotti*

These are times of change for the European Journal of Legal Studies. As life presented them with exciting new opportunities, three of our executive members had to leave their management activities in the Journal. I would like to thank Benedict Wray, Bosko Tripkovic and Maciej Borowicz for the outstanding work done during the period they were in charge. I assume the position of Editor-in-Chief with pride for the accomplishments we achieved so far, but also knowing that there still is much to be done. Joining me in the executive board to help with the task are J. Alexis Galán Ávila, Cristina Blasi and Rebecca Schmidt.

On this issue the EJLS presents topics ranging from legal interpretation to law and economics, and it is divided in two sections. The first section is the outcome of the Young Scholars Lab, an event organized by Professors Miguel Maduro and JHH Weiler that took place at the European University Institute during the first week of June 2012. Devoted to the themes of legal scholarship, doctoral research, legal learning and legal teaching, the Lab also gave young scholars an opportunity to present their work in the workshop 'The Nouvelle Vaugue: A New Generation of Legal Scholarship Questioning Mainstream Assumptions'. Some of those contributions are published here. This is an important initiative that hopefully will be repeated in the following years, allowing for upcoming scholars to have their work scrutinized by their senior colleagues. The three contributions from the lab were written by Stefan Mayr, Patrick Goold and Filippo Fontanelli.

The first article is from Stefan Mayr, where he discusses the doctrine of *effet utile* as a meta-rule of interpretation and its development in the European legal system through the analysis of decisions from the Court of Justice, while at the same time raises important questions on the widespread belief and self-conception of law as a science.

Still within the theme of the role of legal scholarship and legal scholars in legal practice, Patrick Goold explains the change that copyright law scholarship took to overcome its decline. According to his account, by targeting the general public as its audience instead of legal practitioners, copyright law scholars can empower the public, which in turn holds lawmakers accountable for the legislation they enact. This is an interesting

* European University Institute (Italy). Any errors or omissions are entirely my own.

way to think about the position legal scholars have in society.

The last contribution in this section is in the Trade Law area. Filippo Fontanelli dissects the necessity test in the WTO system to answer the question of what regulatory margin of manoeuvre States preserve. By analysing WTO reports on the application of the Weighting and Balancing and the Least Trade Restrictive Measure formula, the author reaches the conclusion that necessity has, at least to a certain extent, killed the GATT.

Outside the Legal Scholars Lab framework we have four contributions. Alessandra Asteriti is concerned with the European Court's disregard of the function of collective bargaining in the Laval judgment and the consequences this may have on the carefully crafted Swedish system of social dialogue between management and labour.

Kushtrim Istrefi examines the approaches European Courts take when applying certain UN Security Council resolutions that may violate fundamental human rights in their own legal orders and suggests further exploration of interpretative techniques that may harmonize the conflicts arising out of art 103 of the UN Charter and national legal orders.

On a comparative study, Davide Strazzari discusses the immigration federalism in the context of the US, Belgium and Italy, proposing a cooperative approach to structure the territorial relations within the immigration field.

Finally, on his contribution F E Guerra-Pujol tries to remedy Coase's Theorem deficiency of being a verbal argument by using the game theory framework to present the Theorem as a formal game. This is an important article for those interested in the field of law and economics.

For our next issue in July we have a call for papers on Sovereignty, which will be made available on our website.

PUTTING A LEASH ON THE COURT OF JUSTICE? PRECONCEPTIONS IN NATIONAL METHODOLOGY V *EFFET UTILE* AS A META-RULE

Stefan Mayr*

The Court of Justice has time and again come under criticism for alleged methodological shortcomings and its dynamic approach towards interpretation. But who determines the boundaries between interpretation and admissible or inadmissible (ultra vires) creation of law? And where does the dividing line lie, given that the Member States have by and large accepted the most obvious creations of the Court of Justice (e.g. direct effect of directives, state liability etc.)? Any answer depends on our understanding of (a) the concept of interpretation as such and (b) the principle of effet utile – in a way the Court's interpretive leitmotif and as I will argue, a meta-rule of interpretation (and as such a small contribution to a genuine European methodology).

TABLE OF CONTENTS

I. INTRODUCTION 4
II. PRELIMINARY CONSIDERATIONS AND TERMINOLOGY 6
1. *The Role of the Court of Justice* 6
2. *The Notion and Concept of Effet Utile* 7
III. METHODOLOGICAL (RE-)CONCEPTION OF THE PRINCIPLE OF EFFET UTILE 8
1. *Key Elements of an Orthodox Perception of Interpretation* 9
a. A Function of Knowledge Focus on the Text 9
b. Legislator's Intention 9
c. Creation of Law 10
2. *Critical Assessment – Positivism Revisited?* 10
a. Kelsen '[T]he court is always a legislator' 11
b. Troper Théorie Réaliste de l'Interprétation 12
c. In a Nutshell 15

* Research Institute for Urban Management and Governance, Vienna University of Economics and Business Nordbergstrasse 15, 1090 Vienna, Austria; stefan.mayr@wu.ac.at.

3. *Consequences – Effet Utile as an Interpretive Meta-Rule* 15
IV. EFFET UTILE IN THE CASE LAW OF THE COURT OF JUSTICE 17
1. *Obvious Creations of Law* 17
2. *Borderline Cases* 18
a. Non-Discrimination on the Grounds of Age 18
b. Fundamental Freedoms and European Citizenship 19
V. CONCLUSION 20

I. INTRODUCTION

> *[W]hoever hath an absolute authority to interpret any written, or spoken Laws; it is He who is truly the Law-giver to all intents and purposes; and not the Person who first wrote or spoke them.*[1]

Time and again, the Court of Justice has been harshly criticised for its dynamic approach towards ensuring 'that in the interpretation and application of the Treaties the law is observed'.[2] [3] In German and Austrian doctrine this criticism mainly concerns alleged methodological shortcomings: Instead of *constructing*, the Court of Justice is said to be *creating* law.

Is such criticism tenable? Some theorists have argued, that it lies in the nature of language as such and of (general and abstract[4]) legal provisions in particular, that legal texts *always* require interpretation. Accordingly, their

[1] Benjamin Hoadly Lord Bishop of Bangor, *The Nature of the Kingdom or Church of Christ. A Sermon preach'd before the King 31 March 1717* (accessed via google.books) 12.
[2] Art 19 (1) Consolidated Version of the Treaty on European Union [2010] OJ C83/13 (TEU).
[3] Eg Roman Herzog and Lüder Gerken, 'Stoppt den Europäischen Gerichtshof' *FAZ* (Frankfurt, 9 September 2008) <http://www.cep.eu/presse/cep-in-den-medien/pressearchiv-2008/> accessed 1 September 2012. For further details: Marco Laudacher, 'Methodenlehre und Rechtsfindung im Gemeinschaftsrecht' [2010] UFSjournal 85, 90; Leslie Manthey and Christopher Unseld, 'Grundrechte vs. „effet utile" – Vom Umgang des EuGH mit seiner Doppelrolle als Fach- und Verfassungsgericht' [2011] ZEuS 323, 324; sceptical against such criticism Bernhard W Wegener, Art 19 EUV, in Christian Calliess and Matthias Ruffert (eds) *EUV/AEUV* (4th edn, CH Beck 2011) para 17.
[4] 'Generality' in this context is not to be mistaken for vagueness: Max Black, 'Vagueness. An Exercise in Logical Analysis' (1937) 4 Philosophy of Science 427, 432.

meaning – and hence, their normative content (ie the norm[5]) – would solely depend on their interpretation.[6] If so, the prevailing opinion in German and Austrian doctrine – assuming that the 'wording' or the 'potential meaning' of the legal text determines the boundary between interpretation (*Auslegung*) and creation of law (*Rechtsfortbildung*) – would be based on circular reasoning, limiting interpretation to the potential meaning of a legal text, which in itself depends on interpretation. However, this sceptical view, which will be referred to as 'realist approach', is highly contentious. Part of the (counter-)criticism is based on Ludwig Wittgenstein's later language philosophy. According to Wittgenstein, '[i]nterpretations by themselves do not determine meaning'[7]; to the contrary, any such attempt would result in the paradox of infinite regress.[8]

Hence, who can determine the boundaries between interpretation and admissible or inadmissible (ultra vires) creation of law? And where could the dividing line lie, given that the Member States have by and large accepted the most obvious creations of the Court of Justice (e.g. direct effect of directives[9], state liability[10] etc)?

Clearly, any assessment of the Court's case law depends on (and more or less openly expresses) an underlying methodological position. Finding an

---

[5] According to Hans Kelsen, *Pure Theory of Law* (University of California Press 1978) 5, '"Norm" is the meaning of an act by which a certain behavior is commanded, permitted, or authorized'.

[6] cf Hans Kelsen, *General Theory of Law and State* (Lawbook Exchange 2009) 146; but also Karl Larenz, *Methodenlehre der Rechtswissenschaft* (2nd edn, Springer 1992) 93: 'Es wäre ein Irrtum, anzunehmen, Rechtstexte bedürften nur dort der Auslegung, wo sie besonders ‚dunkel', ‚unklar' oder ‚widersprüchlich' erscheinen; vielmehr sind grundsätzlich alle Rechtstexte der Auslegung sowohl fähig wie bedürtig.' Jochen Anweiler, *Die Auslegungsmethoden des Gerichtshofs der Europäischen Gemeinschaften* (Peter Lang 1996) 26 f describes interpretation as a process of translation which is a very graphic image with a view to the large number of official languages in EU law.

[7] Ludwig Wittgenstein, *Philosophical Investigations* (3rd edn, Basil Blackwell 1986) para 198; cf Stefan Griller, 'Gibt es eine intersubjektiv überprüfbare Bedeutung von Normtexten?' in Stefan Griller, Karl Korinek and Michael Potacs (eds), *Grundfragen und aktuelle Probleme des öffentlichen Rechts: Festschrift für Heinz Peter Rill zum 60. Geburtstag* (LexisNexis 1995) 543 (550).

[8] Wittgenstein (n 7) para 201; cf James Tully, 'Wittgenstein and Political Philosophy' in Cressida J Heyes (ed), *The Grammar of Politics* (Cornell University Press 2003) 17, 38.

[9] Starting with Case 41/74 *Yvonne van Duyn v Home Office* [1974] ECR 1337.

[10] Starting with Joined Cases C-6/90 and 9/90 *Andrea Francovich and Danila Bonifaci and others v Italian Republic* [1991] ECR I-5357; for state liability in connection with the judiciary cf Case C-224/01 *Gerhard Köbler v Republik Österreich* [2003] ECR I-10239.

orthodox understanding of interpretation not tenable, a (modified) realist approach will bear significant consequences on our further understanding of the Court's interpretive practice in general and the *effet utile* principle – in a way the Court's interpretive leitmotif – in particular.

This paper will focus on sketching out a theoretical and methodological framework which will allow us to re-conceptualise *effet utile* as a meta-rule of interpretation. As such, *effet utile* can enhance the systematic assessment of EU law and be a fragment of a developing genuine European methodology.

II. PRELIMINARY CONSIDERATIONS AND TERMINOLOGY

1. *The Role of the Court of Justice*

According to art 19 (1) TEU, the Court of Justice[11] 'shall ensure that in the interpretation and application of the Treaties the law is observed'. Therefore, it not only decides upon the validity of EU law but also interprets it authoritatively and finally.[12] Moreover, the Court also regards it as its duty to grant comprehensive legal protection, even where it requires going beyond the black letter law of the Treaties.[13] Arguably, the broad concept of 'law' in art 19 TEU (in contrast to the narrower notion of 'the Treaties') and the dynamic approach towards integration reflected in the Preamble of the TFEU[14] justify such creation of law – at least to a certain extent.

In principle the Court of Justice and national (constitutional) courts use

[11] I will only consider the Court of Justice, not the General Court or any specialised courts and will therefore refer to it as the Court of Justice or simply the Court.

[12] cf Franz C Mayer, 'Multilevel Constitutional Jurisdiction' in Armin von Bogdandy and Jürgen Bast (eds), *Principles of European Constitutional Law* (2nd edn, Hart Publishing 2011) 399, 401 ff.

[13] cf eg Joined cases 7/56, 3/57 to 7/57 *Dineke Algera, Giacomo Cicconardi, Simone Couturaud, Ignazio Genuardi, Félicie Steichen v Common Assembly of the European Coal and Steel Community* [1957] ECR 39 para 55: 'The possibility of withdrawing such measures is a problem of administrative law, which is familiar in the case-law and learned writing of all the countries of the Community, but for the solution of which the Treaty does not contain any rules. Unless the Court is to deny justice it is therefore obliged to solve the problem by reference to the rules acknowledged by the legislation, the learned writing and the case-law of the member countries.'

[14] Consolidated Version of the Treaty on the Functioning of the European Union [2010] OJ C83/47. Just like the Preambles of the TEEC and the TEC it expresses the wish of the contracting parties 'to lay the foundations of an ever closer union'; cf Michael Potacs, 'Effet utile als Auslegungsgrundsatz' [2009] EuR 465, 474 ff.

similar methods of interpretation.[15] According to the Court's settled case law 'it is necessary to consider not only [a provision's] wording but also the context in which it occurs and the objects of the rules of which it is part'[16]. However, strikingly, the Court by stressing *effet utile* considerations often goes beyond an isolated purposive interpretation of a specific provision (eg of secondary law) and draws on the aims or purpose of the Treaties as such instead.[17] This openness inherent in the concept of *effet utile* evidently bears a considerable potential for conflict.[18]

2. *The Notion and Concept of Effet Utile*

Although the concept of *effet utile* plays a particularly prominent role in its case law, the Court of Justice has not 'invented' it. Quite the contrary – *effet utile* considerations can be traced back to Roman law (*ut res magis valeat quam pereat*) and have been explicitly codified in numerous modern legal orders.[19]

Also in international law, *effet utile* is regarded as 'one of the fundamental principles of interpretation of treaties'[20]. However, effectiveness considerations are here sometimes counterbalanced by another interpretive rule (*in dubio mitius*), prescribing the 'restrictive interpretation of treaty obligations in deference to the sovereignty of states'[21].

---

[15] Christian Calliess, 'Grundlagen, Grenzen und Perspektiven europäischen Richterrechts' (2005) 58 NJW 929, 929.

[16] Case C-223/98 *Adidas AG* [1999] ECR I-7099 para 23; Case 76/06 P *Britannia Alloys & Chemicals Ltd v Commmission of the European Communities* [2007] ECR I-4443 para 21; with regard to primary law: Case C-156/98 *Germany v Commission* [2000] ECR I-6882.

[17] Case 13/72 *Netherlands v Commission* [1973] ECR 27 paras 29 ff.

[18] Utility considerations aside, it has to be borne in mind that the Treaties also contain 'counterweights' such as the principles of conferral, subsidiarity or proportionality (art 5 TEU) which may also have a certain influence on the interpretation of EU law. cf Potacs (n 14) 476 ff.

[19] Eg art 1157 of the French Code Civil: 'Lorsqu'une clause est susceptible de deux sens, on doit plutôt l'entendre dans celui avec lequel elle peut avoir quelque effet, que dans le sens avec lequel elle n'en pourrait produire aucun.' For numerous further examples: Anna von Oettingen, *Effet utile und individuelle Rechte im Recht der Europäischen Union* (Nomos 2010) 41.

[20] Eg Territorial Dispute (Libyan Arab Jamahiriya/Chad) (Judgment) [1994] ICJ Rep 1994, 6, para 51: 'Toute autre lecture de ces textes serait contraire à l'un des principes fondamentaux d'interprétation des traités, constamment admis dans la jurisprudence internationale, celui de l'effet utile' (references omitted).

[21] At length Christophe J Larouer, 'In the Name of Sovereignty? The Battle over In Dubio Mitius Inside and Outside the Courts' [2009] Cornell Law School Inter-University Graduate Student Conference Papers <http://scholarship.law.cornell.edu/lps_clacp/22> accessed 1 September 2012.

In translations of early decisions of the Court of Justice, the French expression '*effet utile*' appeared in parentheses.[22] English translations now frequently use the terms 'effectiveness'[23] or 'full effectiveness'[24], more rarely also 'full force and effect'[25] or 'practical effect'[26]. In practice, however, these different terms refer to the same concept and lead to the same results.[27]

Whereas distinguishing between *effet utile* in a narrow and broad sense[28] or a weak and strong *effet utile* may be useful, it is also important to keep in mind that it is *one* concept, a continuum between these poles. We do not have to push the idea of e*ffet utile* far to exclude an absurd interpretation or one that would render certain guaranteed rights 'meaningless'[29]. Similarly the Court argues quite often that adopting a different interpretation 'would be tantamount to rendering [a certain right] ineffective and nugatory'[30]. But *effet utile* in an increasingly stronger (or wider) sense also prevents that a legal act or provision is 'deprived of a not insignificant aspect of its effectiveness' [31], or that its effectiveness is 'severely undermined'[32] or even (just) 'impaired'[33].

III. METHODOLOGICAL (RE-)CONCEPTION OF THE PRINCIPLE OF EFFET UTILE

[22] Eg Case 9/70 *Franz Grad v Finanzamt Traunstein* [1970] ECR 825 para 5.
[23] Eg Case 106/77 *Amministrazione delle Finanze dello Stato v Simmenthal SpA* [1978] ECR 629 para 20; Case 8/81 *Ursula Becker v Finanzamt Münster-Innenstadt* [1982] ECR 53 para 23.
[24] Eg *Francovich* (n 10) para 33; *Simmenthal II* (n 23) para 23; Case C-213/89 *The Queen v Secretary of State for Transport, ex parte: Factortame Ltd and others* [1990] ECR I-2433 para 21.
[25] Eg *Factortame* (n 24) para 20.
[26] Eg Case 70/72 *Commission v Germany* [1973] ECR 813 para 13.
[27] cf the *Simmenthal II* case (n 23) where the Court uses 'effectiveness' and 'full effectiveness' interchangeably.
[28] Potacs (n 14) 467.
[29] Case C-438/05 *Internatinoal Transport Workers' Federation and Finish Seamen's Union v Viking Line ABP and ÖU Viking Line Eesti* [2007] ECR I-10779 para 65. This would be a case where a weak *effet utile* 'suffices'.
[30] Case 157/86 *Mary Murphy and others v An Bord Telecom Eireann* [1988] ECR 673 para 10.
[31] Joined cases C-68/94 and C-30/95 *French Republic and Société commerciale des potasses et de l'azote (SCPA) and Entreprise minière et chimique (EMC) v Commission of the European Communities* [1998] ECR I-1375 para 171.
[32] Case C-450/06 *Varec SA v État belge*[2008] ECR I-581 para 39.
[33] *Simmenthal II* (n 23) para 20; *Factortame* (n 24) para 21.

While the flexibility of the concept of *effet utile* has become obvious from what has been said so far, its methodological nature remains somewhat unclear. The prevailing (German and Austrian) opinion regards *effet utile* as an aspect of teleological interpretation[34], a conception that will be challenged subsequently. Critically assessing an orthodox perception of interpretation in the light of a realist approach will bear significant consequences on the further understanding of the *effet utile* principle.

1. *Key Elements of an Orthodox Perception of Interpretation*

a. A Function of Knowledge – Focus on the Text

From a traditional (also referred to as orthodox) point of view, interpretation is a function of knowledge (not the will) and can be described as the text-based and text-bound finding of the correct meaning of a norm. Binding the exercise of state power to a published and accessible text is essential for the underlying understanding of the rule of law (*Rechtsstaat*). Paradigmatically Karl Larenz assumes, that interpretation means to neither add nor omit anything and just make the text speak for itself (by asking the *right* questions).[35] Consequently, proponents of this view argue, that the limits to interpretation can be determined by the meaning of the text itself.[36]

b. Legislator's Intention

Intended deference to the legislator often leads to the question 'What was the legislator's intention?' This question is certainly prone to misapprehension. However, it has been argued quite convincingly that this question should not be understood as aiming at ascertaining any 'psychological' intention. Much more, it should be understood as an attempt to determine, what can be imputed to the legislator according to the general rules and habits of communication and the general linguistic usage.[37]

---

[34] For an overview: Sibylle Seyr, *Der effet utile in der Rechtsprechung des EuGH* (Duncker & Humblot 2008) 103.

[35] Larenz (n 6) 255.

[36] ibid 210; Anweiler (n 6) 29; Franz Bydlinski, *Juristische Methodenlehre und Rechtsbegriff* (2nd edn, Springer 1991) 423; Wegener (n 3) para 12. According to Arthur Meier-Hayoz the wording serves a double purpose 'Er ist Ausgangspunkt für die richterliche Sinnermittlung und steckt zugleich die Grenzen seiner Auslegungstätigkeit ab.' *Der Richter als Gesetzgeber* 42, quote from Larenz (n 6) 210.

[37] Heinz Peter Rill, 'Juristische Methodenlehre und Rechtsbegriff' [1985] ZfV 461, 577; Michael Potacs, *Auslegung im öffentlichen Recht* (Nomos 1994) 24.

These rules can be distinguished into semantics (meaning of words) and pragmatics (context, purpose) which are normally closely interlinked. Still, a purely pragmatic transgression of the semantic meaning can be methodologically admissible (eg analogy, teleological reduction).[38] This shows that the weighting of these methods varies and therefore does not follow a strict rule. Much more, all these criteria for interpretation form a flexible system[39], inevitably open to multiple (value) judgments of the interpreting individual.

c. Creation of Law

According to common opinion, the (admissible) creation of law is regarded less exceptional in EU law due to its gaps and its integrative and dynamic function.[40] Nevertheless the Court of Justice is not to be seen as a 'substitute legislator'[41] and has time and again been harshly criticised in this regard. Even though German doctrine sporadically questions the usefulness of the attempt to draw a line between interpretation and creation of law[42], the practical effect of such scepticism exhausts itself in terminological sophistry. The result remains the same: One can and has to distinguish between admissible and inadmissible (ultra vires) creation of law. The latter comprises any interpretation contra legem and any acts in which the Court fails to maintain political neutrality and strictly refrains from policy making.[43]

2. *Critical Assessment – Positivism Revisited?*

The orthodox view of interpretation described above seems to be rooted in a rather formalistic understanding of law. Contrasting this view with Hans Kelsen's theory of law and Michel Troper's '*théorie réaliste de l'interprétation*'[44] produces some fruitful contradictions, which challenge a

---

[38] Potacs, Auslegung (n 37) 34 ff.
[39] 'Bewegliches System', ibid 40.
[40] Ulrich Everling, 'Richterliche Rechtsfortbildung in der Europäischen Gemeinschaft' [2000] JZ 217, 220.
[41] cf Anweiler (n 6) 54.
[42] Eg Larenz finds interpretation and creation of law 'nicht als wesensverschieden' (n 6 at 254), only the 'kreative Anteil des Subjekts' (ibid 255) varies. See also Seyr (n 34) 334 ff who coins creation of law as 'Weiterentwicklung' (development), but basically argues for the well known traditional limits to the CJEU's jurisdiction; ibid 337 ff.
[43] Calliess (n 15) 932.
[44] cf Michel Troper, 'Constitutional Justice and Democracy' (1995-96) 17 Cardozo L Rev 273, 282 ff; cf also the critical assessment of Otto Pfersmann, 'Contre le neo-réalisme juridique. Pour un débat sur l'interprétation' (2000) 52 Revue Française de Droit Constitutionnel 789 ff.

widespread preconception of legal 'interpretation'. Moreover, finding a considerable dose of realism in Kelsen's pure positivism also challenges the formalistic narrative which is still widely spread e.g. in traditional Austrian legal thinking.[45]

a. Kelsen – '[T]he court is always a legislator'

For Kelsen 'the court is always a legislator'.[46] It 'will always add something new'[47], no matter how detailed a general norm may be. Moreover, insofar as decisions of a court are binding upon future decisions in similar cases (precedents[48]), courts create general norms and are 'legislative organs in exactly the same sense as the organ which is called the legislator'[49].

However, this has nothing to do with any gaps in the legal order. Kelsen finds it logically impossible that the legal order has gaps. For him, the legislator (probably unconsciously) uses the 'fiction of "gaps of law"' to authorise e.g. courts to create new norms to avoid unjust or inequitable results in cases not previously considered (and therefore not covered by any general norm).[50] However, at the same time this fiction restricts the role of courts (as legislators) by creating an artificial pressure of justification.

To make things clearer: According to art 20 III of the German Basic Law, the judiciary is bound 'by law and justice' ('*Recht und Gesetz*'). The difference between interpretation and creation of law in the German context is clear: The consequences of interpretation are accepted as consequences of the legal provision itself, whereas any creation of law requires thorough justification.[51]

---

[45] cf more generally and most insightful John Gardner, 'Legal Positivism: 5 ½ Myths' (2001) 46 American Journal of Jurisprudence 199.
[46] Kelsen, General Theory (n 6) 146; Hans Kelsen, *Reine Rechtslehre* (2nd edn, Österreichische Staatsdruckerei 1992) 252 ff.
[47] Kelsen, General Theory (n 6) 146; Larenz (n 6) insinuates quite the same: „Dabei ist für den Vorgang der Auslegung daß der Ausleger nur den Text selbst zum Sprechen bringen will, ohne etwas hinzuzufügen oder wegzulassen. *Wir wissen freilich, daß sich der Ausleger dabei niemals nur rein passiv verhält* (...)." at 201 (emphasis added).
[48] In this connection it seems noteworthy that an analysis of the ECJ's case law for the year 1999 found that the most frequent method of interpretation was the reference to previous case law (ie precedents). Mariele Dederichs, 'Die Methodik des Gerichtshofes der Europäischen Gemeinschaften' [2004] EuR 345, 347.
[49] Kelsen, General Theory (n 6) 150.
[50] ibid 147.
[51] cf Larenz (n 6) 257.

The Court of Justice is endowed with the authentic and final interpretation of EU law and therefore a veritable lawmaker, creating general norms. Moreover, it is a court of last resort and therefore – according to Kelsen – its decision 'cannot be considered illegal, as long as it has to be considered a court decision at all'[52]. However, like the national courts, the Court of Justice depends on the willingness of the legal community to accept its decisions as binding and authoritative (and act accordingly). As a consequence, applying the yardstick of national methodology increases the (political) pressure of justification for the Court of Justice.

Interestingly, national constitutional courts, and in particular the German *Bundesverfassungsgericht*, still seem to enjoy greater trust in their respective legal communities than the Court of Justice.[53]

b. Troper – Théorie Réaliste de l'Interprétation

At a first glance, the quintessence of Troper's *théorie réaliste* appears quite provocative:

> According to [the realist theory of interpretation] the legal system empowers some authority to produce an interpretation of the text. This authority [...] is free to give any meaning to the text, which therefore has no meaning of its own prior to the interpretation.[54]

Troper's theory is descriptive; its cognitive interest lies in the process of interpretation by legal authorities ('authentic interpretation'), not the outcome or method used in any particular case. His main assumptions – which fundamentally challenge the cognitive (and normative) potential of legal science as such – are that interpretation is a function of the will (not knowledge) and that a legal text does not bear any objective meaning (due to the vagueness of language as a medium for communication).[55]

---

[52] Kelsen, General Theory (n 6) 155.

[53] It is noteworthy in this regard, that according to art 79 III GG not even the legislator could 'correct' certain decisions of the BVerfG; cf Manthey and Unseld (n 3) 324; Everling (n 40) 217 f.

[54] Michel Troper, 'Constitutional Interpretation' (2006) 39 Israel LRev 35, 36. However, Troper notes: 'Of course, the fact that this opportunity exists does not mean that the courts always use it unreservedly.' (n 44) 288.

[55] See also Michel Troper, 'Marshall, Kelsen, Barak and the constitutionalist fallacy' (2005) 3 Int'l J Const L (I.CON) 24, 34 f. A third assumption challenges the equation of meaning and intention. This question is dealt with above 3.1.2. It has to be mentioned however, that the variety of languages and cultures represented in the EU

Surprisingly, and even more remarkably, similar realist ideas have by and large been anticipated by Kelsen:

Traditional theory will have us believe, that the statute, applied to the concrete case, can always supply only one correct decision and that the positive-legal "correctness" of this decision is based on the statute itself. This theory describes the interpretive procedure as if it consisted merely in an intellectual act of clarifying or understanding; as if the law-applying organ had to use only his reason but not his will, and as if by a purely intellectual activity, among the various existing possibilities only one correct choice could be made in accordance with positive law.[56]

Kelsen insinuates interpretation as a function of the will as well as the absence of any objective meaning.[57] For him the indeterminacy of a legal act can be intentional or unintended. Unintended indeterminacy stems from the immanent vagueness of language as a medium.[58]

In order to guard against misunderstandings, a caveat is in place: The realist approach is largely based on the vagueness of language as a medium for (legal) communication. However, '[w]e cannot know that a word is vague, unless we know something about its use.'[59] According to Wittgenstein, '[f]or a *large* class of cases – though not for all – in which we employ the word "meaning" it can be defined thus: the meaning of a word is its use in the language.'[60] However, this use (or meaning) is not equivalent with the orthodox idea of a correct and objectively cognisable meaning of a legal text[61], against which the realist approach outlined above, argues. It is important to differentiate the basic (yet truly fundamental) question of how words can have meanings from the question to what extent such meanings prescribe the outcomes of an authentic interpretation. Arguably legal interpretation presupposes a deep

---

legislation considerably complicates our finding of rules and habits of communication and the general linguistic usage.

[56] Kelsen, Pure Theory (n 5) 351.

[57] For the opposite opinion prominently Larenz (n 6): 'Gegenstand der Auslegung ist der Gesetzestext als "Träger" des in ihm niedergelegten Sinnes, um dessen *Verständnis* es in der Auslegung geht.' 201 (emphasis added).

[58] Kelsen, Pure Theory (n 5) 350.

[59] Jeremy Waldron, 'Vagueness in Law and Language: Some Philosophical Issues' (1994) 82 California LR 509, 511.

[60] Wittgenstein (n 7) para 43.

[61] Griller (n 7) 560.

understanding of the use of the legal language.[62] 'It is only in normal cases that the use of a word is clearly prescribed'[63]; 'interpretation begins when our conventional self-understandings break down and we do not know how to go on.'[64] Arguably courts of last instance which decide complex cases and apply indeterminate ('contestable'[65]) normative standards (like *effet utile*, proportionality etc.) are thus – as the realist approach suggests – relatively free in giving *specific* meaning to legal texts.

Frequently there will exist a number of 'arguable norm-hypotheses'[66]. In the absence of one correct meaning, there can also be no right decision[67], but only a decision taken by an authority granted jurisdiction by the legal order.[68] Consistently, the Court of Justice rejects any limitation to its interpretive competence:

It must also be borne in mind, even where the different language versions are entirely in accord with one another, that Community law uses terminology which is peculiar to it. Furthermore, it must be emphasized that legal concepts do not necessarily have the same meaning in Community law and in the law of the various Member States.[69]

---

[62] Tully (n 8) 39 'Understanding grounds interpretation'; on the issue of infinite regress: According to Wittgenstein (n 7), 'any interpretation still hangs in the air along with what it interprets and cannot give it any support. Interpretations by themselves do not determine meaning' (para 198), much more there is always another interpretation standing behind it, which leads to the paradox of infinite regress (para 201). It should therefore be possible to disrupt the chain of infinite regress by substituting one expression of the rule for another (ie interpretation according to Wittgenstein, para 201), reaching a point where we can *imagine* a doubt, without actually doubting the meaning (cf para 84).

[63] Wittgenstein (n 7) para 142.

[64] Tully (n 8) 38 f.

[65] Waldron (n 59) 526: 'different users disagree about the detailed contents of that normative standard'. In fact Waldron argues that these contestable terms, by inviting us to make value judgments do not 'undermine the determinacy of their meanings. On the contrary, it is part of the meaning of these words to indicate that a value judgment is required ...' 527. Contestable terms thereby become 'a verbal arena in which we fight out our disagreements ...' 530.

[66] Adamovich and others (eds), *Österreichisches Staatsrecht*, vol 1 (2nd edn, Springer 2011) para 03.009.

[67] Again Larenz (n 6): 'Die Absicht, nur das auszusprechen, was der „richtig verstandene" Text von sich aus be-sagt, macht die typische Haltung des Interpreten aus.' (255).

[68] In this sense 'Creation of law is always application of law' and vice versa, Kelsen, General Theory (n 6) 133, 149.

[69] Case 283/81 *Srl CILFIT and Lanificio di Gavardo SpA v Ministry of Health* [1982] ECR 3415 para 19.

Again, this corresponds with Kelsen's understanding of 'authentic [ie law-creating] interpretation'[70] which is strictly limited to law-applying organs.[71] Legal science is much more limited: 'Jurisprudential interpretation can do no more than exhibit all possible meanings of a legal norm. Jurisprudence as cognition of law cannot decide between the possibilities exhibited by it but must leave the decision to the legal organ who, according to the legal order, is authorized to apply the law.'[72]

c. In a Nutshell

Legal science (which is no legal authority) lacks the criteria to evaluate whether an authentic interpretation is right or wrong, whether e.g. the Court of Justice has taken 'the right decision'. This is merely a question of legal policy, not science or theory.[73]

Authentic interpretation is then less of a 'theoretical-cognitive' process but much more the exercise of 'practical power', factually limited (to some extent!) by the acceptance within the legal community, therefore requiring sound reasoning and the adherence to certain argumentative standards.[74] Its validity depends on the authority of the organ, not the content of this result. Once a decision has become final (*res iudicata*), the validity of the created norm can no longer be rescinded. Therefore Kelsen notes: 'It is well known that much new law is created by way of such authentic interpretation, especially by courts of last resort.'[75]
However, there is a second side to it: whereas (almost) any interpretation by a court of last instance may be valid[76], its effectiveness depends on the acceptance within the legal community. Hence there is a certain pressure to actually stick to general rules and habits of communication and the general linguistic usage.

3. *Consequences – Effet Utile as an Interpretive Meta-Rule*

If we accept the above established conception of interpretation, ie that there is virtually no grey area between interpretation and (inadmissible)

[70] Kelsen, Pure Theory (n 5) 354.
[71] ibid 354 ff.
[72] ibid 355. cf Adamovich and others (n 66) para 03.010.
[73] Kelsen, Pure Theory (n 5) 352 eg reminds us that there exists no legal criterion to decide whether to apply an *argumentum e contrario* or an analogy.
[74] At length Ralph Christensen and Markus Böhme, 'Europas Auslegungsgrenzen' [2009] Rechtstheorie 285, 294 ff.
[75] Kelsen, Pure Theory (n 5) 355.
[76] Apart from ultra vires or non-existent acts.

creation of law, the latter appears as a catchword, aiming at constraining the Court of Justice substantially – but in a methodological (and seemingly objective) *disguise*. If the threat is serious enough, the Court may even change its interpretation – however for political, not legal reasons.

Similarly, *effet utile* might serve as a political slogan, aiming at convincing the legal (or much more political) community of the Court's rationality.

However, if we can derive interpretive meta-rules from the Court's argumentative patterns we can assess the *process* against the standards of such meta-rules and the Court's consistency (or arbitrary deviance) in their application.

Arguably *effet utile* is one of these meta-rules. It is not a rule of general linguistic usage and neither part of semantics nor pragmatics but logically comes into play, once potential meanings (or norm-hypotheses) have been established. It serves as a guideline without being an interpretive method itself.[77] As mentioned above it is a flexible concept, varying gradually but not qualitatively.[78] An absurd interpretation undermines the utility of a norm fundamentally. Hence preserving the validity of a provision can be understood as a very basic expression of furthering its utility.

Effet utile functions as a rule of choice and therefore logically presupposes a variety of arguable norm-hypotheses.[79] It cannot be applied in the absence of any doubt or alternatives. Even the avoidance of an absurd interpretation presupposes at least one arguable alternative (not absurd) meaning. Exceptionally this may not be the case: 'The law simply prescribes something nonsensical. Since laws are man-made, this is not impossible.'[80]

In a situation where the Court has to choose from alternative arguable norm-hypotheses, *effet utile* emphasises or even prioritises the teleological aspects. Among alternative meanings, it favours those furthering the effectiveness of EU law, putting a twofold emphasis on teleological aspects (with a view to the provision but also EU law in its entirety). It functions as an exclusionary rule of choice between tentative interpretive results.

This may sound rather trivial but the added value of this conception lies in an enhanced rationalisation:

---

[77] Marcus Mosiek, *Effet utile und Rechtsgemeinschaft* (LIT 2003) 7.
[78] Differently von Oettingen (n 19) 34 f.
[79] cf von Oettingen (n 19) 91.
[80] Kelsen, Pure Theory (n 5) 250.

According to the prevailing opinion a result that – semantics and pragmatics considered – cannot convincingly be found as consistent with the intention of the legislator, amounts to creation of law.[81] What may be considered 'convincing' has to be decided on a case by case basis with regard to the result of this interpretive act, ie the meaning, hence the norm itself. However, according to the view advanced here, legal science – lacking any legal authority – can develop norm hypotheses but not *decide* which result may be regarded as convincing.

What legal science can offer instead is an evaluation of the process of interpretation. Assuming that the traditional methods of interpretation form a flexible system, their relative weight in a specific case is governed by meta-rules of interpretation.[82] Hence, the admissibility of a certain interpretation depends on the plausibility of this weighing process. According to the opinion put forward here, any result that is arguable within a given set of meta-rules must be considered 'convincing'. Contrariwise, an inadmissible creation of law is one that violates (ie transgresses) these meta-rules *in toto*.

IV. EFFET UTILE IN THE CASE LAW OF THE COURT OF JUSTICE

1. *Obvious Creations of Law*

Due to the vast number of judgments referring to *effet utile*, the following section can only present a very small selection of relevant cases. It has been mentioned in the beginning that the Member States have by and large accepted the most obvious creations of law, many of them can be seen as cornerstones of the EU legal order.[83] In accepting these decisions, the Member states have accepted arguments based on weak and strong *effet utile* considerations alike. However, this shows that the Court's own effectiveness strongly depends on the acceptance of its reasoning in the legal community and some authors argue that only this acceptance can

[81] Potacs, Auslegung (n 37) 41. Either the result cannot be established at all or *better* reasons indicate a different result: ibid 277.

[82] Another meta-rule was formulated by the Court of Justice in the CILFIT decision (n 69), namely 'that legal concepts do not necessarily have the same meaning in Community law and in the law of the various Member States' (para 19).

[83] Rather obvious creations of law can also be found in other fields, concerning eg questions of residence and access to social benefits and have been criticised particularly harshly; cf Case C-184/99 *Rudy Grzelczyk v Centre public d'aide sociale d'Ottignies-Louvain-la-Neuve* [2001] ECR I-6193 and the criticism of Calliess (n 15) 932 and below.

prove the Court right or wrong.[84]

The direct effect of directives[85] serves as a prime example: The Court found it 'incompatible with the binding effect attributed to a directive by [art 288 TFEU] to exclude, in principle, the possibility that the obligation which it imposes may be invoked by those concerned. In particular, where the Community authorities have, by directive, imposed on Member States the obligation to pursue a particular course of conduct, the *useful effect* of such an act would be weakened if individuals were prevented from relying on it before their national courts and if the latter were prevented from taking it into consideration as an element of Community law.'[86] Once the Court's argument had been accepted it dropped the reference to *effet utile*[87] and applied direct effect without reiterating a detailed reasoning.[88]

Subsequently some of the more contentious decisions will be discussed. These borderline cases show argumentative shortcomings but may also serve as departure points for future considerations on refining *effet utile* as an interpretive meta-rule.

2. *Borderline Cases*

a. Non-Discrimination on the Grounds of Age

At the time the Court's *Mangold* decision[89] was among the particularly sharply criticised decisions.[90] The Grand Chamber held that the principle of non-discrimination on grounds of age constitutes a (directly effective) general principle of EU law. The difficulty with its reasoning was that the Treaty of Amsterdam had actually introduced the prohibition of discrimination on the grounds of age (Art 13 TEC), but only as a provision authorising the enactment of secondary law. The critics argued that such an authorisation implies that the Member States particularly did not want to introduce any directly effective prohibition of discrimination on grounds of age. However, recalling Kelsen's view, this is a typical situation where the Court *has to* take a policy decision. There exists no legal criterion to decide between an *argumentum e contrario* and an analogy.[91]

---

[84] cf Everling (n 40) 227.
[85] *van Duyn* (n 9).
[86] *van Duyn* (n 9) para 12 (emphasis added).
[87] Eg Case 152/84 *M. H. Marshall v Southampton and South-West Hampshire Area Health Authority (Teaching)* [1986] ECR 723 para 47.
[88] Eg *Becker* (n 23); cf von Oettingen (n 19) 133 ff.
[89] Case C-144/04 *Werner Mangold v Rüdiger Helm*[2005] ECR I-9981.
[90] Most prominently Herzog and Gerken (n 3).
[91] Kelsen, Pure Theory (n 5) 352.

b. Fundamental Freedoms and European Citizenship

The Court of Justice regularly relies on *effet utile* in cases involving Fundamental Freedoms, often broadening their scope or limiting exceptions[92] and other restrictions which obstruct the exercise of a Fundamental Freedom or simply make it less attractive[93]. Finally, it also constructs potential justifications for exceptions narrowly.[94]

Interestingly the Court takes a similar approach towards European Citizenship. The Grand Chamber decision in Ruiz Zambrano[95] is quite illustrative in this regard: The Court held that 'Article 20 TFEU precludes national measures which have the effect of depriving [minor, and dependent] citizens of the Union of the genuine enjoyment of the substance of the rights conferred by virtue of their status as citizens of the Union.'[96] The 'genuine enjoyment of the substance of the [citizen] rights' arguably means the full effectiveness, the *effet utile*, of these rights. The refusal to grant a right of residence but also a work permit to a third country national with such 'dependent minor children in the Member State where those children are nationals and reside'[97] is likely to produce exactly this effect:

It must be assumed that such a refusal would lead to a situation where those children, citizens of the Union, would have to leave the territory of the Union in order to accompany their parents. Similarly, if a work permit were not granted to such a person, he would risk not having sufficient resources to provide for himself and his family, which would also result in the children, citizens of the Union, having to leave the territory of the Union. In those circumstances, those citizens of the Union would, in fact, be unable to exercise the substance of the rights conferred on them by virtue of their status as citizens of the Union.[98]

---

[92] Eg Case 2/74 *Jean Reyners v Belgian State* [1974] ECR 631 concerning the right to establishment (and exceptions thereto) with regard to the profession of lawyers.
[93] cf Case 33/74 *Johannes Henricus Maria van Binsbergen v Bestuur van de Bedrijfsvereniging voor de Metaalnijverheid* [1974] ECR 1299, with regard to the freedom to provide services.
[94] cf Case C-355/98 *Commission v Belgium* [2000] ECR I-1221 paras 28 ff with regard to public policy and security.
[95] Case C-34/09 *GerardoRuiz Zambrano v Office nationale de l'emploi (ONEm)* [2009] (ECJ 8 March 2011).
[96] ibid para 42.
[97] ibid para 43.
[98] ibid para 44.

Already in Zhu and Chen[99] the Court argued that the refusal to grant a right of residence to the carer would deprive a child's right of residence of any useful effect, ie *effet utile*. Terminological differences aside, the Ruiz Zambrano judgment increases this *effet utile* in two regards: Firstly, third party nationals with dependent minor children who are Union Citizens gain unrestricted access to the European labour market. Secondly, the Court deduces this right from the status of the children as Union Citizens, regardless of any cross boarder reference and independently from the exercise of any Fundamental Freedom.[100]

Indeed, the Court first acknowledges that Directive 2004/38[101] explicitly limits its scope to Union Citizens 'who move to or reside in a Member State other than that of which they are a national'[102]. However the Court subsequently bases its decision directly on Art 20 TFEU, without mentioning, that these rights of Union Citizens 'shall be exercised in accordance with the conditions and limits defined by the Treaties and by the measures adopted thereunder [ie Directive 2004/38!]'[103].[104]

With a view to methodological stringency, the failure to mention this clause is a regrettable weakness. The Court's extremely short, if not erratic reasoning does not contribute to any further development of *effet utile* as a meta-rule. However, such contentious cases fuel the discourse and may therefore serve as points of departure for further considerations on *effet utile*.

V. Conclusion

The fiction of *the one* correct meaning of a legal provision serves at least two purposes. On the one hand, it is a basic pre-condition for (the ideal or illusion of) legal certainty. On the other hand it legitimises (political) value

---

[99] Case C-200/02 *Kunqian Catherine Zhu and Man Lavette Chen v Secretary of State for the Home Department* [2004] ECR I-9925 para 45.

[100] cf Krzysztof Chmielewski, 'Das Aufenthaltsrecht von drittstaatsangehörigen Familienangehörigen von Unionsbürgern nach dem Urteil des EuGH C-34/09' [2011] migralex 74, 77.

[101] Council Directive 2004/38/EC of 29 April 2004 on the right of citizens of the Union and their family members to move and reside freely within the territory of the Member States [2004] OJ L158/77.

[102] ibid art 3 (1).

[103] Art 20 (2) TFEU.

[104] Loic Azoulai, '"Euro-Bonds" The Ruiz Zambrano judgment or the Real Invention of EU Citizenship' (2011) 3 Perspectives on Federalism 31, 35.

judgments as scientific truth or legal necessity.[105] It is definitely more sophisticated – yet similarly flawed – to translate value judgments into methodological necessities.

A realistic approach towards interpretation fruitfully challenges widespread (more orthodox) beliefs concerning the process of interpretation but also the self-conception of law as a science (and legal scholars as scientists). What is more, a close reading of Kelsen – the proto-positivist – finds that he anticipates a lot of the realistic input and therefore also challenges the formalistic-positivistic narrative (e.g. in traditional Austrian legal thinking). Keeping in mind that language as a medium for law is vague but also based on conventions, I argued for a moderate realistic point of view. Even in the absence of an objective meaning a law-applying organ cannot arbitrarily ascribe *any* meaning to a norm. What is more the Court of Justice depends on the acceptance of its decisions within the legal community – not in terms of validity but in terms of effectiveness. Provocatively it could be argued that a decision which lacks such acceptance may be valid but *wrong*.

What does this mean for the (re-)conception of *effet utile*?
With a view to the Court of Justice as a court of last resort, the particularities of interpretation in the context of EU law and the sheer amount of case law referring to *effet utile* the analysis of argumentative patterns proves somewhat intricate. *Effet utile* could then merely be the leitmotif in the re-narration of a never-ending story – case by case. *Effet utile* can also be regarded as a political slogan aiming at convincing the legal (or much more political) community of the Court's rationality.

And finally, to the extent that the Court of Justice enters into a discourse and makes its understanding of *effet utile* and the aims and purposes of the Treaties transparent, we can derive meta-rules of interpretation from these argumentative patterns. Recalling that the Court ultimately depends on the acceptance of its decisions in the legal community, it has to be said that some of the Court's decisions, like *Ruiz Zambrano*, are disappointing in terms of argumentative style.

Whereas it appears untenable to criticise the Court of Justice on grounds of national methodological preconceptions, a re-conception of *effet utile* as a meta-rule of interpretation may enhance the systematic assessment of EU law and be a small fragment of a developing genuine European methodology.

[105] Kelsen, Pure Theory (n 5) 356.

THE EVOLUTION OF NORMATIVE LEGAL SCHOLARSHIP: THE CASE OF COPYRIGHT DISCOURSE

Patrick R Goold*

Legal scholarship's central function is to provide normative advice about the law. However, some academics have challenged the importance of such scholarship. Pierre Schlag argues that this function of legal scholarship is "unravelling" because judges and legislators do not listen to academic opinions. This unravelling would seem to be present in the field of copyright law where numerous instances suggest that normative legal scholarship is ignored. However, copyright scholarship has evolved to overcome this problem. Today the most influential copyright scholarship comes not in law reviews or similar traditional academic outlets, but through publicly oriented books and social media. Rather than aim normative advice to lawmakers, scholars give their advice to the public generally. The public then hold the lawmakers accountable for enacting bad laws. In this way, academics can retain their position as normative advice givers.

TABLE OF CONTENTS

I. INTRODUCTION 22
II. COPYRIGHT FUNDAMENTALS 24
III. ACADEMIC UNEASE 25
IV. ACADEMICS IGNORED 27
V. POPULARIZATION OF COPYRIGHT 31
VI. ACADEMIC RESPONSE: NEW TARGET AUDIENCES AND NEW DISTRIBUTION METHODS 33
VII. CONCLUSION 37

I. INTRODUCTION

Legal scholarship generally consists of normative statements about the way that government decisions should be made. These statements can be understood as

* Patrick R Goold, LL.B., LL.M., Esq., Doctoral Candidate (International Max Planck Research School for Competition and Innovation), Microsoft Research Fellow 2012-2013 (University of California, Berkeley, School of Law), patrick.goold@imprs-ci.ip.mpg.de. For comments on this paper I am grateful to Korbinian Reiter and participants of the EUI "Nouvelle Vague" Workshop. Any errors are purely the fault of the author.

prescriptions addressed to the relevant decision maker: most frequently a judge, but also a legislator or administrator.[1]
- Edward Rubin

By and large, neither judges nor any other bureaucratic decision makers are listening to academic advice that they are not already prepared to believe.[2]
- Pierre Schlag

Often the function of legal scholarship is to provide normative advice about the law. It is unlike the natural or social sciences, which aim to describe how the world is. Instead, legal scholars aim to show how the world ought to be. They endeavour to demonstrate what the optimal law on a given issue is. These suggestions are directed towards the legal decision makers (eg judges and legislators) in the hope that good law will be created. This is the rationale presented in the first opening quote from Edward Rubin.

But there is a problem with that role. That problem is summed up in the quote from Pierre Schlag. Arguably lawmakers do not consider the advice of legal scholars when it does not suit them. Schlag argues that this presents an 'unravelling' or 'decomposing' of the function of the legal scholar as normative advice giver. This belief is a serious challenge to legal scholars. If normative legal advice is routinely ignored, then does it have any justifiable place in modern legal discourse? Alternatively, is normative legal scholarship a relic from a past time, merely clinging to life within today's higher education system? This essay asks whether Schlag is correct and whether the normative advice-giving role of scholars is decomposing. It answers this question through the case study of copyright discourse.

Copyright law demonstrates features of Schlag's belief. One can argue that often lawmakers are not concerned with the writings of copyright scholars. However, on closer observation, one can see an evolution in normative copyright scholarship. As lawmakers increasingly ignore the views of copyright academics, scholars have changed the target audience for whom they write normative advice. Rather than aim normative advice directly to lawmakers, scholars now frequently write advice for the general public to read; they aim to persuade the public about what the law should accomplish. Once that is performed, the public can express their

[1]Edward L Rubin, 'On Beyond Truth: A Theory for Evaluating Legal Scholarship' (1982) 80 Cal L Rev 889-963, 900.
[2] Pierre Schlag, 'Pre-Figuration and Evaluation', (1992) 80 Cal L Rev 965-977, 972.

disapproval at undesirable copyright law through the democratic process. Scholars engage the public in this way by turning away from traditional forms of legal scholarship and instead distributing their ideas through social media and publicly oriented books. As a result, these scholars are shaping the way society and law makers view copyright despite an atmosphere that is arguably unresponsive to traditional academic opinion. In the arena of copyright therefore, Schlag's fears appear misguided; the function of the scholar is not dead. In this regard, copyright may be an atypical area of scholarship. Copyright perhaps concerns highly particularized issues of concern only to a specific group in society. Nevertheless, this case study may still provide an important message to those scholars working in other areas: if public engagement in legal discourse is possible, it can provide an efficacious tool for ensuring the ultimate creation of good law.

This paper shall begin by recapping the fundamental theory of copyright law. It will then go on in part 3 to highlight the academic views surrounding that theory. Part 4 will demonstrate how and why those views are often ignored. Part 5 shall however demonstrate the growing impact of the public will on copyright policy, and part 6 will show how copyright academics are successfully fuelling that public voice with their normative legal advice. The paper concludes by discussing some of the significance and limitations of this insight.

II. COPYRIGHT FUNDAMENTALS

Copyright law provides authors with the exclusive right to copy their literary and artistic works (eg books, music, film etc).[3] The reason for doing so, particularly in the Anglo-American tradition, is a matter of economics and is known as the Incentive Theory of Copyright.[4]

Artistic works have high fixed costs. This means that substantial resources (typically time and money) must be used to create the first copy of the work. But they also have low marginal costs. Once the first copy is in existence, it is cheap and simple to make further copies. This stems from the fact that these works are public goods.[5] They are non-rival, meaning

[3] See eg Berne Convention for Protecting Literary and Artistic Works (adopted 9 September 1886) art 9, (as amended on 28 September 1979) para 9.
[4] William Landes and Richard Posner, *The Economic Structure of Intellectual Property Law* (Harvard University Press 2003) 37-70; See also Reto M Hilty, 'Rationales for the Legal and Protection of Intangible Goods and Cultural Heritage' [2009] IIC 883-911.
[5] ibid 14.

that one person's consumption does not affect the ability of others to consume the good. They are also non-excludable, meaning that possession by one person does not prevent others from possessing the same good at the same time.

This leads to a particular market failure.[6] If there were no copyright, an author would spend significant resources creating the first copy of a work. For example, an author would spend time creating a book, in which time he still would have to expend money on food and shelter to maintain his existence. If he publishes the book, it could then be quickly and easily copied. Now there would be two versions in the market: the original and the copy. The copyist could then sell the book to a third party. Price competition between the two works would ensue. This is a competition that the copyist would be likely to win. The copyist would have no fixed costs to recover; he could therefore sell the work cheaper than the original author. The consumer would buy the copied version, and not the original author's, and as a result the original author would not recover his fixed cost investment. He therefore would lose money. If this scenario was routine, it would arguably be unlikely that he or any other author would invest their time creating new works in the future, despite the fact that doing so would be positive for social welfare.

Copyright aims to alleviate this market failure. By providing exclusivity in the market place, the copyright allows the author to raise prices above marginal cost without encouraging price competition. This supra-competitive pricing allows him to recover his costs. He (and other authors) therefore has an incentive to produce new works.

III. Academic Unease

However, despite this positive economic theory, copyright has numerous costs. Firstly, the supra-competitive pricing is poor for social welfare. The existence of consumers who are willing to pay a price above marginal cost, but not prepared to pay the supra-competitive price, means unfulfilled demand. This leads to deadweight loss and allocative inefficiency.[7] In addition, there are significant enforcement costs to the copyright.[8] Finally, copyright has potential non-economic costs. Particularly it has the potential to harm freedom of expression because it limits citizens' ability

[6] See Wendy Gordon, 'Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors' (1982) 82 Colum L Rev 1600-57.
[7] Landes and Posner (n 4) 71-84.
[8] ibid

to duplicate informational works.[9]

And it is also known that alternatives exist to ensure the production of new artistic works.[10] Even without the profit incentive, many authors would still create a certain number of new works. This is due to artists' general enjoyment from producing art and literature. Alternatively, the author could rely on his market lead-time. Copying successfully often requires a certain amount of time. In which time, the author has market exclusivity and can charge supra-competitive prices. During this time, no enforcement costs are incurred by the state. The government may also choose to actively encourage work production in other forms. Government bodies may provide subsidies to artists to produce works on commission. Or money could be awarded through prizes, allocated for works that are the most objectively impressive, or popular. Finally, there are also private contractual arrangements that could work. The author could contract with various actors prior to creating the work. They would provide him with money and he would use that to create the work. A modern equivalent of this is online crowd-sourcing, whereby Internet users pool money and allocate it to artists with original artistic ideas.

So far, however, this discussion is confined to theory. Theoretical benefits exist to copyright, but equally theoretical disadvantages exist, as do theoretical alternatives. There is little empirical evidence to suggest in reality whether the copyright is necessary.[11] And as a result of this lack of knowledge, scholars have often demonstrated uncertainty about whether copyright is indeed desirable. This tradition of academic uncertainty can be traced back at least to Arnold Plant in 1934[12] but has had much more modern iteration. Hurt and Schuman concluded that the 'traditional assumption that copyright enhances the general welfare is at least subject

---

[9] See generally Robert C Denicola, 'Copyright and Free Speech: Constitutional Limitations on the Protection of Expression' (1979) 67 Cal L Rev 283-316; Melville B Nimmer, 'Does Copyright Abridge the First Amendment Guarantees of Freedom of Speech and Press' (1969) 17 UCLA L Rev 1180-204; L.Ray Patterson, 'Free Speech, Copyright and Fair Use' (1987) 40 Vand L Rev 1-66.

[10] See eg Stephen Breyer, 'An Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs' (1970) 84 Harv L Rev 281-351; Lior Zemer, 'Rethinking Copyright Alternatives' (2006) 14 Int J Law Info Tech 137-45; Arnold Plant, 'The Economic Aspects of Copyright in Books' (1934) 1 Economica 167-95; Landes and Posner (n 4) 37-70.

[11] Ivan PL Ping, 'Copyright: A Plea for Empirical Research' (2006) 3 Review of Research on Copyright Issues 3-13.

[12] Plant (n 10).

to attack on theoretical grounds';[13] Steven Breyer (now US Supreme Court Justice Breyer) came to an 'ambivalent position on the question of whether copyright protection – considered as a whole – is justified';[14]and more recently Richard Watt concluded that 'some copyright piracy is highly likely to be socially efficient.'[15] The purpose of demonstrating this is not to suggest that all copyright academics wholly disagree with the necessity of copyright. Rather it is to show that a large number of academics are uncertain on this question, and because of the paucity of understanding, they do not wish to see copyright unjustifiably extended.

IV. Academics Ignored

Such academic perturbations are however often met with sanguine responses from lawmakers. Despite well documented theoretical deficits surrounding copyright, the law has expanded drastically throughout history. The first copyright statute, passed in Great Britain in 1709, allowed the authors of books the right to copy their work for a maximum of 28 years.[16] Today's copyright looks very different. US copyright, for example, lasts for the life span of the author plus an additional seventy years. [17] The right attaches to almost all forms of creative work [18] demonstrating a 'spark' or 'minimal degree' of originality.[19] And finally, the law no longer merely provides an exclusive right to make copies, but also confers exclusive rights to make adaptations, to perform the work publicly, to display the work publicly, and to distribute the work.[20]

Given the academic inability to prove copyright's necessity, why has it expanded so greatly? One answer is that private lobbying has successfully driven the legislative agenda. Historically, the impact of vested interests is familiar within copyright. The first copyright statute was fuelled by the

[13] Robert M Hurt and Robert M Schuman, 'The Economic Rationale of Copyright' (1966) 56 The American Economic Review 421-32, 432.
[14] Breyer (n 10) 322.
[15] Richard Watt, *Copyright and Economic Theory: Friends or Foes* (Edward Elgar 2000) 201.
[16] An Act for the Encouragement of Learning, by Vesting the Copies of Printed Books in the Authors or Purchasers of such Copies during the Times therein mentioned, 1709.
[17] 17 U.S.C. §302.
[18] 17 U.S.C. §102 (Copyright subsists in literary works (including computer programs), musical and accompanying works, dramatic and accompanying works, pantomimes, choreographic works, pictorial works, graphic works, sculptural works, motion pictures and other audio-visual works, sound recordings, and architectural works).
[19] *Feist Publications, Inc., v. Rural Telephone Service Co.*, 499 U.S. 340, 365 (1991).
[20] 17 U.S.C. §106.

bequests of the Stationers' Company, a collection of private booksellers.[21] And today this aspect of copyright is still well understood. The following passage from William Patry, former Copyright Counsel to the U.S. House of Representatives, is illustrative:

Copyright interest groups hold fundraisers for members of Congress, write campaign songs, invite members of Congress (and their staff) to private movie screenings or sold-out concerts, and draft legislation they expect Congress to pass without any changes. In the 104th Congress, they are drafting the committee reports and haggling among themselves about what needs to be in the report. In my experience, some copyright lawyers and lobbyists actually resent members of Congress and staff interfering with what they view as their legislation and their committee report. With the 104th Congress we have, I believe, reached a point where legislative history must be ignored because not even the hands of congressional staff have touched committee reports.[22]

Two modern anecdotes seem to add weight to the claim that academic views will often be overlooked when countered by the interests of lobbyists. Firstly, consider the Copyright Term Extension Act (CTEA).[23] This piece of US legislation was enacted in 1998. Prior to this date, the copyright term lasted the life of the author plus an additional fifty years. The CTEA extended that to life plus seventy years. The influence of lobbying was relatively clear. The Walt Disney Company's copyright over the lucrative Mickey Mouse character was due to expire in 2003. As Robert Merges relays, this company then went on a mission to prevent the character from falling into the public domain.[24] Merges is not alone in

---

[21] See eg Ronan Deazley, *On the Origin of the Right to Copy: Charting the Movement of Copyright Law in Eighteenth Century Britain (1665-1775)* (Hart Publishing 2004) 31-51; Adrian Johns, *Piracy: The Intellectual Property Wars From Gutenberg to Gates* (University of Chicago Press 2010) 17-40; Mark Rose, *Authors and Owners*: *The Invention of Copyright* (Harvard University Press 1993) 31-48.

[22] William F Patry, 'Copyright and the Legislative Process: A Personal Perspective' [1996] 14 Cardozo Arts & Ent 139, 141; See also Reto M Hilty, 'The Expansion of Copyright Law and its Social Justification' in Christopher Heath and Kung-Chung Liu (eds), *Copyright Law and the Information Society in Asia* (Hart Publishing, 2007) 1-31.

[23] Public Law 105 - 298 - An Act To Amend The Provisions Of Title 17, United States Code, With Respect To The Duration Of Copyright, And For Other Purposes.

[24] See eg Robert Merges, 'One Hundred Years of Solicitude: Intellectual Property Law 1900-2000' (2000) 88 Cal L Rev, 2187, 2235, n 218: 'Walt Disney was a company with a mission. With its copyright for Mickey Mouse up in 2003, Disney wanted to keep the character and the royalties for as long as it could. The company pushed for a law in the 105th Congress that would grant a 20-year extension on all copyrighted works. Congressional Quarterly reported that Disney CEO Michael Eisner made the

describing how Disney's concerns were at the root of the copyright extension. Other notable academics have made similar statements.[25] Even those inclined to support copyright expansion have noted the hand of private lobbying in this legislation.[26]

The law was passed and subsequently challenged on constitutionality grounds in the Supreme Court.[27] During the trial, seventeen famous economists, including 5 Nobel Prize winners, presented an amicus curiae brief to the court. In the brief the economists explained how the extension of copyright protection 'made little economic sense'.[28] They argued that any beneficial impact on author's incentive to create new works was insignificantly small.[29] At the same time they acknowledged that increasing the length of copyright protection has negative effects for economic welfare – due to longer monopoly pricing and enforcement costs. Despite these arguments, the court upheld the law. It dismissed the academic claims in a fashion that many have found unsatisfactory.[30] It appeared that

---

entertainment giant's position known at an informal June 9 meeting with Senate Majority Leader Trent Lott (R-Miss). A week later, Lott signed on as a co-sponsor to copyright extension legislation-and the very day Walt Disney's political action committee made a $1,000 contribution to Lott's campaign committee. On June 25, Disney made another donation-$20,000 in soft money to the National Republican Senatorial Committee.'; Center for Responsive Politics, 'No Lights, No Camera, Lots of Action: Behind the Scenes of Hollywood's Washington Agenda' (Oct. 11, 1998).

[25] See eg Chris Sprigman, 'The Mouse That Ate the Public Domain: Disney, the Copyright Term Extension Act, and Eldred v. Ashcroft', (*Findlaw,* 5 March 2002) <http://writ.news.findlaw.com/commentary/20020305_sprigman.html> accessed 30 May 2012.

[26] See eg Richard Epstein, 'The Dubious Constitutionality of the Copyright Term Extension Act' [2002] 36 Loy L.A. L Rev 123; Richard Posner, 'The Constitutionality of the Term Extension Act: Economics, Politics, Law and Judicial Technique in Eldred v. Ashcroft' (2003) 55 Sup Ct. Rev. 143.

[27] *Eldred v. Ashcroft*, 537 U.S. 186 (2003).

[28] Stan J Liebowtiz and Stephen Margolis, 'Seventeen Famous Economists Weigh in on Copyright: The Role of Theory, Empirics, and Network Effects' (2005) 18 Harv J L & Tech 435, 437.

[29] ibid 438.

[30] Epstein (n 26); Posner (n 26); Joy Rillera, 'Eldred v. Ashcroft: Challenging the Constitutionality of the Copyright Term Extension Act' (2003) 5 Vand J Ent L & Prac 23; Thomas R. Lee, 'Eldred v. Ashcroft and the (Hypothetical) Copyright Term Extension Act of 2020' (2003) 12 Tex Intel Prop LJ 1; Michael Jones, 'Eldred v. Ashcroft: The Constitutionality of the Copyright Term Extension Act' (2004) 19 Berkeley Tech LJ 85; Arlen W Langvard, 'Unwise of Unconstitutional? The Copyright Term Extension Act, the Eldred Decision, and the Freezing of Public Domain for Private Benefit' (2004) 5 Minn Intel Prop Rev 193; Sue Ann Mota, 'For Limited Times: The Supreme Court Finds the Copyright Term Extension Act as

lobbying beat the views of academia and that neither the court nor legislators were prepared to listen to a view that they did not already support.[31]

A similar process is occurring today with the proposed Anti-Counterfeiting Trade Agreement (ACTA).[32] This is a multilateral agreement between the USA and many other nations. The final text for which it has been drafted[33] and signed by most parties.[34] This law will further expand the protection for copyright by strengthening the enforcement power of the right holders.[35] Not only is it the strongest international law on the civil and criminal enforcement of copyright, it requires many nations to implement novel anti-piracy aids such as border measures (i.e. searching at ports for counterfeit or pirated goods) and technological protection measures (i.e. digital technologies designed to restrict copying). The proposed law was received badly by many legal scholars. Over 90 law professors gathered in June 2010 at the Washington College of Law to discuss the matter.[36] They concluded that the law "threatens numerous public interests"[37] including freedom of speech and privacy on the Internet. Later, more than 75 legal professors sent a letter to President Barrack Obama suggesting the law is harmful and should be

---

Constitutional in Eldred v. Ashcroft, But When Does it End?' (2005) BC Intel Prop & Tech F 110501.

[31] Equally, one US Supreme Court justice has notably denounced the value of legal scholarship. Chief Justice Roberts has recently made he following comment: "Pick up a copy of any law review that you see, and the first article is likely to be, you know, the influence of Immanuel Kant on evidentiary approaches in 18th Century Bulgaria, or something, which I'm sure was of great interest to the academic that wrote it, but isn't of much help to the bar.", see Richard Brust, 'The High Bench vs. The Ivory Tower' (*ABA Journal*, 1 Feb. 2012). <http://www.abajournal.com/magazine/article/the_high_bench_vs._the_ivory_tower/> accessed 14 July 2012

[32] The Anti-Counterfeiting Trade Agreement (ACTA), (Final Proposed Text 15 November 2010,) <http://www.ustr.gov/webfm_send/2379> accessed 30 May 2012.

[33] ibid.

[34] See generally, Michael Blakeney and Louise Blakeney, 'Stealth Legislation? Negotiating the Anti-Counterfeiting Trade Agreement (ACTA)' (2010) 16(4) Int TLR 87; Margot Kaminski, 'On the Origin and Potential Impact of the Anti-Counterfeiting Trade Agreement (ACTA)' (2009) 34 Yale J Int'l L 247.

[35] ACTA (n 32) ch 2.

[36] American University Washington College of Law, International Experts Find that Pending Anti-Counterfeiting Trade Agreement Threatens Public Interests (23 June 2010), <http://www.wcl.american.edu/pijip/go/acta-communique> accessed 30 May 2012.

[37] ibid.

substantially altered.[38] This has been echoed in the EU where 182 academics signed a letter to the EU commission criticizing the law in equally forceful terms[39]. However, these letters did not fundamentally alter the direction of the law.[40]

This ignorance of academic views is grist to the mill for Schlag. These anecdotes suggest a certain futility of academic normative advice. But, there is a countervailing force yet to discuss.

V. POPULARIZATION OF COPYRIGHT

Today, copyright is a matter of general public interest and debate. And a number of recent incidents suggest that the public voice is becoming increasingly important in dictating legal policy.

Continuing with the theme of ACTA, while academic opinion did not greatly influence the issue, public engagement did. During the early part of 2012 numerous wide-scale public demonstrations against the treaty occurred in Europe.[41] Since then, the Commission has asked the Court of Justice of the European Union to decide on whether ACTA is in line with fundamental human rights.[42] Neelie Kroes, the Digital Agenda Commissioner has suggested strongly that this was a response to the public protests. According to Kroes, the commission has 'recently seen how many thousands of people are willing to protest against rules which they see as constraining the openness and innovation of the Internet' and she

[38]Letter from 75 law professors to President Barack Obama (28 October 2010), available at <http://www.wcl.american.edu/pijip/go/blog-post/academic-sign-on-letter-to-obama-on-acta> accessed 30 May 2012.

[39] Opinion of European Academics on Anti-Counterfeiting Trade Agreement (3 December 2010), http://www.iri.uni-hannover.de/tl_files/pdf/ACTA_opinion_200111_2.pdf; For latest signatories see Institut für Rechtsinformatik, <http://www.iri.uni-hannover.de/subscriber.html> accessed 30 May 2012.

[40] The EU Commission responded but pushed ahead with the law, see EU Commission Comments on Opinion of European Academics on Anti-Counterfeiting Trade Agreement, (27 April 2011), <http://trade.ec.europa.eu/doclib/docs/2011/april/tradoc_147853.pdf> accessed 30 May 2012.

[41] See eg Dave Lee, 'ACTA Protests: Thousands take to the streets across Europe' *BBC News* (London, 11 February 2012) <http://www.bbc.co.uk/news/technology-16999497> accessed 30 May 2012.

[42] EU Commission Press Release, EU Commission Officially Referred ACTA to ECJ (Brussels, 11 May 2012) <http://trade.ec.europa.eu/doclib/press/index.cfm?id=799> accessed 30 May 2012.

acknowledged this as 'strong new political voice'.[43]

The same story could be made surrounding the Stop Online Piracy Act and the Protect IP Act (SOPA/PIPA).[44] These were two bills laid before the US Congress in early 2012. Like ACTA each aimed to increase the enforcement powers of IP holders. Despite initial momentum, the bills lost support after widespread dissatisfaction from the public as well as some of the world's most popular websites eg Wikipedia.org.[45]

These examples conform to a more general trend. It is far more common today to see the public engage in copyright issues. In the last decade a number of organizations have founded in order to facilitate this. Of primary importance is the Free Culture Movement. It refers to an ideological perspective advocating that copyright be less restrictive and allow the general public more freedom to use copyrighted works. This ideology translates itself into a number of real world activist groups. Students for Free Culture, for example, is an international organization, consisting of many different university chapters upholding the free culture ideals.[46] And the Free Culture Forum[47] is a coalition of various actors who produce white papers on copyright issues such as the 'Charter for Innovation, Creativity, and Access to Knowledge'.[48]

Beyond that there are also licencing organizations such as Creative Commons.[49] This is an international non-profit organization that aims to facilitate the licensing of copyrighted material. When a good is licensed under traditional copyright law, the copyright holder maintains all the rights over the work. The copyright holder is still the only person who can

---

[43] Neelie Kroes, 'The European Public on the Net, (Berlin, 4 May 2012) <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/12/326&format=HTML&aged=0&language=EN&guiLanguage=en> accessed 30 May, 2012; see also John Clancy, EU Trade Spokesman (note 42) 'The Court's opinion is vital to respond to the wide-ranging concerns voiced by people across Europe on whether ACTA harms our fundamental rights in any way.'.

[44] H.R. 3261, 112th Cong. (2012); S. 968, 112 Cong. (2012).

[45] See eg Jonathan Weisman, 'In Fight Over Piracy Bills, New Economy Rises Against Old' *New York Times* (New York, 18 January 2012) <http://www.nytimes.com/2012/01/19/technology/web-protests-piracy-bill-and-2-key-senators-change-course.html?pagewanted=all> accessed 30 May 2012.

[46] Students for Free Culture, <http://freeculture.org/> accessed 30 May 2012.

[47] The Free Culture Forum, <http://fcforum.net/> accessed 30 May 2012.

[48] The Free Culture Form, 'Charter for Innovation, Creativity, and Access to Knowledge' <http://fcforum.net/charter_extended> accessed 30 May 2012.

[49] Creative Commons, <http://creativecommons.org/> accessed 30 May 2012.

copy, adapt, perform or display, and distribute the work.[50] Creative Commons licensing is different. It allows the author to retain 'some' rights.[51] For example, the copyright holder may allow users to create adaptations to his work. Or he may allow users to copy the work freely for certain purposes. Which rights the author retains depends on which license he uses.[52]

A final point could also be made about copyright advocacy groups, such as the Electronic Frontier Foundation (EFF). The EFF is a non-profit organization that advocates the rights of users in the digital world. It describes itself as 'the first line of defense'[53] when user freedoms come under attack. In pursuing these goals, the EFF funds a number of court cases[54] and the production of whitepapers on copyright issues.[55] In doing so, it has had a number of successes in changing the direction of the law.[56]

VI. Academic Response: New Target Audiences and New Distribution Methods

There is therefore a public engagement in copyright issues. What is even more interesting is the relationship between this public audience and copyright academics. In a world where traditional academic opinion often falls on deaf ears, frequently copyright academics write directly for this public audience. And, although the nature of this relationship is undoubtedly complex, one can advance the hypothesis that the work of these academics is one causal factor in generating public discussion.

The clearest example is that of Lawrence Lessig. Lessig is a professor of law at Harvard. In addition, he is the founder of Creative Commons, a former board member of the EFF, and arguably the figurehead of the Free Culture Movement. And, particularly in relation to the latter movement, it is interesting to note how Lessig has helped to develop this public

[50] This is the case unless the rights are expressly transferred, e.g. under17 USCS Sect. 106A(e) in the US context.
[51] Creative Commons, 'About' <http://creativecommons.org/about> accessed 30 May 2012.
[52] Creative Commons, 'Licenses' <http://creativecommons.org/licenses/> (last visited 30 May 2012).
[53] Electronic Frontier Foundation, 'About' <http://www.eff.org/about> (last visited 30 May 2012).
[54] Electronic Frontier Foundation, 'Cases' <http://www.eff.org/cases> (last visited 30 May 2012).
[55] Electronic Frontier Foundation, 'Whitepapers' <http://www.eff.org/wp> (last visited 31 May 2012).
[56] Electronic Frontier Foundation, 'Victories' <http://www.eff.org/victories> (last visited 31 May 2012).

engagement.[57] As a legal academic and professor at Harvard, one would expect to see a long list of lengthy, footnote laden articles (perhaps fairly describable as esoteric and arcane) published in traditional legal journals and law reviews. These articles would make normative statements about the correct shape of the law. The target audience would be legislators and judges. This would be consistent with Rubin's view of legal scholarship. That is what one would expect but not what one will find. Although some such works still exist,[58] Lessig has conveyed his most influential legal thoughts by writing books designed for the general public to read.

Some of Lessig's most prominent works on copyright law are: *Code*,[59] *The Future of Ideas*,[60] *Free Culture*,[61] and *Remix*[62]. Most of his books are free for download under creative commons licenses as e-books. Alternatively, they can be found in paper back at most book retailers. The central message of all these books is that copyright law is too restrictive and has negative effects on the creation and spread of creative works in society.[63] And much of the Free Culture Movement is founded directly upon these ideas.[64] The movement employs the Lessig-coined phrase 'Free Culture' as its central theme and uses much of Lessig's terminology and arguments. In doing so, these publications have given shape to the entire copyright discourse in the digital age.

---

[57] Lawrence Lessig, Harvard Law Faculty Directory, <http://www.law.harvard.edu/faculty/directory/index.html?id=888&show=bibliography > accessed 31 May 2012

[58] See eg Lawrence Lessig, 'What Everybody Knows and What Too Few Accept' (2009) 123 Harvard LR 104; Lawrence Lessig, 'In Support of Network Neutrality' (2007) 3 ISJLP 185.

[59] Lawrence Lessig, *Code: And Other Laws of Cybersprace* (Basic Books, 2000).

[60] Lawrence Lessig, *The Future of Ideas: The Fate of the Commons in a Connected World* (Random House, 2003).

[61] Lawrence Lessig, *Free Culture: How Big Media Uses Techology and the Law to Lock Down Culture and Control Creativity* (Penguin Books, 2005).

[62] Lawrence Lessig, *Remix: Making Art and Commerce Thrive in the Hybrid Economy* (Penguin Book, 2008).

[63] See eg Lessig (n 61) 28: 'There has never been a time in history when more of our 'culture' was as 'owned' as it is now. And yet there has never been a time when the concentration of power to control the uses of culture has been as unquestionably accepted as it is now'.

[64] See eg Richard Poynder, 'Interview with Lawrence Lessig' <http://poynder.blogspot.com/2006/04/interview-with-lawrence-lessig.html> (describing Lessig as the "de facto leader" of the Free Culture Movement) accessed 30 May 2012; Siva Vaidhyanathan, 'The Anarchist in the Coffee House: A Brief Consideration of Local Culture, The Free Culture Movement, and Prospects for a Global Public Sphere' (2007) 70 L & Contemporary Problems 205.

Lessig's scholarly strategy has not stopped at writing books. He has adopted other innovative ways of distributing his advice. His use of television and film is one such example. Lessig appeared and discussed his ideas in popular television shows such as *The West Wing*,[65] *The Colbert Report*,[66] *The Daily Show*[67] and in popular documentaries such as *RiP: A Remix Manifesto*.[68] In addition Lessig employs a private blog, a twitter feed, and a wiki (a website that allows the creation and editing of any number of interlinked web pages via a web browser using some simple tools) to distribute his ideas. He is also a frequent blogger on various other sites, such as the influential news-blog *The Huffington Post.*[69]

When one looks at Lessig's work, one sees a legal scholar that has had a strong impact on how society views copyright policy. But rather than speak to lawmakers, who seem unlikely to listen, he has addressed his advice to the public generally. And Lessig is not alone in this process. While he is perhaps the clearest example, numerous other copyright scholars have also changed their target audience and distribution methods. In the footsteps of Lessig, well-established academics have with increasing frequency produced copyright literature for the general masses. This essay mentioned William Patry above. In addition to writing one of the leading copyright treatises, Patry has produced two popular books entitled *Moral Panics and the Copyright Wars*[70] and *How to Fix Copyright.*[71] In the former Patry discusses how copyright expansionists have resorted to metaphors that demonize copyright infringers just as is often the case with moral panics. And in the latter, Patry discusses the interplay between copyright law and technology. Neil Netanel, professor of law at UCLA Law School, published *Copyright's Paradox*.[72] This work details the complicated relationship between copyright law and free speech. Adrian Johns, professor of history at the University of Chicago has produced *Piracy: The Intellectual Property Wars from Gutenberg to Gates*.[73] This is an historical account of the term copyright 'piracy'. And, there are many more examples of these books.[74] It would take too much time to detail them all here.

[65] 'The West Wing: The Wake-Up Call' *NBC Television Broadcasts* (9 February 2005).
[66] 'The Colbert Report' *Comedy Central Television Broadcast* (8 January 2009).
[67] 'The Daily Show' *Comedy Central Television Broadcast* (13 December 2011).
[68] Brett Gaylor, 'RIP: Remix Manifesto!' (2008).
[69] *The Huffington Post*, 'Blog Entries by Lawrence Lessig,' <http://www.huffingtonpost.com/lawrence-lessig> accessed 30 May 2012.
[70] William F Patry, *Moral Panics and the Copyright Wars* (OUP 2009).
[71] William F Patry, *How to Fix Copyright* (OUP 2011).
[72] Neil Netanel, *Copyright's Paradox* (OUP 2010).
[73] Johns (n 21).
[74] See eg Siva Vaidhyanathan, *Copyrights And Copywrongs: The Rise Of Intellectual Property And How It Threatens Creativity* (New York University Press 2001); Joanna

Needless to say, these books are relatively cheap[75] and they are distributed to the public in the same manner that other public books are. They can be found online at Amazon.com or a local bookstore. Many are even downloadable as e-books to facilitate the new generation of technology savvy digital-book readers such as the Kindle. In addition, these professors also employ the use of social media. This often comes in the form of blogs, some of which are individually run[76] while others chose to contribute to collaborative blogs such as the *Huffington Post;*[77] many use Twitter as well.[78] By doing so, these scholars distribute their normative legal suggestions directly to the general public, rather than to judges and legislators; they then rely on the public to demand that good laws be created in the routine democratic fashion, as has happened in the ACTA and SOPA/PIPA controversies.

Demers, *Steal This Music: How Intellectual Property Law Affects Musical Creativity* (University of Georgia Press 2006); Jonathan Zittrain, *The Future Of The Internet: And How To Stop It* (Penguin Books 2008); Michele Boldrin, *Against Intellectual Monopoly* (CUP 2008); James Boyle, *The Public Domain: Enclosing The Commons Of The Mind* (Yale University Press 2010); John Palfry, *Born Digital: Understanding The First Generation Of Digital Natives* (Basic Books, 2010); Patricia Aufderheide and Peter Jaszi, *Reclaiming Fair Use: How To Put Balance Back In Copyright* (University of Chicago Press 2011); Paul K. Saint-Amore, *Modernism And Copyright* (OUP 2011); John Tehranian, *Infringement Nation: Copyright 2.0 And You* (OUP 2011); Madhavi Sunder, *From Goods to a Good Life: Intellectual Property and Global Justice* (Yale University Press, 2012).

[75] See eg Boyle (n 74); Amazon.com Price §11.23 <http://www.amazon.com/The-Public-Domain-Enclosing-Commons/dp/0300158343/ref=tmm_pap_title_0> accessed 30 May 2012; Patry (n 70); Amazon. Com Price §11.98 <http://www.amazon.com/Moral-Panics-Copyright-William-Patry/dp/B0062GK70O/ref=sr_1_1?s=books&ie=UTF8&qid=1338648930&sr=1-1> accessed 30 May 2012.

[76] See eg James Boyle, 'The Public Domain: Enclosing the Commons of the Mind' <http://www.thepublicdomain.org/blog/> accessed 30 May 2012; William Patry, 'The Patry Copyright Blog' <http://williampatry.blogspot.com/> accessed 30 May 2012; Jonathan Zittrain, 'The Future of the Internet and How to Stop It' <http://futureoftheinternet.org/blog> accessed 31 May 2012.

[77] See eg *The Huffington Post*, 'Blog Entries by Pamela Samuelson' <http://www.huffingtonpost.com/pamela-samuelson> accessed 30 May, 2012; *The Huffington Post*, 'Blog Entries by Edward Lee, The Huffington Post' http://www.huffingtonpost.com/edward-lee accessed 30 May 2012; *The Huffington Post*, 'Blog Entries of Yochai Benkler', <http://www.huffingtonpost.com/yochai-benkler> accessed 30 May 2012.

[78] See e.g. Michael Boldrin, MichaleBoldrin@michaeleboldrin <http://twitter.com/#!/micheleboldrin> accessed May 30 2012; James Boyle, JamesBoyle@thepublicdomain <http://twitter.com/#!/thepublicdomain> accessed 30 May 2012; Jonathan Zittraine, Jonathan@Zittrain, http://twitter.com/#!/zittrain accessed 30 May 2012.

The author of this essay has in the past had the opportunity to speak with some of these scholars and ask them their opinions on this idea. Lawrence Lessig particularly agreed that by writing books designed for the public he could maximize his impact on society purely by reaching more people.[79] Whereas law review articles would be read by a small number of people made up of mostly other law professors, as well as some judges and legislators, books such as *Free Culture* and *Remix* are read by a far greater number of people. This maximizes the dispersion and the impact of the normative advice. The same idea was endorsed by Michael Geist who, when asked about the impact of his traditional academic articles compared to his well known blog on ACTA,[80] felt that the latter had a much greater impact on how law would develop.[81]

VII. Conclusion

This paper has discussed solely normative legal scholarship. It is true that there are perhaps other aspects of legal scholarship, such as doctrinal study, which may not be as well suited to popularization. Nevertheless, the example of copyright is significant for any scholar wishing to retain an impact on the fundamental policy objects of the law. On the most important questions in copyright, private interests often override traditional academic opinion. This supports Schlag's 'unraveling' theory. But in response to this, copyright scholarship has evolved. Scholars frequently choose to address their normative advice, not to lawmakers, but to the general public. This public, as seen in the examples of the ACTA and SOPA/PIPA, is capable of influencing law making where arguably traditional academics are not.

This case study of copyright provides a message to academics working in other areas of law. The message is indeed tentative currently and in need of thorough empirical study. Nevertheless, some anecdotal evidence suggests that when academic opinion appears to be routinely ignored, then trying to engage the public is a strategic move towards ensuring the creation of good laws. Therefore, let academics discover knowledge about what is good law, give that knowledge to the public and allow people to make the normative decisions that lawmakers should follow.

---

[79] Telephone Interview with Lawrence Lessig, Professor of Law, (Harvard Law School, 8 July 2011).

[80] Michael Geist, 'Michael Geist Blog' <http://www.michaelgeist.ca/> accessed 31 May 2012.

[81] Interview with Michael Geist, Professor of Law at University of Ottawa, (Berkeley, California, USA, 21 April 2012).

Will all legal subjects equally benefit from such popularization in the same way that copyright has? It is difficult to say from this early vantage point. But one could easily envision the polemic issues found in constitutional law, public international law, and criminal law (amongst others) equally engaging the public's imagination. And arguably law professors working in these areas are granted job security via tenure in order to encourage academic risk taking. Notably, Pierre Schlag has elsewhere called for tenured legal professors to take more risks and to reinvigorate legal scholarship.[82] Talking to the public may be part of that new future. In doing so, scholars will retain their positions as normative advice givers and this aspect of scholarship will remain justifiable.

[82] Pierre Schlag, 'Spam Jurisprudence, Air Law, and the Rank Anxiety of Nothing Happening (A Report on the State of the Art)' (2009) 97 Geo L J 803-835.

NECESSITY KILLED THE GATT
ART XX GATT AND THE MISLEADING RHETORIC ABOUT 'WEIGHING AND BALANCING'

Filippo Fontanelli*

Art XX GATT, listing the policy grounds available to WTO Members that wish to deviate from their GATT obligations, makes some of them conditional on a requirement of necessity in relation to the pursued interest. In their reports, Panels and the AB have developed the analysis of this element in two separate but interlaced tests: one whereby they allegedly perform an exercise of 'weighing and balancing' of the interests involved (a value-judgment), the other ascertaining the trade-restrictiveness of the measures challenged (an optimization analysis). It is submitted that an appraisal of the case-law demonstrates that this distinction is artificial, and most importantly, that no real balancing is ever performed - or in any event, relied on - to determine the outcome of a dispute (Claim 1). However, a diffuse trend of 'strict proportionality' is discernible in the case-law, not so much within the 'weigh and balance' analysis, but within the trade-restrictiveness test. The latter, therefore, is arguably less value-neutral than the quasi-judicial bodies would claim it to be, and then WTO Members tend to understand, when construing the necessity requirement (Claim 2).

TABLE OF CONTENTS

I. INTRODUCTION 40
II. THE NECESSITY TEST IN ART XX GATT 41
III. NECESSITY IN THE GATT-DAYS: ENTER THE LTRM 44
IV. THE ABSOLUTE FREEDOM TO SET THE LEVEL OF PROTECTION 47
V. THE ADDITIONAL CHALLENGE OF ART XX(D) GATT 49
VI. ENTER THE WAB 51
VII. PAYING LIP-SERVICE TO THE WAB: *GAMBLING* AND *CIGARETTES* 55
VIII. THE WILDCARDS: COMPLEMENTARITY AND QUANTITATIVE CONTRIBUTION 59
IX. THE PRACTICE OF EXCEPTION-SHOPPING, AND THE REVIVAL OF THE WAB 61

* f.fontanelli@surrey.ac.uk Lecturer in Law, University of Surrey. Thanks to Renginee Pillay, Giuseppe Martinico and Theodore Konstadinides for their helpful comments, and to the organizers of, and participants to, the workshop 'The Nouvelle Vaugue: A New Generation of Legal Scholarship Questioning Mainstream Assumptions', Fiesole, 8 June 2012. Special thanks to Miguel P Maduro.

X. MAKING SENSE OF THE *TYRES* GUIDELINES ON CONTRIBUTION 63
XI. CONCLUSION 65

I. INTRODUCTION

Art XX of the General Agreement on Tariffs and Trade (GATT) entitles Members of the World Trade Organization (WTO) to adopt WTO-inconsistent measures, provided that they fall into one of the categories listed therein, each related to a different policy objective, and they are applied in a non-discriminatory way. In particular, exceptions designed to promote public morals, human (and animal and plant) health, and compliance with GATT-consistent national norms must be 'necessary' to achieve the sought objective. This article is concerned with the interpretation and application of this necessity factor by WTO quasi-judicial bodies (Panels and Appellate Body).

The necessity test is but one of the typical devices used to govern the interplay of overlapping regulatory regimes in a situation of legal pluralism. Not unlike other doctrines, such as subsidiarity, margin of appreciation, comity, *Solange* etc, it aims to limit the scope and application of a regime that would normally enjoy priority over norms of other concurring regimes. The purpose of the said doctrines is to provide the 'yielding' norms with enough margin to operate, if certain conditions are/are not met, or if some subject-matters are/are not touched upon.[1]

An analysis of the interpretation and application of the necessity test, therefore, provides an optimal vantage point to take stock of the WTO's impact on the regulatory autonomy of Member States and, accordingly, on the judicial review of national policies performed by the (quasi)judicial branch of a specific international legal regime. As such, the findings of this article can be easily compared with the analogue operation performed by other courts or tribunals (amongst others, the Court of Justice of the European Union, the European Court of Human Rights, investment tribunals).

[1] For an insightful account of these doctrines as the normal toolkit to perform 'constitutional interpretation and adjudication' see, for instance, Ernst-Ulrich Petersmann, 'De-Fragmentation of International Economic Law Through Constitutional Interpretation and Adjudication with Due Respect for Reasonable Disagreement' (2008) 6 Loyola U Chicago Intl L Rev 209-248. See also Michel Rosenfeld, 'Rethinking Constitutional Ordering in an Era of Legal and Ideological Pluralism' (2008) 6 Intl J Con L 415-456.

The rationale of these techniques is closely related to the general regulatory design of the supra-national institution concerned: organizations whose purpose is the creation of common standards for, or the regulatory harmonization of, national regimes in certain areas are naturally inclined to discourage regulatory diversity and fragmentation (and to allow for State discretion only subject to certain conditions).[2] On the other hand, other organizations tend to acknowledge ample freedom in relation to domestic policies (the means): the focus is rather on the attainment of the agreed objectives (the ends), and regulatory diversity is the default standard ('*laissez-régler*'). The WTO system, in particular, is precisely premised on the principle of de-regulation, or of negative integration.[3] States retain their sovereign power to choose and implement their regulatory policies as they deem fit, as long as they do not interfere with the international commitments under the WTO.[4]

In this article, the WTO reports on the necessity test will be analyzed, to ascertain whether States preserve a considerable regulatory margin of manoeuvre in the WTO system, and whether such margin has a predictable scope. A negative answer to either of these questions would suggest that the spirit of negative integration has given way, at least in part, to normative harmonization and centralized cost-benefit assessment. To put it bluntly, the de-regulatory inspiration of the GATT 1947 is maybe under the wearisome attack of the necessity test, as performed by the Panels and the AB. Could it be that necessity has, to some extent, killed the GATT?

II. THE NECESSITY TEST IN ART XX GATT

[2] The European Union is one example in many fields of regulation, although it is suggested that the enlargement of its membership and competences might cause a shift from a model of positive integration to one of negative integration, see Giandomenico Majone, 'Liberalization, Re-Regulation, and Mutual Recognition: Lessons from Three Decades of EU Experience' (2009) Scottish Jean Monnet Centre Working Paper Series, Vol. 1, No. 1, <http://www.gla.ac.uk/media/media_111516_en.pdf>.

[3] Among the multiple works expounding this premise, Petros C Mavroidis, 'Market Access in the GATT' (2008) STALS Research Paper 7/2008 <http://www.stals.sssup.it/site/files/stals_Mavroidis.pdf> stands out for clarity. See also Joel P Trachtman, 'The Constitutions of the WTO' (2006) 17 EJIL 623-646.

[4] The classic view is encapsulated in the following passage in Armin von Bogdandy, 'Law and Politics in the WTO – Strategies to Cope with a Deficient Relationship', in Jochen A Frowein and Rüdiger Wolfrum (eds.), *Max Planck Yearbook of United Nations Law* 5/2001 (2001) 609, 657: 'The [European Court of Justice]'s jurisprudence is based on the premise that legislative correction is possible at the supranational level. That possibility does not obtain within the WTO'.

The necessity test, as it stands now, is briefly but comprehensively enounced in the following recital of the report of the Appellate Body (AB) in the *Brazil Tyres* case:

In order to determine whether a measure is 'necessary' within the meaning of art XX(b) of the GATT 1994, a panel must assess all the relevant factors, particularly the extent of the contribution to the achievement of a measure's objective and its trade restrictiveness, in the light of the importance of the interests or values at stake. If this analysis yields a preliminary conclusion that the measure is necessary, this result must be confirmed by comparing the measure with its possible alternatives, which may be less trade restrictive while providing an equivalent contribution to the achievement of the objective pursued. It rests upon the complaining Member to identify possible alternatives to the measure at issue that the responding Member could have taken. ... [I]n order to qualify as an alternative, a measure proposed by the complaining Member must be not only less trade restrictive than the measure at issue, but should also 'preserve for the responding Member its right to achieve its desired level of protection with respect to the objective pursued'. ... If the responding Member demonstrates that the measure proposed by the complaining Member is not a genuine alternative or is not 'reasonably available', taking into account the interests or values being pursued and the responding Member's desired level of protection, it follows that the measure at issue is necessary.[5]

It is possible to break this composite test down into single elements, each amenable to either of the two sub-tests which can be referred to, respectively, using the 'weighing and balancing' (WAB) formula[6] and the LTRM acronym (which stands for Least Trade-Restrictive Means). From the passage above, it transpires that both these tests aim to assess whether a certain measure is indeed necessary, the difference being that whereas the WAB test yields a 'preliminary' conclusion, the 'confirmation' comes from the LTRM test.

---

[5] WTO, *Brazil – Measures Affecting Imports of Retreaded Tyres—Report of the Appellate Body* (17 December 2007) WT/DS332/AB/R [156].

[6] Which was first stated in WTO, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef—Report of the Appellate Body* (11 December 2000) WT/DS161/AB/R, WT/DS169/AB/R, at [164]. The AB referred to the process 'of weighing and balancing a series of factors which prominently include the contribution made by the compliance measure to the enforcement of the law or regulation at issue, the importance of the common interests or values protected by that law or regulation, and the accompanying impact of the law or regulation on imports or exports.' More on this below.

Briefly, the necessity test routine comprises the following steps:

Weighing and Balancing (WAB)

I. Assessment of the importance of the value at stake (the Value);
II. Assessment of the contribution that the challenged measure makes to the Value;
III. Assessment of the trade-restrictiveness of the measure.

Least-Trade Restrictive Means (LTRM) test

I. Ascertainment of the correspondence between the Value and one of the prongs of Art. XX GATT;
II. Annotation of the level of protection of the Value sought by the respondent;
III. Ascertainment that no alternative measure can achieve the same level of protection, while being less trade-restrictive;
IV. Ascertainment that alternative measures identified under 3) are reasonably available.

A first claim of this article is that the WAB-half – at least in the way it has operated so far before the Panels and the AB – brings no added value to the LTRM-half, other than serving as a gateway filter for unacceptable measures. However, this does not mean that the necessity test is reduced to a mechanical analysis, because some discretion-laden pattern is indeed discernible in the use that Panels and AB make of the LTRM test. The second claim, it follows, is that an element of *stricto sensu* proportionality (or cost-benefit analysis) guides at times the necessity test performed by Panels and AB, but this exercise of appreciation is not embedded in the balancing effort (as it would be normal to assume), but in the loose application of the LTRM analysis (which would, in principle, bar discretionary evaluation).

This article takes stock of the WTO's *grands arrêts* on necessity, by tracing the development of the test in a rigorous chronological perspective. This analysis permits to appreciate how the test was repeatedly integrated and adjusted over time, and reveals the process that led to the over-elaborated version described above. Such a retrospective will lead to the conclusion that, in essence, some elements of the necessity test as it stands now are less an essential part thereof than a residue of accumulation, and could be interpreted away without being too concerned with their distinct *effet utile*.

The interpretation and application of art XX GATT, and the necessity

test in particular, have attracted a fair amount of scholarly attention;[7] however, the following overview focuses on some original elements that have been generally disregarded in the literature. Namely, this article intends to substantiate the claim that the WAB-test is irrelevant and that the case-law reveals an unavowed pattern of judicial interference into States' policies.

III. NECESSITY IN THE GATT-DAYS: ENTER THE LTRM

As anticipated, the word 'necessary' in art XX(a) (b) and (d) GATT gradually unfolded into a complex legal test. It is noteworthy that, in the words of Schoenbaum, this provision led to a semantic shift, since in the test currently in use 'necessary no longer relates to the protection of living things, but to whether or not the measure is a 'necessary' departure from the trade agreement."[8] Hence, the idea that all deviations from the trade obligations should be minimized is at the basis of the LTRM paradigm.

The LTRM principle was first used in 1990 by the GATT Panel *US – Section 337*,[9] which seemingly took cues from the EC's suggestion to the

[7] Among the most recent and complete works, see Glyn Ayres and Andrew D. Mitchell, 'General and Security Exceptions under the GATT and GATS' in Indira Carr, Jahid Bhuiyan and Shawkat Alam (eds), *International Trade Law and WTO* (Federation Press 2012); Gisele Kapterian, 'A Critique of the WTO Jurisprudence on 'Necessity' (2010) 59 ICLQ 89; Alec Stone Sweet and Jud Mathews, 'Proportionality Balancing and Global Constitutionalism' (2008) 47 Columbia J Trans L 73; Mads Andenas and Stefan Zleptnig, 'Proportionality: WTO Law: in Comparative Perspective' (2007) 42 Texas Intl L J 371; Donald H. Regan, 'The meaning of 'necessary' in GATT Article XX and GATS Article XIV: the myth of cost–benefit balancing' (2007) 6 WTR 347. For a detailed analysis of the single disputes relevant to the definition and application of this test see also Filippo Fontanelli, 'Whose Margin Is It? State Discretion and Judges' Appreciation in the Necessity Quicksand' in Filippo Fontanelli, Giuseppe Martinico and Paolo Carrozza (eds), *Shaping Rule of Law Through Dialogue. International and Supranational Experiences* (European Law Publishing 2009) 377. For an earlier article that served as a stepping stone for the 2009 article and the present one, see Alan O. Sykes, 'The Least Restrictive Means' (2003) 70 U Chicago L Rev 403.

[8] Thomas J Schoenbaum, 'International Trade and Protection of the Environment: The Continuing Search for Reconciliation' (1997) 91 AJIL 269, 276, mentioned in Kapterian (n 7) 103.

[9] *United States – Section 337 of the Tariff Act of 1930* (1990) GATT BISD 36S/345, 392-93 [5.26]: '[A] contracting party cannot justify a measure inconsistent with another GATT provision as 'necessary' ... if an alternative measure which it could reasonably be expected to employ and which is not inconsistent with other GATT provisions is available to it'.

Panel,[10] an all the more reasonable hypothesis since Pierre Pescatore (former judge at the European Court of Justice and passionate advocate of the process of European integration) was sitting on the Panel. At that (pre-WTO) time, the LTRM analysis was the only selection device used to check the GATT-compliance of measures allegedly falling under one of the prongs of art XX(a), (b) and (d) GATT (besides the application of the *chapeau*), and other Panels adopted it after its first appearance.[11]

However, this early version of the LTRM test was still relatively under-developed. For instance, Panels tended to accept alternative less-restrictive measures without careful consideration of the level of protection set by the respondents; neither did they spend particular efforts to make sure that the alternative measure was reasonably available to them.[12] Nevertheless, in the early 90s Panels became familiar with the idea that GATT-inconsistency could be measured and arranged on a scale of gravity, and it was possible to identify the measure that was *less* GATT-inconsistent than the others.[13]

The LTRM test, on its face, is a formula of (Pareto) efficiency,[14] and it did

[10] According to Stone Sweet and Mathews (n 7) 156, the EC put forward this test bearing in mind the doctrine of proportionality in use in EU and ECHR law. The US, on its part, had proposed a stricter test based on rational analysis of the measure (the so-called 'strict in theory, fatal in fact' strict scrutiny test). Seemingly, 'each side was proceeding on the basis of their understanding of how Least Restrictive Means tests are used *in their own systems*'.

[11] *Thailand – Restrictions on Importation of and Internal Taxes on Cigarettes* (1991) GATT BISD 37S/200, 223 [74-75] (inconsistencies with GATT obligations arising from a national measure were deemed to be legitimate only as far as they were 'unavoidable'; the word 'necessary' has the same meaning in art XX(b) and (d)); *United States – Restrictions on Imports of Tuna* (1991) GATT BISD 39S/155 (unadopted); *United States – Measures Affecting Alcoholic and Malt Beverages* (1992) GATT BISD 39S/206 [5.52].

[12] On this, see Kapterian (n 7) 103, and the literature referred to therein, in particular Deborah A Osiro, 'GATT/WTO Necessity Analysis: Evolutionary Interpretation and Its Impact on the Autonomy of Domestic Regulation' (2002) 29 Legal Issues of Economic Integration 123, 127-128.

[13] Kapterian (n 7) 105, holds that not only is this assessment difficult to make, due to the absence of a shared view on how to measure GATT-inconsistency, but also it does not seem to be allowed by art XX GATT, because the meaning of the word 'exception' 'does not provide space for shading.' However, this argument does not appear compelling: the LTRM test is an interpretative elaboration of the word 'necessity,' therefore, it does not relate to the (indeed monolithic) exceptional nature of the measure, but to the conditions precedent for it to arise, which might well be dependent on a value judgment.

[14] By this we mean that, since it keeps one of the variable fixed (achievement of the regulatory purpose), it is not a full-fledged cost-benefit analysis, of the kind used to

not take long for Contracting Parties to wonder why efficiency had become a standard of review in a regime that, purportedly, should eminently care about non-discrimination and negative integration, rather than regulatory positive harmonization. Moreover, the strict LTRM test understandably disconcerted the WTO Members, as it seemed to 'require dispute settlement panels to dictate the specific measure to be adopted by a WTO Member, since presumably there was only one measure among all the alternatives that was the 'least inconsistent' with the GATT 1994.'[15]

When the US acted as responding party in the *US – Shrimps* dispute, it fought at length with the received interpretation of necessity deriving from the *US – Section 337* report, notably protesting that the intricate LTRM test and the steps that it required could not be inferred from the normal meaning of the art XX(b) provision in light of the standards of the Vienna Convention on the Law of Treaties:

The use of one word, 'necessary', was a slender reed indeed on which to hang such an extensive and complex set of obligations. Rather than attempt to impose a reading of the text that no reader could be expected to know, it would be wiser to interpret the language in accordance with its normal meaning.[16]

The backlash against the LTRM test was not simply an element of the US' defensive strategy, but more generally an instance of the Parties' distrust of the Panels' and AB's activism.[17] According to the US, the *chapeau* of art XX (mandating that domestic measures be *applied* non-discriminatorily and non-arbitrarily, and not disguise a trade restriction) would have been sufficient to ensure that protectionist measures could not stand scrutiny, and the LTRM test was, in short, uncalled-for and intrusive.

For a while, certain States simply could not come to terms with the LTRM test, as illustrated by Argentina's vehement complaint[18]:

---

maximize global welfare, but a truncated version thereof. On this, see Chad P Bown and Joel P Trachtman, 'Brazil – Measures Affecting Imports of Retreaded Tyres: A Balancing Act' (2009) 8 WTR 85.

[15] WTO, *United States – Import Prohibition of Certain Shrimp and Shrimp Products—Report of the Panel* (15 May 1998) WT/DS58 [3.228] summarizing the position of the US.

[16] ibid [3.225].

[17] ibid [3.226]: 'After all,' the US stated 'the basic thrust of the GATT was to prevent protectionism, not to intrude on the decision making of the contracting parties when pursuing legitimate policy objectives such as environmental protection'.

[18] WTO, *Argentina – Measures Affecting the Export of Bovine Hides and the Import of Finished Leather—Report of the Panel* (19 December 2000) WT/DS155 [8.251-252].

Where does subparagraph (d) prescribe that the government of a State must analyze an array of options, and choose the least restrictive? What is the yardstick for defining what is less restrictive? Accepting this approach would mean supplanting the sovereignty of governments by a panel's evaluation. ... A certain degree of discretion must therefore be allowed to the member invoking the exception in determining which measure is necessary for securing observance of laws and regulations that are not inconsistent with the general agreement.

IV. THE ABSOLUTE FREEDOM TO SET THE LEVEL OF PROTECTION

In the first case of the WTO era, *US – Gasoline*,[19] the Panel applied art XX(b) GATT, and significantly expanded the necessity test used hitherto.[20] Firstly, it took cognizance of the different words that the Contracting Parties used in art XX GATT to indicate the link between the measure and the various values pursued, stating that a different meaning must be attached to each formulation. In particular, the 'necessity' word (letters a, b and d) postulated a closer connection between the measure and the policy objective than that required by the 'related to' formula (letters c, g and e).

The Panel set up a three-tiered test to perform the judicial review under art XX(b) GATT, requesting the responding party to establish

(1) that the policy in respect of the measures for which the provision was invoked fell within the range of policies designed to protect human, animal or plant life or health; (2) that the inconsistent measures for which the exception was being invoked were necessary to fulfill the policy objective; and (3) that the measures were applied in conformity with the requirements of the introductory clause of Article XX.[21]

The Panel ran the LTRM test and even suggested an alternative measure, less restrictive than (and as effective as[22]) the US one. In so doing, the

[19] WTO, *United States – Standards for Reformulated and Conventional Gasoline—Report of the Panel* (29 January 1996) WT/DS2, WT/DS4.
[20] For the record, the US invoked the art XX(g) GATT defense as well (relating to the protection of limited natural resources), but the AB found that the measure under review was applied in a discriminatory way, and therefore breached the *chapeau* of art XX GATT.
[21] See ibid [6.20].
[22] Ibid [6.22-29]. In particular, [6.27]: 'slightly stricter overall requirements applied to both domestic and imported gasoline could offset any possibility of an adverse

Panel took upon itself the burden of proof regarding the research of the alternative, and seemingly deviated from the general principle that the party who wants to invoke an exception must prove that preconditions for its application are met.[23]

More importantly, the newly established Appellate Body clarified that the subject of the judicial review of national regulatory measures is the measures themselves, not the value that they pursue (the Value) and the expected level of attainment thereof.[24] As Mavroidis lucidly puts it, '*a WTO adjudicating body* [...] *can extend its judicial review only with respect to the means used to achieve the ends: ends are not justiciable, means are.*'[25]

The *US – Gasoline* dispute, ultimately, established the untouchable nature of the level of protection set unilaterally by the State (let alone the choice of the value to protect),[26] although this alleged autonomy has come under

---

environmental effect from these causes, and allow the United States to achieve its desired level of clean air without discriminating against imported gasoline. Such requirements could be implemented by the United States at any time'. According to Kapterian (n 7) 103, the test was applied somehow loosely, since the Panel was content with an alternative capable to achieve one of the objectives of the measure 'often,' but presumably not always, as sought after by US. On this loose version of the LTRM (where the alternative is less trade-restrictive, but also slightly less effective than the one quashed), see Donald H Regan, 'Judicial Review of Member-State Regulation of Trade within a Federal or Quasi-Federal System: Protectionism and Balancing, *Da Capo*' (2001) 99 Mich L Rev 1853, 1899–1900.

[23] Moreover, the AB snubbed US' attempt to use the costliness of the alternative measure as a proof of its non-availability, by using art 27 of the Vienna Convention, see WTO, *US – Gasoline—Report of the Appellate Body* (29 April 1996) WT/DS2/AB/R, 27: 'The fact that the United States Congress might have intervened, as it did later intervene, in the process by denying funding, is beside the point: the United States, of course, carries responsibility for actions of both the executive and legislative departments of government'.

[24] See [7.1]: 'It was not [the Panel's] task to examine generally the desirability or necessity of the environmental objectives of the Clean Air Act or the Gasoline Rule. ... Under the General Agreement, WTO Members were free to set their own environmental objectives, but they were bound to implement these objectives through measures consistent with its provisions, notably those on the relative treatment of domestic and imported products'.

[25] Petros C Mavroidis, *The General Agreement on Tariffs and Trade: A Commentary* (OUP 2005), 191. See also the similar dictum in WTO, *Canada – Certain Measures Concerning Periodicals—Report of the Panel* (14 March 1997) WT/DS31/R [5.9]: 'we are neither examining nor passing judgment on the policy objectives of the Canadian measure regarding periodicals; we are nevertheless called upon to examine the instruments chosen by the Canadian Government for the attainment of such policy objectives'.

[26] See Report of the Panel (n 19), [6.22]: 'it was not the necessity of the policy goal that was to be examined, but whether or not it was necessary [to adopt the

scrutiny over time, and it is now controversial whether States are actually free to choose their preferred level of protection.[27]

V. THE ADDITIONAL CHALLENGE OF ART XX(D) GATT

Shortly after that, the Panel in *Canada – Periodicals* rejected Canada's invocation of art XX(d) GATT. The Canadian measure fell even before making it to the necessity test, because Canada failed to prove that it 'secured compliance' with the designated law.[28] This outcome suggested that, although 'ends' are safe from judicial review (see above), it is not guaranteed that *all* measures will get undisturbed to the necessity stage, especially if they are allegedly covered by art XX(d) GATT (as opposed to letter (a) and (b)).

Indeed, measures falling under art XX(d) GATT pursue a Value (compliance with a national law, ie enforcement of its obligations) that is not an abstract one like 'public morals' or 'human health,' hence a judicial body can reasonably assess whether the trade-restrictive measure is *prima facie* instrumental to the enforcement of the national norm invoked, even before getting to the LTRM phase, where that contribution is examined and measured. Incidentally, neither of these tests implies any review of the aim of the national laws itself or of the policy the latter are designed to promote. In other words, Panels must initially verify whether the domestic measures do actually 'secure compliance' with a wider discipline, simply assessing *prima facie* the existence of a means-ends relationship between the two. Afterward, the LTRM test examines the efficiency of the measures with respect to the national policy (as opposed to the general aim

---

challenged measures]. It was the task of the Panel to address whether these inconsistent measures were necessary to achieve the policy goal under Article XX(b). It was therefore not the task of the Panel to examine the necessity of the environmental objectives of the Gasoline Rule, or of parts of the Rule that the Panel did not specifically find to be inconsistent with the General Agreement.' On the absolute freedoms of Members to set their appropriate level of protection, see WTO, *Australia – Measures Affecting Importation of Salmon—Report of the Appellate Body* (6 November 1998) WT/DS18/AB/R [199].

[27] Michael Ming Du, 'Autonomy in Setting Appropriate Level of Protection under the WTO Law: Rhetoric or Reality?' (2010) 13 JIEL 1077-1102.

[28] See *Canada – Periodicals*, Report of the Panel (n 25) [3.5]: 'The general objective of these measures is to help the Canadian periodical industry raise advertising revenues. Tariff Code 9958 ensures the achievement of this goal, with Section 19 of the Income Tax Act'. The Panel reaches its conclusion (see [5.10]) through the test set by a GATT Panel, whereby the 'to secure compliance' formula means 'to enforce obligations under laws and obligations,' not 'to ensure the attainment of the objectives of the laws and regulations' (see Report of the Panel in *EC – Regulations on Imports of Parts and Components* (1990) GATT BISD 37S/132 [5.14-5.18]).

it pursues, such as fighting evasion, or securing efficient border control). These two steps are not concerned with questioning the appropriateness of the Value pursued, and are based on seemingly technical evaluations.[29]

An application of the XX (d)-specific preliminary test is visible in the *Mexico – Soft Drinks* case, where the Panel and the AB did not get as far as examining whether the challenged measures were 'necessary' under art XX(d) GATT, since the latter did not 'secure compliance with [the relevant national] laws and regulations,' and therefore fell outside the scope of the exception. To appreciate how the same preliminary analysis does not apply to exceptions other than the XX(d) ones, suffice it to recall the Panel's Report of the *EC – Tariff Preferences* case. In that case, the Panel found that the challenged measure did not fall under the health heading, since it was 'not one designed for the purpose of protecting human life or health'.[30] Nevertheless, the panellists, rather than stopping the review, went on–*arguendo*–to demonstrate that the necessity test and the *chapeau* requirements were not met.

Intuitively, as seen above, the different approach is due to the different degree of confidence that Panels have when dealing with the review of measures allegedly covered by art XX(b) or XX(d) GATT. Even before entering the necessity test, the Panel can refuse to apply the art XX(d) GATT justification just by focusing on the 'securing compliance' parameter, and without questioning the legitimacy of the domestic policy indicated by the State. To sum up, the preliminary test applicable under art XX(d) GATT ('is the measure *prima facie* capable of securing compliance with the national law?') adds a layer to the review, but does not threaten the neutrality of the analysis with respect to the Value. On the contrary, it takes some temerity for a Panel to state that a measure does not fall under the category of art XX(b) GATT and, as a consequence, does not even deserve to reach the necessity test. Such a finding implies an appraisal of the declared Value and a *prima facie* understanding of the measure's contribution to it. Therefore, it is not surprising that the Panel in *EC – Tariff Preferences*[31] was self-conscious about its preliminary finding,

[29] Contrarily, when it comes to measures allegedly covered by the art XX(b) GATT exception, not only is the LTRM test virtually always granted, but it must also be performed solely as regards to the 'abstract' value (health promotion), irrespectively of whether the measures are necessary to enforce any wider national regulation scheme. It goes without saying that in such cases it is easier for the responding Party to argue that the measure brings at least some contribution to the (even prospective) attainment of the public interest pursued.

[30] See WTO, *European Communities – Conditions for the Granting of Tariff Preferences to Developing Countries—Report of the Panel* (1 December 2003) WT/DS246/R [7.210].

[31] The same holds true with respect to the *China – Raw Materials* case, see below.

and preferred to render it more solid showing that the measure would have been struck nevertheless, even if its initial decision on the non-subsumption under art XX(b) GATT were ultimately wrong.[32]

Another notable example in this respect is *Colombia – Entry Ports*, in which Colombia tried to defend some border measures invoking art XX(d) GATT, namely compliance with national regulations aimed at the prevention of under-pricing techniques and smuggling. The Panel, relying on statistical data, concluded that the measures were virtually unable to reduce smuggling. Therefore, they were not necessary, for they did not contribute to the enforcement of the relevant national policy.[33]

VI. ENTER THE WAB

In this well-known dispute, the claimants held that Korea's measures requiring that imported beef be sold only in specialized imported beef stores, as well as Korean laws and regulations restricting the resale and distribution of imported beef, resulted in a violation of art III.4 GATT (national treatment). Korea objected, *inter alia*, that these measures were necessary to comply with its Unfair Competition Act (a domestic regulation providing for consumers' protection), for the purpose of preventing retailers from deceiving consumers by selling imported beef as domestic beef. In the course of this controversy, the necessity test underwent a momentous mutation, possibly due to Korea's incisive defence, which sought to hamper the Panel's review of necessity, invoking the mantra of regulatory autonomy (and sending out the veiled threat that activism accusations could follow):

Korea noted that so far GATT/WTO case law has not explored the link between regulatory diversity, on the one hand, and the necessity

[32] This difference is efficiently encapsulated in Panama's remark in the WTO, *Colombia – Indicative Prices and Restrictions on Ports of Entry—Report of the Panel* (27 April 2009) WT/DS366/R [7.495]: 'whereas the Art. XX(b) exception is purpose-oriented, the Art. XX(d) exception is «functional»'.

[33] ibid [7.588]. In WTO, *Canada – Measures Relating to Exports of Wheat and Treatment of Imported Grain* (6 April 2004) WT/DS276/R, the Panel was dismissive of the possibility to justify the challenged measure under art XX(d) GATT, see [4.371-374]. It was enough for the Panel to note that Canada had not proven that the grain segregation measures it adopted contributed to the enforcement of the national policies on competition and on fair commercialization of grain. The half-hearted invocation of art XX(d) GATT was sweepingly rejected by the AB (in WTO, *Thailand – Customs and Fiscal Measures on Cigarettes from The Philippines—Report of the Appellate Body* (17 June 2011) WT/DS371/AB/R, [175-180]), for Thailand's failure to make a *prima facie* defense and demonstrate that the measures were necessary under the general exception.

requirement, on the other. Korea submitted that there is a correlation between the two in the sense that were a regulatory objective to be sought in a very strict manner, the choice of instruments would consequently be influenced. Since the level of protection sought cannot be put into question, the choice of instrument will have to be appreciated in the same context.[34]

The Panel rejected Korea's defence and quashed the challenged measures, pointing at less restrictive alternatives,[35] and to an inconsistency of Korea's policies.[36] These findings were subsequently upheld on appeal.
The AB (unlike the *Argentina – Hides and Leather* Panel) took upon itself the task of expounding the meaning of the term 'necessity' and to draw the limits of the LTRM test. It clarified that art XX GATT did not cover only 'indispensable' measures (which are 'certainly' allowed protection[37]), but also measures bearing a slightly less direct link with the Value, provided that they are not just 'making a contribution' thereto.[38]

---

[34] WTO, *Korea – Measures Affecting Imports of Fresh, Chilled and Frozen Beef—Report of the Panel* (31 July 2000) WT/DS161/R, WT/DS169/R [242].
[35] As for the LTRM test, see Report of the Panel [672]: 'For instance, a generally applied record-keeping requirement backed with sanctions would constitute a WTO consistent alternative to the WTO inconsistent dual retail system. If foreign beef shops can keep book-records, it is difficult to see why the same could not be requested from domestic shops'.
[36] In fact, the AB correctly stresses the legitimacy of the reasoning by which the panel had noted the absence of similar measures in other market sectors. Korea had alleged that this fact could not imply that the stricter measures adopted for the beef sector were not necessary, as this finding would amount to an interference in Korea's right to set the level of protection at its sole discretion, and to set different levels of protection in different market sectors. The AB stated that this comparative analysis had in fact the different purpose of highlighting that efficient alternative measures were available, as the one Korea used to enforce in the non-beef sectors of the market. See *Korea – Beef*, Report of the Appellate Body (n 6) [175-178]. Benn McGrady, 'Necessity Exceptions in WTO Law: Retreaded Tyres, Regulatory Purpose and Cumulative Regulatory Measures' (2009) 12 JIEL 153, 159, notes that the panel and the AB characterized Korea's goal in different ways, somewhat contrary to the principle that each State has the power to define it autonomously, and that the loose formulation of the goal adopted by the AB made it easier to find equally efficient alternatives.
[37] This is also a controversial statement. By putting indispensible measures in a safe haven, the AB makes it unlikely that the balancing is applied at all, see Regan, The Meaning (n 7) 354.
[38] *Korea – Beef*, Report of the Appellate Body (n 6), [161]. See the similar holding of the ECtHR in the *Handyside* case (*Handyside v. the United Kingdom*, App. no. 5493/72 (ECHR, 7 December 1976) 5, where the judges observed that 'the adjective 'necessary,' ... is not synonymous with 'indispensable' [and] neither has it the flexibility of such expressions as ... 'admissible,' ... 'useful,' 'reasonable,' or 'desirable'.

More importantly, the AB's reasoning on the necessity of measures that are not 'indispensable' encouraged Panels and AB to embark on the review of the Values at stake, opening the floodgates to the WAB test and to cost-benefit analysis. Firstly, the AB hinted at a graduation of importance of the Values, suggesting that 'the more vital' the value, the easier it would be for the measure to prove 'necessary.'[39]

It is remarkable to learn from the AB, keeping in mind the *Section 337* paradigm, that the outcome of the necessity test is not solely a matter of efficiency, but also depends on a value-judgment (that is, how important the pursued Value is). This was but the first crack in the LTRM building. The AB went on to add other elements that should provide guidance in the 'process of weighing and balancing a series of factors,' an exercise that implies, on its face, a significant degree of discretion by the reviewer.[40] Specifically, it lays down two additional guidelines: the greater the contribution of the measure to the enforcement of the national policy, and/or the lighter its trade-restrictiveness, the more easily it will pass the necessity test. Enter the WAB test:

In sum, determination of whether a measure, which is not 'indispensable', may nevertheless be 'necessary' within the contemplation of Article XX(d), involves in every case a process of weighing and balancing a series of factors which prominently include the contribution made by the compliance measure to the enforcement of the law or regulation at issue, the importance of the common interests or values protected by that law or regulation, and the accompanying impact of the law or regulation on imports or exports.[41]

This move signalled that the AB is into the business of looking into the merits of the measures under review (not simply into their efficiency and their functional design) and of embarking on a review of proportionality.[42]

---

[39] *Korea – Beef*, Report of the Appellate Body (n 6), [162]. To help the adjudicator in handling such an indistinct test, the AB argues that it is helpful to 'take into account the relative importance of the common interests or values that the law or regulation to be enforced is intended to protect. The more vital or important those common interests or values are, the easier it would be to accept as 'necessary' a measure designed as an enforcement instrument'.

[40] ibid [163].

[41] ibid [164].

[42] See Regan, The Meaning (n 7) 355-356: 'there is nothing in the text of Article XX(d) to suggest that different regulatory purposes are accorded different values by Article XX(d). *A fortiori*, there is nothing to suggest that it is appropriate for the Appellate Body to rank Members' regulatory purposes according to the Appellate Body's

There is an inherent contradiction between this balancing activity and the oft-repeated assumption that Member States have the right 'to determine for themselves the level of enforcement of their WTO-consistent laws,' that is, the level of protection of the Value.[43] However, the AB declared that the fully fledged WAB test was already 'encapsulated' in the LTRM *Section 337* test, packed within the 'reasonable expectation that the contracting party employs' alternative measures. In other words, the AB allegedly did nothing new, and simply unpacked the 'reasonable' element, so as to obtain the WAB test.

In the *EC – Asbestos* case[44] the WAB/LTRM compound test of *Korea – Beef* was applied again, although ultimately the French ban at bar was found to be *indispensable* to achieve the Value (a zero-risk protection against asbestos-related illness[45]). Accordingly the AB spared the ban from the WAB assessment,[46] after noting that that the preservation of human life and health is 'vital and important in the highest degree.'[47] On this occasion, the EC's invocation of a zero-risk policy proved successful, as it made the LTRM test an uphill battle for the claimant. Since then, many responding parties have tried to mimic this strategy when invoking art XX GATT, but Panels and AB have countered this strategy, by somehow assessing the veracity (not the appropriateness, of course) of zero-risk declarations.[48]

---

intuitions about their value'. See Andrew Lang, *World Trade Law after Neoliberalism - Reimagining the Global Economic Order* (OUP 2011) 323, noting that the AB's statement implies a strong test of *stricto sensu* proportionality, and Peter Van den Bossche, 'Looking for Proportionality in WTO Law' (2008) 35 Legal Issues of Economic Integration 283-294.

[43] *Korea – Beef*, Report of the Appellate Body (n 6) [176]. This contradiction is lucidly described in Joseph H.H. Weiler, 'Comment on Brazil – Measures Affecting Imports of Retreaded Tyres' (2009) 8 WTR 137, 141. See also Regan, The Meaning (n 7) 353 ff.

[44] WTO, *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products—Report of the Appellate Body* (12 March 2001) WT/DS135/AB/R.

[45] WTO, *European Communities – Measures Affecting Asbestos and Asbestos-Containing Products—Report of the Panel* (18 September 2000) WT/DS135/R [8.217]: 'controlled use does not constitute a reasonable alternative to the banning of chrysotile asbestos that might be chosen by a decision-maker responsible for developing public health measures, bearing in mind the objectives pursued by France [absolute halt to risk-spreading]'.

[46] In this respect, see Regan, The Meaning (n 7), and Robert Howse and Elisabeth Türk 'The WTO Impact on Internal Regulations—A Case Study of the Canada–EC Asbestos Dispute', in Graínne de Búrca and Joanne Scott (eds), *The EU and the WTO. Legal and Constitutional Issues* (Hart 2002) 283, 324.

[47] See *Report of the Appellate Body* (n 44), [172].

[48] Since the notion of zero-risk 'is an abstraction,' it is understandable that the adjudicators feel entitled to reshape it as a 'de minimis' risk-tolerance, see Damien J Neven and Joseph HH Weiler, 'Japan - Measures Affecting the Importation of

Famously, the AB in *Korea – Beef* second-guessed Korea's declared objective to eliminate 'all fraud,' noticing that such 'unlikely' objective would probably require a ban on all imports, hence Korea's policy objective was toned down to a more modest 'considerable reduction' of fraud, an aim that could be conveniently achieved also by less-restrictive measures than those adopted.[49] Conversely, the Panel and the AB showed more deference to Brazil's declaration, in the *Tyres* case, that the purpose its measure intended to achieve was the reduction of the risks of waste tyre accumulation 'to the maximum extent possible.' Brazil arguably got away with that because it managed to convince the Panel that reduced tyre-accumulation was *the* Value, whereas it actually was a means to protect health (the real Value). In so doing, it benefitted from an *Asbestos*-treatment with respect to necessity. This is further developed in part 8, below.

VII. PAYING LIP-SERVICE TO THE WAB: *GAMBLING* AND *CIGARETTES*

The *US – Gambling* and *Dominican Republic – Cigarettes* cases[50] added nothing to the *Korea – Beef* test (apart from the *Gambling* one inaugurating the case-law on a new Value, ie morals and public order[51]), but it is worthwhile to examine how the WAB test played out in these disputes. In the reports *Korea – Beef* and *US – Asbestos*, in spite of the large amount of reasoning devoted to its formulation, the balancing moment hardly contributed to the *dispositifs* (in *Asbestos*, the measure was indispensable, therefore no balancing was needed; in *Korea – Beef* the conclusion was reached through the LTRM analysis, even if a WAB balance would have been very easy to assess: since Korea had an outright ban in place, the 'trade-restrictiveness' score was clearly at its maximum).[52]

Apples: One Bad Apple?' in Henrik Horn and Petros C Mavroidis (eds), *The WTO Case Law of 2003* (CUP 2006) 289–290.

[49] *Korea – Beef*, Report of the Appellate Body (n 6) [172] and [178].

[50] WTO, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services—Report of the Appellate Body* (7 April 2005 WT/DS285/AB/R; *Dominican Republic – Measures Affecting the Importation and Internal Sale of Cigarettes—Report of the Appellate Body* (25 April 2005) WT/DS302/AB/R.

[51] WTO, *United States – Measures Affecting the Cross-Border Supply of Gambling and Betting Services—Report of the Panel* (10 November 2004) WT/DS285/R [3.279]: 'Remote supply of gambling raises significant concerns relating to the maintenance of public order and the protection of public morals.' See also [3.273-277]. Morals and public order are protected under art XIV(a) GATS, the avatar of art XX(a) GATT.

[52] See Regan, The Meaning (n 7) 361, referring to *Asbestos*, *Gambling* and *Cigarettes*: 'when it comes to actually deciding the case, all three rely on the principle that

As it turned out, the WAB did not play a significant role in these two cases either. In *Gambling*, the AB found that the US measures were necessary in the abstract, but were applied in violation of the *chapeau* of art XX GATT.[53] In *Dominican Republic – Cigarettes*, instead, the challenged measures[54] were found to be unnecessary because they were so ineffective that many other GATT-consistent alternatives could be foreseen, and keeping the zero-tolerance level as a constant would have not been reasonable.[55] The respondent party attempted to lure the Panel into issuing a good cost-benefit report, describing the high importance of the public interest pursued (compliance with tax laws) and the minimal impact of the measures on the imports, but even if the Panel did not challenge this reconstruction the measure was not spared.[56]

Arguably, the *Korea – Beef* bit where the AB maintained that the balancing test is 'encapsulated' in the LTRM analysis (see above) might be revealing of the real stance of WTO judicial bodies towards the balancing task. The reason why the WAB is never really used to balance values between them and to assess their proportionality is that the WAB, in the particular WTO scenario, is of no practical use. Of its three elements, one is virtually untouchable[57] (the importance of the Value), and the LTRM test already

---

Members get to choose their own level of protection.' See also Caroline E Foster, 'Public Opinion and the Interpretation of the World Trade Organisation's Agreement on Sanitary and Phytosanitary Measures' (2008) 11 JIEL 427, 437; Howse and Türk (n 46) 326.

[53] The Panel had held that since the US had failed to negotiate with Antigua, it could not be sure that the measure was actually the LTRM available. The AB overturned this part of the Panel report.

[54] Which were allegedly taken to enforce the obligations under the national Tax Code, and were useful in preventing cigarette smuggling, see WTO, *Dominican Republic – Measures Affecting the Importation and Internal Sale of Cigarettes – Report of the Panel* (26 November 2004) WT/DS302/R [4.88].

[55] ibid [7.228]: 'Dominican Republic has not proved why, for example, providing secure tax stamps to foreign exporters ... would not be equivalent to the current tax stamp requirement in terms of allowing it to secure the same high level of enforcement with regard to tax collection and the prevention of cigarette smuggling.' More boldly, the AB found 'no evidence to conclude that the tax collection requirement secures a zero tolerance level of enforcement...' (see Report of the Appellate Body (n 50) [72]).

[56] On the difficulty of understanding the rationale and the functionality of the WAB test, see Steve Charnovitz, 'The WTO's Environmental Progress' (2008) 10 JIEL 685.

[57] The fact is, in any event, that 'in no case to date has a Panel or the Appellate Body found that a measure pursues values of only moderate or negligible importance.' See Ayers and Mitchell (n 7) 18 (of the preview available online). Similarly, Regan, The Meaning (n 7) 363: 'the Appellate Body has yet to say that any specific legitimate

takes care of the other two (the 'less-restrictive but equally effective' quality of the sought-after alternative postulates that efficiency and restrictiveness are already known variables, and are decisive in appraisal of necessity).[58]

The classic balancing test, therefore, has little in common with a real proportionality test, nor does it allow for express cost-benefit analysis.[59] The 'weighing and balancing,' all things considered, must be seen as a preparatory exercise, a propaedeutic to the LTRM test. It is possible to get a glimpse of this unconfessed approach in *Dominican Republic – Cigarettes*, where the Panel, after running the WAB assessment, bridges to the LTRM as follows: 'having said that [referring to the WAB], the Panel will focus its analysis on whether [the measure] ... is in fact necessary [to achieve the Value],'[60] clearly suggesting that only the LTRM test is apt to ascertain the necessity of a measure, the WAB serving merely as a warm-up test.

---

regulatory purpose is less valuable than any other.' For some examples, see for instance the Report of the Panel in *US – Gambling* (n 51) [6.492], acknowledging that the interests and values protected by the challenged measures serve very important societal interests that can be characterized as 'vital and important in the highest degree' (see also [6.558]). Likewise, see *Dominican Republic – Cigarettes*, Report of the Panel (n 54) [7.215]: 'The Panel finds no reason to question the Dominican Republic's assertions in the sense that the collection of tax revenue ... is a most important interest for any country and particularly for a developing country such as the Dominican Republic.' See also WTO, *Brazil – Measures Affecting Imports of Retreaded Tyres—Report of the Panel* (12 June 2007) WT/DS332/R [7.112]. Even with respect to China's mundane attempt to invoke art XX(d) in WTO, *China – Measures Affecting Imports of Automobile Parts—Reports of the Panel* (18 July 2008) WT/DS339/R, WT/DS340/R, WT/DS342/R [7.360] the Panel could not help acknowledging the importance of the proclaimed interest (tax collection). See also WTO, *United States – Measures Relating to Shrimp from Thailand—Report of the Appellate Body* (16 July 2008) WT/DS343/AB/R and WT/DS345/AB/R [313], and *Colombia – Entry Ports*, Report of the Panel (n 32) [7.566] (fighting under-invoicing and money laundering). In this respect, see also McGrady (n 36) 162.

[58] See Regan, The Meaning (n 7) 357: 'the only consideration in the Appellate Body's list that is relevant to a cost–benefit balancing test and not to a less-restrictive alternative test is the value of the regulatory purpose, which as we have already seen is a seriously suspect consideration'.

[59] This cost-benefit analysis, in fact, was merely proclaimed in *Korea – Beef* and never applied, see Joel P Trachtman, 'Regulatory Jurisdiction and the WTO' (2007) 10 JIEL 631, 647. For an enlightening analysis of the necessity test under cost-benefit terms that takes into account the Learned Hand test and other similar formulas, see David Collins, 'Health Protection at the World Trade Organization - The J-Value as a Universal Standard for Reasonableness of Regulatory Precautions' (2009) 43 JWT 1071.

[60] *Dominican Republic – Cigarettes*, Report of the Panel (n 54) [7.215], quoted also in Kapterian (n 7) 122.

It is just argued, here, that the 'balancing' result is rarely spelled out in clear terms, and virtually never relied upon to decide on the WTO-legality of the measure.[61] Take for instance the *Colombia – Entry Ports* case. Formally, the Panel held that since one of the three factors of the 'balancing' was irremediably flawed (the measure made an insignificant contribution to the policy objective), the art XX(d) GATT defence did not stand.[62] Although seemingly the case was decided on the WAB, this outcome could have been the result of the least-restrictive test as well: given the low level of effectiveness, many better alternatives were available to the defendant (like in *Korea – Beef*). At most, the WAB is a simpler version of the LTRM, filtering out measures that are *prima facie* untenable.[63]

This might sound fair: after all, it was not clear in the first place how the Panels and the AB could be entitled to perform any sort of balancing between values, given the presumption for regulatory autonomy that reigns in the WTO. Balancing and proportionality are a prerogative of constitutional adjudication,[64] and are at variance with the negative integration paradigm described above.[65] However, as Sykes first showed in 2003, there is some discernible pattern in the practice of Panels and AB, whereby certain Values are treated more deferentially than others (in particular, the protection of human health[66]). The following section intends to account for this trend in the case-law, and explain how it pervades the application of the LTRM test (whereas the WAB slowly turned into what it actually is, i.e. little more than a boilerplate section of the reasoning), and in particular the search for 'reasonably available alternative measures.'

---

[61] Note, for instance, how the utmost importance of the public interest is a variable that did not affect the result in the *Cigarettes* case, or how the ruinous effect on trade of the remote-gambling ban was not, per se, sufficient to prevent the AB from finding it 'necessary,' and pass on to the *chapeau* test, in *Gambling*.

[62] *Colombia – Entry Ports*, Report of the Panel (n 32) [7.619].

[63] This use of WAB is consistent with the evidentiary regime: it is for the responding party to propose a *prima facie* case of necessity, and this is where the WAB should operate, see Christopher Doyle, 'Gimme Shelter: the 'Necessary' Element of GATT Article XX in the Context of the China-Audiovisual Products Case' (2011) 29 Boston U Intl L J 143, 159.

[64] On this, see extensively Stone Sweet and Mathews (n 7) 138 and *passim*.

[65] On the WTO incompetence to rule on similar conflict of values, see Steve Charnovitz, 'The Law of Environmental PPMs in the WTO: Debunking the Myth of Illegality' (2002) 27 YJIL 59, 101.

[66] See Petros C Mavroidis, George A Bermann and Mark Wu, *The Law of the World Trade Organization (WTO): Documents, Cases and Analysis* (West Group 2010) 693.

One clear example of this trend is that Panels and AB, from time to time, do not shy away from taking an exploratory detour to look into the consistency of the respondent party's policies, with respect to the chosen level of protection for values *other* than the Value.[67] This is expressly provided for in art 5.5 SPS, for the purpose of encouraging States to adopt sanitary and phytosanitary policies that are at least roughly homogeneous.[68] On the contrary, nothing in the GATT or in the basic formulation of the LTRM test suggests that a measure is *per se* less necessary if the State has set a lower level of protection for *other* Values, or if it seems fit to implement them using less-restrictive measures.

VIII. THE WILDCARDS: COMPLEMENTARITY AND QUANTITATIVE CONTRIBUTION

In the *Brazil – Tyres* dispute, Brazil's ban of foreign re-treaded tyres (other than those from MERCOSUR countries) was purportedly aimed at securing a better level of health protection.[69] The AB hung to the WAB test, describing it as 'a holistic operation that involves putting all the

[67] In the *Asbestos* case (no 44) the AB had refused to take into account the fact that the EC enforced less rigid measures with respect to other dangerous substances; in *Korea – Beef* (no 6) it looked at the less restrictive policies adopted by Korea in other sectors in the market, but allegedly only for the purpose of finding reasonable alternatives to the dual-retail system. In the *Gambling* case, instead, the Panel went further and seemed to review the US conduct in a parallel sector of the services market (namely, non-remote gambling services) in order to question the overly high level of protection of public morals pursued, see *US – Gambling*, Report of the Panel (n 51) [6.493]: '[we ought to] determine whether particular aspects associated with the remote supply of gambling and betting services will justify a prohibition, particularly in light of the tolerant attitude displayed in some parts of the United States to the non-remote supply of such services.' The Panel determined that on-line gambling entails some specific risks that could require a different regulation from the one governing non-remote gambling services (see [6.521]). Finally, in *China – Raw Materials* (see below) the Panel noticed that China's invocation of health policy objectives attached only to some of its export duties. The Panel inferred from the China's failure to invoke art XX(b) GATT with respect to other equivalent measures an adverse inference as to the genuineness of its defense, see [7.496] ff.

[68] On the obligation under art 5.5 SPS, and in particular on how this provision compares to the obligation under art XX GATT, see Michael Ming Du, 'Autonomy in Setting Appropriate Level of Protection under the WTO Law: Rhetoric or Reality?' (2010) 13 JIEL 1077, 1083 ff.

[69] At [4.11] of the Report of the Panel (n 57), Brazil mentions, *inter alia*, the risks related to cancer, dengue (and other mosquito-borne diseases), reproductive problems and environmental contamination that would be aggravated by permitting that non-reusable tyres are disposed and amassed in large landfills that might harbor mosquito colonies.

variables of the equation together and evaluating them in relation to each other after having examined them individually, in order to reach an overall judgement.'[70] In truth, the Panel diligently considered and weighed the three WAB factors: human life and health are very important (good), the ban is extremely trade-restrictive (bad), but likely to make a certain contribution towards the overall policy of disposed tyres reduction (average).

As mentioned above, there is a difference in the Value (health) and the purpose that the measure is supposed to achieve (reduction of disposed tyres). Arguably,[71] Brazil's insistence on the latter was a smart move in the LTRM perspective: interlacing health and tyre-waste reduction within one policy objective left the Panel and the claiming party with a truncated LTRM review to perform. If Brazil had declared that health protection was *the* Value, it would have been easier for the Panel to point at alternative less-restrictive measures that could ensure a similar or better result, and had nothing to do with disposed tyres. By focusing on tyre-disposal as the ultimate objective, instead, Brazil managed to limit the Panel's review to the tyre-reduction effect of the measure, drastically narrowing down the Panel's margin of discretion in looking for alternative measures.

The AB also confirmed the Panel's loose evaluation of the third factor, that is, the assessment of the contribution made by the measure in 'qualitative' (*lege*: rough) terms, rather than on the basis of quantitative measurable data.[72] This quantitative appraisal is especially likely to be justified when the contribution is not observable immediately or in the short term, or when it forms part of an aggregate contribution made by several cumulative measures.[73]

When the Panel performed the LTRM, it essentially discarded all alternative measures proposed by the EC, because they were either

[70] *Brazil – Tyres*, Report of the Appellate Body (n 5) [182].
[71] This is also the central view in Bown and Trachmann (n 14).
[72] *Brazil – Tyres*, Report of the Appellate Body (n 5) [147]. See also [210]. Note how this assessment of the rough contribution of the measure is apparently at variance with the AB's statement that a necessary measure is 'located significantly closer to the pole of 'indispensable' than to the opposite pole of simply 'making a contribution to,' in *Korea – Beef*, Report of the Appellate Body (n 6) [161]. This method to assess the contribution of the measure, as Bown and Trachmann (n 14) rightly note, is similar to the 'suitability test' advocated in the 90's by the US. However, this takes place within the virtually irrelevant WAB, so it does not substitute the LTRM.
[73] *Brazil – Tyres*, Report of the Appellate Body (n 5) [151].

unfeasible or already in place, *together* with the challenged ones.[74] It should be noted that in this case, since the assessment of the contribution was conducted without looking at its magnitude, or, as the AB put it more elegantly, 'qualitatively,' the LTRM was affected accordingly. As seen above, the WAB is a preparatory exercise, which does not substitute the LTRM test; but if the LTRM relies on the information collected in the WAB, any flaw in the latter would transmit to the former.

In *Brazil – Tyres*, the 'qualitative' assessment of the contribution of the ban to the sought objective had a double consequence. It impaired the balancing phase (non-measurable entities can hardly be weighed against each other) and affected the LTRM test, because it is impossible to look for equally-effective alternatives when the effectiveness of the original measure is not known to begin with, at least in objective terms.[75]

IX. The practice of exception-shopping, and the revival of the WAB

In this case,[76] China invoked the art XX(a) GATT exception in order to justify several measures targeting the sale and distribution of imported audiovisual products. These measures were directly or indirectly aimed at ensuring that the Chinese authorities perform some control review over the imported material. The Panel accepted the subsumption under the art XX(a) GATT, and expressed its customary praise for the policy objective and the (legitimately) high level of protection sought.[77] The claimant (the US) did not challenge this qualification, limiting itself to claim that the measures were not necessary. In so doing, it somehow conceded implicitly that Chinese censorship on foreign audiovisual materials was a perfectly legitimate policy (that only needed to be performed efficiently and non-discriminatorily), and that its exported materials could actually harm

[74] On this particular aspect, see McGrady (n 36) 155–60.
[75] See Bown and Trachtmann (n 14): 'the Appellate Body's approach also makes impossible the use of a LTIARA test, for such a test must determine equivalence of contribution, and equivalence of contribution requires assessment of magnitudes. So, in effect, the Appellate Body has now implicitly backed away not only from balancing, but also from the traditional LTIARA test'.
[76] WTO, *China – Measures Affecting Trading Rights and Distributional Services for Certain Publications and Audiovisual Products—Report of the Appellate Body* (21 December 2009) WT/DS363/AB/R. A thorough comment is provided in Joost Pauwelyn, 'Squaring Free Trade in Culture with Chinese Censorship: The WTO Appellate Body Report on *China – Audiovisuals*' (2010) 11 Melbourne J Intl L 119.
[77] WTO, *China – Audiovisuals—Report of the Panel* (12 August 2009) WT/DS363/R [7.817-818].

Chinese public morality.[78]

The Panel used the two-step analysis (WAB and LTRM), 'concluding' at first that the measures were necessary (under the WAB), then that they were not, because reasonable alternatives were available. In an attempt to clarify, the AB definitively certified the preparatory ('intermediate') role of the WAB:

the Panel's use of the word 'conclude' in setting out its *intermediate* findings risks misleading a reader, as does its characterization of certain requirements as 'necessary' before it had considered the availability of a less restrictive alternative measure.[79]

The AB also fine-tuned the *Brazil – Tyres*' opening to the 'qualitative' assessment of a measure's contribution, maybe realizing that a loose evaluation of this factor would falsify both the WAB and the LTRM tests. It recalled that the contribution is to be assessed primarily with the support of evidence and factual information, and only residually is a qualitative assessment possible.[80] Moreover, it criticized the analysis of the Panel, for relying too much on assumptions and failing to do as promised, ie assessing the 'actual contribution' of the Chinese measures to the protection of public morals.[81]

Whereas the Panel seemingly engaged in an accurate WAB test[82] and used

[78] On the unfortunate implications of this strategy, that seemed hinge upon the care with which all parties involved tried to avoid a head-on clash on the Chinese censorship regime, see Pauwelyn (n 76) 132–135.
[79] *China – Audiovisuals*, Report of the Appellate Body (n 76), [248] (emphasis added).
[80] ibid [253].
[81] ibid [294]: 'In reaching its finding regarding the contribution made by the State plan requirement to the protection of public morals in China, the Panel simply stated that limiting the number of import entities 'can make a material contribution' to the protection of public morals in China. Yet, the Panel neither addressed quantitative projections nor provided qualitative reasoning based on evidence before it to support that finding.'
[82] *China – Audiovisuals*, Report of the Panel (n 77) [7.828], [7.836], [7.863], [7.868]. For the Panel those measures imposing requisites for national importing enterprises were likely to be effective, and they did not restrict imports a priori, therefore they were legitimate. Other measures, to the contrary, were found not to be reasonably contributing to the attainment of the overall policy, and raised protectionism concerns, therefore they were reviewed more strictly. However, in light of their low trade-restrictiveness (and of the importance of the value pursued) some of them passed the necessity test, whereas others affecting the importing rate more significantly (or qualitatively, setting *a priori* prohibitions) were rejected by the Panel.

the results thereof to pronounce on the necessity of the measures, the AB reversed the analysis, and held that China failed to prove that *any* of the measures was apt to make an actual contribution. However, this did not lead the AB's report to a sudden conclusion (as one might expect: if a measure makes no contribution to the stated policy, it certainly fails under the WAB test, but it also renders the LTRM test moot). Instead, the AB entered the LTRM anyway, and confirmed the Panel's assessment (other measures were reasonably available[83]). However, the premise on the contribution was so different that it is hard to understand what the AB meant when it said that 'United States has demonstrated that the proposed alternative would ... make a contribution that is at least equivalent to the contribution made by the measures at issue to securing China's desired level of protection of public moral.'[84] It seems that the AB, like in the *Dominican Republic – Cigarettes* precedent, summoned an eighth member, Monsieur Jacques de la Palice, the only one who could subscribe without embarrassment that the measure made *no* contribution, and that accordingly any of the alternative proposals was (of course) *as effective*, or even *more*.[85]

X. MAKING SENSE OF THE *TYRES* GUIDELINES ON CONTRIBUTION

In the *China – Raw Materials* dispute,[86] several complaining parties challenged Chinese measures setting export restrictions on certain raw materials. China invoked, among other things, art XX(b) and (g) GATT

See Fontanelli, Whose Margin (n 7), 399, noting that this was the as close to a real balancing as one could hope to find in the WTO case-law.

[83] *China – Audiovisuals*, Report of the Panel (n 77) [7.898]: 'It emerges ... that implementing the US proposal would make a contribution that is at least equivalent to that of the relevant [China measures]. At the same time, the US proposal would have a significantly less restrictive impact on importers – in fact, it would have no such impact – without there being any indication that it would necessarily have a more restrictive impact on imports of relevant products than the [measures] at issue'.

[84] *China – Audiovisuals*, Report of the Appellate Body (n 76) [335].

[85] Note that the AB expressly insists that the LTRM is the dynamic combination of the values collected during the static WAB analysis, making it all the more weird, if one thinks that the AB itself had denied that the measures could made *any* contribution. See ibid., [310]: 'if a Member chooses to adopt a very restrictive measure, it will have to ensure that the measure is carefully designed so that the other elements to be taken into account in weighing and balancing the factors relevant to an assessment of the 'necessity' of the measure will 'outweigh' such restrictive effect'.

[86] WTO, *China – Measures Related to the Exportation of Various Raw Materials—Report of the Panel* (5 July 2011) WT/DS394/R, WT/DS395/R, WT/DS398/R. The AB Report was published on 30 January 2012.

(on the preservation of exhaustible resources[87]), with arguments displaying various levels of conviction and convincingness. The Panel found that these measures were adopted in violation of the China's Accession Protocol to the WTO,[88] and therefore the general exceptions of art XX GATT could not apply, since there was no reference to the GATT discipline in the applicable WTO instrument.[89] The AB later confirmed this view.[90]

However, the Panel performed the review of the measures at issue under art XX GATT, to ensure the completeness of its Report had the AB chosen to reverse the finding on the application of this provision. Leaving the art XX(g) GATT-defence aside,[91] we should focus on the Panel's reasoning on the health-related argument (which is similar to Brazil's one in *Tyres*: in essence, limiting exports of both scrap and raw materials, China would favour the transition of its industrial economy to a 'recycle' or 'circular' model, causing the increase of health protection standards that follows naturally from the adoption of an environmentally sustainable model).[92]

The claimants contended that the health-friendly description of the export duties was a mere *ex post facto* rationalization of measures that were not originally designed to protect health. The Panel upheld this complaint,[93] but decided to assess whether the measures could nevertheless make some material (although unintended) contribution to that end. The Panel concluded that the evidence submitted did not evince that the export restrictions made a material contribution to the protection of health (for one thing, because China, while highlighting the beneficial effects of said policies, omitted to account for their health-adverse effects).[94] Moreover the Panel, mindful of the *Brazil – Tyres dictum* about the 'aptness' of the

[87] *China – Raw Materials*, Report of the Panel (n 86) [7.356]: 'China's argument is that refractory-grade bauxite and fluorspar are exhaustible natural resources; they are scarce, are not easily substitutable, and thus need to be managed and protected'.
[88] See art 11, paragraph 3.
[89] *China – Raw Materials*, Report of the Panel (n 86), section VII.B.5.
[90] *China – Raw Materials*, Report of the Appellate Body (n 86) [307].
[91] Primarily, because the standard required is not one of necessity, but of 'relation to' the policy objective. Moreover, the defense failed because China did not prove to be in compliance with the even-handedness condition of art XX(g) GATT, whereby measures restricting exports must be made effective 'in conjunction' with restrictions on domestic production or consumption.
[92] *China – Raw Materials*, Report of the Panel (n 86) [7.471-472].
[93] ibid [7.516].
[94] ibid [7.538], [7.604].

measure to make some 'future contribution' to the policy objective, determined that it was not enough for China to simply claim that these measures could increase national growth and welfare, and consequently raise the level of health protection.[95] After declaring China's failure to demonstrate that the measures fell under art XX(b) GATT, like in the *EC – Tariff Preferences* case, the Panel went on *arguendo*, to prove that in any event the measures could not pass the LTRM test.

XI. CONCLUSION

The first claim of this article is that, as it emerged repeatedly in the case-law, the balancing test filters measures that would have failed the least-restrictive analysis upfront, for being both ineffective *and* significantly restrictive. No actual balancing is ever performed through the 'weighing.' The WAB is similar to the weighing-in session in boxing: fighters are weighed, but the real confrontation occurs later,[96] and somewhere else (in the LTRM ring, as it were).

The second claim is, however, that some proportionality might be spotted here and there, in the use of the LTRM routine, under the radar of the reports' reasoning. A list of these instances, without pretence to exhaustiveness, is below:

- As mentioned above,[97] sometimes the Panel takes the liberty to look into other policy areas regulated by the State, so as to get a sense of what could be an appropriate level of protection for similar Values, and whether the measure under analysis is so unusually restrictive that it might harbour a protectionist design. Obviously, when the measure is designed to achieve a relatively 'less vital' Value, it will be easier to find out that the State has in place less rigid policies regarding equivalent values.
- When the Value is human health, the 'zero risk' (or 'maximum possible enhancement') level of protection can be accepted (see *Asbestos*, *Brazil – Tyres*), whereas in connection with other Values it is routinely

[95] ibid [7.553]: 'For the Panel, even if growth makes environmental protection statistically more likely, this does not prove that export restrictions are necessary for environmental gains. For example, to the extent that a higher income per capita generates citizens' preferences for a better quality of environment, income redistribution policies may serve the environmental objective just as well as it is claimed that export restrictions do'.

[96] This brings to mind Bown and Trachtman's lament (n 14) 88: 'Yet, one might ask, if you consider these factors, but you do not compare them with each other ... how do you determine which domestic measures are acceptable and which are not?'.

[97] See above, particularly notes 67-68 and accompanying text.

toned down by the AB (see *Korea – Beef*, *Dominican Republic – Cigarettes*, *Apples I* and *Apples II*).[98] More generally, it is not unheard of that adjudicators, when ascertaining whether the less-restrictive alternative can meet the level of protection of the original measure,[99] lower the 'appropriate level of protection' predetermined by the State, so as to make the alternative eligible.[100]

- When the Value is health protection, at least in one case it was acceptable to evaluate the contribution of the measure 'qualitatively,' prospectively, and cumulatively with other policy measures (*Brazil – Tyres*). This opening was unprecedented, and was somehow shut down when, dealing with the policy objective of public morals, the AB required again that the measure be evaluated relying on objective evidence of the actual contribution (*China – Audiovisuals*).[101]

- Likewise, the qualitative approach of *Tyres* fits into the habit of relaxing the scientific boundaries of the assessment of health-related protection. In a similar vein, see how the AB held in *EC – Hormones* and *Asbestos* that governments are not obliged to base their health policies on the mainstream scientific opinion, as long as the minority views that they espouse come from 'qualified and respected sources'.[102]

Although these trends are hardly disputable, especially in their cumulative effect, their existence does not add to the predictability of the necessity

[98] WTO, *Japan – Measures Affecting the Importation of Apples—Report of the Appellate Body* (10 December 2003) WT/DS245/AB/R, and WTO, *Australia – Measures Affecting the Importation of Apples from New Zealand—Report of the Appellate Body* (29 November 2010) WT/DS367/AB/R. The *Apples* cases are not discussed in this article, as the necessity test applied therein is derived from art 5.6 of the SPS Agreement, rather than art XX GATT. A brief discussion of both disputes is provided in Filippo Fontanelli, 'When SPS applies to apples. The Japan – Apples and Australia – Apples WTO disputes' in Sabino Cassese et al. (eds), *Global Administrative Law: Cases, Materials, Issues, third edition*, (IRPA-IIJL 2012), Vol. IV, 23-29 <http://www.irpa.eu/wp-content/uploads/2012/08/The-Casebook-Chapter-4.pdf>.

[99] On the difficulty of this exercise, see WTO, *Australia – Measures Affecting Importation of Salmon - Recourse to Article 21.5 by Canada—Report of the Panel* (18 February 2000) WT/DS18/RW [7.128–7.131] (in which the LTRM was governed by art 5.6 SPS).

[100] On this, see Ming Du (n 68) especially 1097 ff.

[101] In my view, the *China – Raw Materials* Panel Report does not disprove this distinction, at least because China's demonstration about the contribution to health are *prima facie* untenable. The main problem of China's measures, with respect to the art XX(b) GATT justification, is that apparently they were not designed to pursue higher levels of health protection: they would have failed even at the rational analysis soft test advocated in the 90's by the US.

[102] WTO, *EC – Measures Concerning Meat and Meat Products – Report of the Appellate Body* (16 January 1998) WT/DS26/AB/R, WT/DS48/AB/R [194]. *EC – Asbestos*, Report of the Appellate Body (n 44) [178]. See Mavroidis, Bermann and Wu (n 66) 699-700.

test. They are not amenable to the text of art XX GATT, nor are they clearly derived from the reasoning of Panels and AB on the correct way to interpret and apply this provision. In other words, these trends are under the radar, and so are the reasons and the conditions of their operation, the 'necessity' standard of review 'enables the AB to keep maximum adjudicatory flexibility; but it leaves Members uncertain of the legality of their measures'[103] or, to put it more graphically, leaves Members and (judicious) judges 'wandering in deserts of uncharted discretion.'[104]

Weiler pointed out that the AB in *Korea – Beef* blurted out its genuine take on the WAB (a real proportionality test), only to reassume a moderate (but impenetrable) attitude, later in the same *Korea* report and in following ones.[105] This assumption of discretion, given the AB's mandate to 'complete' the WTO contract, is not pernicious *per se*. After all, it is a matter of jurisdictional allocation, and it might be acceptable that States devolve to the WTO (and to its judiciary) the competence to bring down not only discriminatory measures, but also inefficient measures, as it is normally the case under the TBT and SPS.[106] One can easily draw a comparison between the Apples cases and some of the art XX GATT cases described above. Japan and Australia's measures to prevent the slightest risk of plant disease were not discriminatory, but were disproportionate in light of the remoteness of the risk. Likewise, think of Korea's concern for commercial fraud in the meat sector, Dominican

[103] See Ming Du (n 68) 1096. In Bown and Trachtman's words: 'The result ... is so incoherent as to leave states unsure as to what types of measures may withstand scrutiny' (n 14) 88. Similarly, Kapterian (n 7) 118.

[104] *ExxonShippingCo. v. Baker*, 2008, 128 S.Ct.2605, citing Frankel (1973) 'Criminal Sentences: Law without order.' This quote is used in Fontanelli, Whose Margin (n 7), to exemplify the main claim of that work, that the margin of action that Members should be afforded has turned into a margin of adjudication in the hands of the judges, through the misuse of the necessity test.

[105] Weiler (n 43) 144. According to Ming Du (n 68) 1101: 'The AB's approach is pragmatic in the sense that it both retains de jure regulatory autonomy, but de facto allows balancing scrutiny to root out indefensible, haphazardly set risk levels'.

[106] See Trachtman (n 59) 647: 'The WTO's negative integration 'trade-off devices,' including national treatment, least-trade restrictive alternative testing and balancing testing, may be understood simply as mandates to judges to exercise discretion in the allocation of jurisdictional authority. [...] they leave much discretion to judges, they may plausibly be understood to orient and constrain judges towards, if not to, an approximation of efficiency. They do so under circumstances where it is difficult to imagine an alternative approach, other than one of positive integration. Positive integration has its own costs.' For an earlier formulation and a wider discussion of this view, see Id., 'Trade and ... Problems, Cost-Benefit Analysis and Subsidiarity' (1998) 9 EJIL 32, 82 and Id., 'Institutional Linkage: Transcending 'Trade and ...' – an Institutional Perspective' (2002) 96 AJIL 77.

Republic's apprehension about illegal border transactions, China's alleged interest in monitoring cultural material that could threat its cultural identity: the implementation of these Values did not necessarily result in discriminatory measures, but their impact on trade was disproportionate, and WTO DSM bodies used an augmented LTRM test to strike them off.

The LTRM test, being narrowly devoted to ensure Pareto optimization,[107] fails to represent an open and flexible test for the evaluation of policies, therefore it is understandable that some deal of reasonableness and good governance[108] finds its way in the reasoning of the Panels and AB.
However, the haphazard accumulation of redundant and wearisome tests related to the necessity requirement of art XX(a) (b) and (d) GATT does not seem the optimal way to ensure that a bit of reasonableness underpins the Reports of the Dispute Settlement Body. As things stand now, Panels and AB are more likely to appear activist rather than reasonable when they soften the LTRM test: maybe it is time to dust the WAB and start embracing, very cautiously, a bit of proportionality *proprement dite*.

In sum, it is fair to note that the mandate of WTO quasi-judicial bodies is such that no real proportionality can control the outcome of a case. [109] This is visible in the truncated WAB (where the first factor is never really weighed), and in the obstinate use of the LTRM. There is some subterranean 'constitutional' trend, traceable in a 'loose' use of the LTRM and the statistical evidence showing that certain values and 'more Values' than the others.

[107] According to Trachtman, Trade and... (n 106) 72, it can be overbroad and under-inclusive at the same time: '[n]ecessity testing engages in truncated maximization, or truncated comparative cost-benefit analysis, by keeping the regulatory benefit relatively constant and working on the trade detriment side. It thus evaluates a much more limited range of options, ignoring other groups of options that may be superior'.
[108] Lang (n 42) 325.
[109] Nor could the AB perform this constitutional test. See Lang (n 42) 320 ff; P. Van den Bossche (n 42) 283; Jan Neumann and Elisabeth Türk, 'Necessity Revisited: Proportionality in World Trade Organization Law after Korea – Beef, EC – Asbestos and EC – Sardines' (2003) 37 JWT 199, 214, 233, and bibliography referred to therein. For a definition of the narrow proportionality test, see Trachtman, Trade and... (n 106) 35, and bibliography referred to therein.

SOCIAL DIALOGUE, *LAVAL*-STYLE

Alessandra Asteriti*

The article provides an analysis of the Laval Judgment in light of Habermas' theory of discursive practice and compares the European social model and the Swedish system of collective agreements in light of this theory. In this context, the article argues, the comprehensive dismissal by the European Court of the carefully constructed and balanced Swedish system of social dialogue between management and labour is truly the most disturbing aspect of this controversial judgment. For all the supposed importance placed on discursive practices and social dialogue for the European social model, when confronted with the Swedish system of social dialogue, the Court retreats in the familiar territory of hard law and statutory obligations. In doing so, it wilfully misunderstands the function of collective bargaining, by effectively decoupling its process from its function, and leaving social dialogue with the hollow role of a deliberative practice devoid of any finality.

TABLE OF CONTENTS

I. INTRODUCTION 70
II. COLLECTIVE BARGAINING, DELIBERATION AND SOCIAL DIALOGUE 71
1. *The Theory – Jürgen Habermas* 72
2. *The Practice at EU Level: The European Social Dialogue* 74
III. THE SWEDISH SOCIAL MODEL 77
IV. THE LAVAL CASE 79
1. *The Facts* 80
2. *The Law* 80
V. THE DECISION OF THE COURT 83
1. *The Background* 83
2. *The Judgment of the Court* 86
VI. THE AFTERMATH 92
VII. CONCLUDING REMARKS 97

* University of Glasgow. E-mail: asteriti.alessandra@gmail.com. I am very grateful to Maria Fletcher and Charles Woolfson for their comments on an earlier draft of this article and to the anonymous reviewers for their attentive and helpful review. The usual disclaimer applies.

I. INTRODUCTION

Ever since the ECJ delivered its *Laval* judgment on 18 December 2007, the name of this small Latvian company has become notorious. The sole mention of '*Viking* and *Laval*' has become short-hand for those critical of a certain idea of Europe giving primacy to economic considerations to the detriment of 'social Europe'. This article intends to go back to the original *Laval* judgment to reconstruct its history and deconstruct its myth.

The *Viking* and *Laval* judgments have been criticised for using freedom of establishment and freedom to provide services respectively as 'trumps' against the fundamental right of freedom of association and collective action[1]. What protection for the right to strike after what the Court decided, one was inclined to ask? Were we going to see social dumping become the norm, a race to the bottom that would see Eastern European workers compete against their Western counterparts by offering their low labour cost as their best asset? These are crucial questions and they have justly been discussed extensively elsewhere. This article will only consider the *Laval* judgment, and will explore a different angle, by taking as its starting point Habermas' theory of discursive practices as guarantees for a democratic outcome and offering the Swedish system of collective agreements as a substantiation of such practices. In this context, the article will argue, the comprehensive dismissal by the Court of the carefully constructed and balanced system of social dialogue between management and labour is truly the most disturbing aspect of this controversial judgment. For all the supposed importance placed on discursive practices and social dialogue for the European social model, when confronted with a successful example of such model, the Court retreated in the familiar territory of hard law and statutory obligations. In doing so, it wilfully misunderstood the function of collective bargaining, by effectively decoupling its process from its function, and leaving social dialogue with the hollow role of a deliberative practice devoid of any finality, the very openness of which both signifies and nullifies its democratic credentials.

The article is structured as follows: Part I provides the theoretical grounding for the argument, by considering how discursive practices have

[1] More comprehensively, *Laval* belongs to a 'quartet' of cases decided in rapid succession by the ECJ along similar lines, comprised of C-341/05 *Laval un Partneri* [2007] ECR I-11767; C-438/05 *The International Transport Workers' Federation and The Finnish Seamen's Union* [2207] ECR I-10779; C-346/06 *Rüffert* [2008] ECR I-1989; and C-319/06 *Commission vs Luxembourg* [2008] ECR I-4323, hereafter referred to as *Laval, Viking, Rüffert* and *Luxembourg*.

influenced our conception of democracy, including in the area of industrial relations. Part II focuses on the development of social dialogue in European law, from its introduction in the Single European Act in 1986 to arts 154-155 TFEU. Part III considers how the Swedish social model puts in practice the theory of social dialogue in its system of collective agreements, heavily dependent on deliberative practices between management and labour, and with a minimal statutory framework. Part IV summarises the facts of the case brought by the Latvian company Laval un Partneri against the Swedish building and public works trade union. Part V analyses the decision of the Court, concentrating on the value judgment made by the Court of the system of collective agreements described above and its continued viability following Directive 96/71 (the Posted Workers Directive). Finally, Part VI considers the aftermath of this decision at the national level, with the passing of the 'Laval Law' by the Swedish government in 2010, and at the European level, with the issuing by the Commission of a draft new Directive on the Enforcement of Directive on Posted Workers in March 2012, before offering some concluding remarks.

II. COLLECTIVE BARGAINING, DELIBERATION AND SOCIAL DIALOGUE

According to John Dunlop's system of industrial relations[2], a tripartite structure, including workers, employers and the State, is engaged in a framework of collective bargaining where the speaking positions reflect opposing, and to a certain extent, irreconcilable viewpoints, and the goal is to reach a compromise where all partners engage in the discussion using the 'weapons' at their disposal; in the case of the workers or employees, this is the tool of the withdrawal of labour, or the threat of industrial action.

In contrast, deliberation as a form of discursive practice in Habermasian terms, or 'civil dialogue', can be conceptualised as a more open framework, where a consensus can be reached by actors engaged in the dialogue in a non-confrontational form, 'through exchanges of arguments accepted as valid by the participants in the public debate'[3].

Social dialogue, defined as the 'institutionalised consultation procedure involving the European social partners, [or also] the processes between

[2] John Dunlop, *Industrial Relations Systems* (Harvard Business School Press 1993).
[3] Claude Didry and Annette Jobert, 'Deliberation: a New Dimension in European Industrial Relations', in Jean De Munck and others, *Renewing Democratic Deliberation in Europe, The Challenge of Social and Civil Dialogue* (Peter Lang 2012) 171, 171.

social partners at various levels of industrial relations[4], seems to sit uneasily between these two extremes, sharing elements of both, and, seemingly, failing to provide any of the benefits of its two more established predecessors. Its openness to deliberative processes is coupled with an integrative thrust to the given of the Single Market project to the exclusion of alternative paths, in a misguided effort to accommodate the 'social' to the reality of the market, never the other way around. This logic of commitment, through social dialogue, to a predetermined outcome risks undermining any advantages conferred by the deliberative ethos to the bargaining process[5].

The following table summarises the differences between the frameworks. Attention is called particularly to the difference in mode, means, goals and outcome; in these, social dialogue distinguishes itself by its *soft* nature, with reference to the lack of bindingness of the outcome as well as the way in which deliberation is structured; its *targeted deliberation*, in the sense that the dialogue is not 'free-flowing' but channelled through approved paths and rigidly constructed 'givens'; and its *predetermined consensus*, because the rupture of the framework is not an option, as exemplified by the *outcome* category, where the potential conflict is neither defused, nor resolved, but simply *denied*.

Table I

FRAME-WORK	PARTICI-PANTS	MODE	MEANS	GOALS	OUTCOME
Collective bargaining	Employers Workers State	Rigid	Industrial action	Compromise	Conflict defusal and deferral
Civil dialogue	All affected parties	Flexible	Deliberation	Freely obtained consensus	Conflict resolution
Social dialogue	Tripartite or Bipartite	'Soft'	Targeted deliberation	Predetermined consensus	Conflict denial

1. *The Theory – Jürgen Habermas*

The influence of Jürgen Habermas in democratic theory cannot be underestimated. His communicative model provides the testing ground

[4] From the Eurofound website, <www.eurofound.europa.eu/>, accessed 12 October 2012.

[5] The point is made also by Ruth Dukes and Emilios Christodoulidis, 'Habermas and the European Social Dialogue: Deliberative Democracy as Industrial Democracy?'(2012) 18(4) Industrial LJ 21.

and legitimisation tool for normative statements in a democratic context[6]. His co-originality theory of private and public autonomy, whereby rights and democracy are seen as reflexively underpinning each other, is in itself dependent on a working framework where discursive practices involve all participants under ideal speech conditions[7]. In these, the openness of the discourse guarantees a democratic outcome and the inclusiveness of the participation results in the development of what he calls the 'social perspective of the first-person plural'[8], in which all affected persons are given a stake in the result of the dialogue and at the same time, bind themselves to that result. Arguably, the bindingness of the result constitutes the problematic element in the model, being more prone to capture. However, in the case of industrial relations, where the expressed *telos* of bargaining between the social partners is defusal or deferral of the conflict by means of a binding agreement, this bindingness is organic to the system[9], and it is other elements upon which one should concentrate the critical attention, and these are the procedural guarantees and the substantive rights within that procedural framework. Indeed, it is important to note that, in the context of industrial relations, it is crucial not to lose the capacity of the social partners to create binding agreements, and furthermore, not to lose the bargaining tools that allow that bindingness to be established (in the case of workers, the right to undertake industrial action).

Habermas' model of participatory democracy is predicated on three essential elements: an effective framework, equal speaking positions for all participants – effective participation – and openness of outcome. It is not the place here to comprehensively critique the viability of this model, when faced with the relentlessness of predetermined structures and their power to close down possibilities, which is at its strongest precisely in a functioning democratic framework, as counterintuitive as this might seem. Rather, this brief introduction to Habermas' discursive practices theory serves to illustrate the convergence between this theory and the practice in the Swedish model of industrial relations. To my knowledge, this convergence has not been noted before, which is particularly surprising

[6] Jürgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (William Rehg tr, Polity Press 1996). Also, more recently, Jürgen Habermas, *Between Naturalism and Religion* (Ciaran Cronin tr, Polity Press 2008).
[7] ibid 118 ff, especially 122; also 106: 'Just those action norms are valid to which all possibly affected persons could agree as participants in rational discourses'.
[8] ibid 92.
[9] As part of the recognition of the fact that the conflict is organic to the model, and therefore necessarily present.

given the contractual nature of his co-originality thesis[10]. On the contrary, the influence of his theories on the development of the European Social Dialogue (ESD) is evident and has received much attention[11].

2. *The Practice at EU Level: The European Social Dialogue*

The history of the creation of 'social Europe' as a project in parallel to 'economic Europe' has been interpreted as an exemplification of Karl Polanyi's 'social embeddedness theory', whereby 'the initial decoupling of the economic from the social economic constitution in the design of the integration project and the later strive for competitiveness through the "completion" of the internal market programme can be interpreted as disembedding moves [which] .... provoke countermoves directed at a re-embedding of the market'[12].

Others have remarked on the 'dysfunctional relationship' between the European Social Model and the Single Market project[13]; regardless of how ownership of the social is interpreted (as an internal move by the market to pre-empt disruption, or as a genuine countermove, still subject to the risk of appropriation by the market), the development of the European Social Dialogue took place precisely when the social model and the market model came to confront each other in what seemed like a case of binary and irreversible choice.

---

[10] As explicitly stated by Habermas (n 6) 122, referring to '[...] a horizontal association of free and equal persons [...] *prior to* any legally organized state authority from whose encroachments citizens would have to protect themselves.'

[11] An interesting theoretical approach to ESD, especially in light of the well known controversy between Habermas and Luhmann on societal structures (which started following their joint work in Niklas Luhmann and Jürgen Habermas, *Theory of Society or Social Technology: What Does Systems Research Accomplish?* [Suhrkamp 1971]) is by Christian Welz, *The European Social Dialogue under Article 138 and 139 of the EC Treaty* (Kluwer Law International 2008). In it, Welz adopts Luhmann's and Teubner's theories in order to argue for ESD to be understood as an autopoietic subsystem of the European Union.

[12] Christian Joerges and Florian Rödl, 'On De-formalisation in European Politics and Formalism in European Jurisprudence in Response to the "Social Deficit" of the European Integration Project – Reflections after the Judgments of the ECJ in *Viking* and *Laval*', (2008) 4(1) Hanse Law Review 3; the reference is obviously to Karl Polanyi's seminal work *The Great Transformation: The Political and Economic Origins of Our Time* (Beacon Press 1944).

[13] John Foster, 'The Single Market and Employment Rights: From a Dysfunctional to an Abusive Relationship?', Institute of Employment Rights Conference, 21 March 2012, *Developments in European Labour Law*. Thanks to Professor Charles Woolfson for having brought this contribution to my attention.

It was at the *Val Duchesse* talks, organised by the Delors Commission in 1985, that the ESD between employers and trade unions was launched, under the auspices of the Commission[14]. Shortly thereafter, the ESD was given statutory presence by art 21 of the Single European Act[15], which amended the EEC Treaty via the addition of art 118(a) and 118(b). Art 118(a) established the possibility of adopting directives by Qualified Majority Voting[16]; art 118(b) recited as follows:

The Commission shall endeavour to develop the dialogue between management and labour at European level which could, if the two sides consider it desirable, lead to relations based on agreement.

The policy was expanded and embedded further in the Treaties with the Protocol on Social Policy annexed to the Maastricht Treaty[17], where the ESD is mentioned in art 1 (programmatic article) and arts 3 and 4:

> Article 3
> 1. The Commission shall have the task of promoting the consultation of management and labour at Community level and shall take any relevant measure to facilitate their dialogue by ensuring balanced support for the parties.
> 2. To this end, before submitting proposals in the social policy field, the Commission
> shall consult management and labour on the possible direction of Community action.
> 3. If, after such consultation, the Commission considers Community action advisable, it shall consult management and labour on the content of the envisaged proposal. Management and labour shall forward to the Commission an opinion or, where appropriate, a recommendation.
> 4. On the occasion of such consultation, management and labour may inform the Commission of their wish to initiate the process provided for in Article 4. The duration of the procedure shall not

[14] Information on the Val Duchesse process is available on Eurofound website <www.eurofound.europa.eu> accessed 12 October 2012.
[15] Single European Act [1987] OJ L 169/1.
[16] 'Article 118A (now Article 137(1) EC) was inserted by the Single European Act, which allowed for qualified majority voting for proposals 'encouraging improvements, especially in the working environment, as regards the health and safety of workers'. From the Eurofound website <www.eurofound.europa.eu/> accessed 12 October 2012.
[17] Treaty on European Union Protocol on Social Policy [1992] OJ C191/1.

> exceed nine months, unless the management and labour concerned and the Commission decide jointly to extend it.

Article 4
1. Should management and labour so desire, the dialogue between them at Community level may lead to contractual relations, including agreements.
2. Agreements concluded at Community level shall be implemented either in accordance with the procedures and practices specific to management and labour and the Member States or, in matters covered by Article 2, at the joint request of the signatory parties, by a Council decision on a proposal from the Commission.
The Council shall act by qualified majority, except where the agreement in question contains one or more provisions relating to one of the areas referred to in Article 2(3), in which case it shall act unanimously.

These arts became incorporated in the EC Treaty as arts 138 and 139[18] and are now arts 154 and 155 TFEU[19].

Since its inception in 1985, the ESD has created a substantial amount of literature dedicated both to its initiatives, including its successes and failures, and to critical reflections and analysis, both country-specific and more general in approach[20]. Issues of process and result intersect with domestic patterns of industrial relations, raising several questions: what is ESD for, and how is it supposed to interact with national models? Are we confronted with substitution, where ESD comes to replace industrial relations conducted at the domestic level, or validate them at the European level, or something else? This article is premised on the assumption that the Swedish model of industrial relations successfully accomplishes what Habermas envisioned as the function of discursive practices in guaranteeing a democratic outcome in the shadow of the law. The assumption holds, one would like to think, if there is a balance between the democratic nature of the discursive practices and the framing and the binding provided by the law. Crucial for this balance is that the discursive practices cannot just be a procedural value, but have to have substantive content, and that this has to be reflexively present at practice

[18] Treaty Establishing the European Community [2006] OJ C321 E/5.
[19] Treaty on the Functioning of the European Union [2010] OJ C83/49.
[20] The output is considerable: 'institutional' information is available from several EU websites, such as <www.eurofound.europa.eu>, and <http://ec.europa.eu> accessed 12 October 2012; a recent study by the Policy Department of the DG for Internal Policies was published in 2011, *Cross-border Collective Bargaining and Transnational Social Dialogue*, IP/A/EMPL/ST/2010-06; see also Welz (n 12) and Jean De Munck and others (n 3).

level: in other words, the social partners have to have the power to determine the content of the binding rules, and they have to be aware of this power[21]. For now, it will suffice to note that, with all its limitations, in the Swedish model this substantive reflexive power is conferred on the social partners.

III. THE SWEDISH SOCIAL MODEL

The Swedish model of industrial relations is based on a collective agreements framework with a robust procedural structure and extensive powers granted to the social partners to come to collective decisions as to their substantive rights and obligations under private law contracts, with minimal legislative involvement[22]. The *telos* of this model is exemplified by the absence of a law on minimum wage in Sweden, since the rate of pay is agreed within the collective agreements negotiated by the employers and trade unions at the sectoral level. The main statutory provisions are contained in the 1974 Employment Protection Act (LAS) and the 1976 Co-determination Act (MBL)[23], and include the obligation to maintain 'industrial peace' when a collective agreement is entered into (s 41 MBL). These pieces of 1970s legislation have been seen as an attempt to crystallise in statutory form (and therefore constitutionalise) certain substantive and procedural advantages for unions, while maintaining the traditional system of collective negotiated agreements[24]. Collective agreements are applicable to trade union members directly (in Sweden about 70% of workers belong to a trade union and 90% of working relationships are covered by a collective agreement[25]) and indirectly

[21] I am here conflating authorship, as intended by Habermas (n 6) 120, and power in the sense of creative legislative power.

[22] For a historical review, see Ole Hasselbach, 'The Roots – the History of Nordic Labour Law', (2002) 43 Scandinavian Studies in Law 11.

[23] The *Lag om Anställningsskydd* (*Official Gazette* 1982:80) and the *Medbestämmandelagen* (*Official Gazette* 1976:580).

[24] Hasselbalch (n 23) 32. A divergent look at the history of social relations in Sweden by Svante Nycander, with more emphasis placed on the shift from a model of 'collective *laissez faire*' as described by Otto Kahn-Freund (who believed Sweden to be the most accomplished example of this model), accompanied by the 'spirit of *Saltsjöbaden',* to a much more interventionist and State-led policy, exemplified by the legislative activity of the 1970s. See Svante Nycander, 'Misunderstanding the Swedish Model', in *Collective Bargaining, Discrimination, Social Security and European Integration*: *Papers & Proceedings of the 7th European Regional Congress of the International Society for Labour Law and Social Security Law, Stockholm, September 2002* (Kluwer Law International 2002) 437.

[25] As sources for the data, see *The Swedish Model – The Importance of Collective Agreements in Sweden*, leaflet produced by the Swedish Trade Union Confederation (LO) <www.lo.se> accessed 12 October 2012; and the Report produced by the

through subsidiary agreements to non-unionised workers and employees. Additionally, the social partners enter into basic agreements establishing the procedural rules to be followed in the negotiations; these are modelled on the *Saltsjöbaden Agreement*, signed in 1938 between the then Swedish Employers' Association and the largest Swedish Trade Union Confederation, LO, and still applicable to most negotiated agreements. The law gives the trade unions exclusive powers to conclude agreements and a powerful negotiating tool in the constitutional protection granted to the right to engage in industrial action[26]. Once an agreement is reached, there is, as noted, an obligation on the parties to a social truce. This obligation is given statutory strength in Section 42 MBL:

Employers' or workers' associations shall not be entitled to organise or encourage illegal collective action in any way whatsoever. Nor shall they be entitled to participate in any illegal collective action by providing support or in any other way. An association which is itself bound by a collective agreement shall also, in the event of a collective action which its members are preparing to take or are taking, seek to prevent such action or help to bring it to an end.
If any illegal action is taken, third parties shall be prohibited from participating in it.

The Swedish Labour Court (*Arbetsdomstolen*) interpreted para 1 of Section 42 to apply also to industrial action taken in Sweden against foreign undertakings; the judgment[27] concerned a company that owned a ship, *M/S*

Swedish Government, *Action in Response to the Laval Judgment – Summary,* Swedish Government Official Reports, SOU 2008:123. The coverage for the building sector is even higher, with collective agreements covering 96% of workers, of which 77% belonged to a trade union. All data refer to the year 2007. However reassuring or frankly enviable these numbers might seem from a UK perspective, there has been a downward trend, from a high of 85% in the early 1990s to the current numbers, as reported by Charles Woolfson, Christer Thörnqvist and Jeffrey Sommers, "The Swedish Model and the Future of Labour Standards after Laval', (2010) 41(4) Industrial Relations J 333.

[26] Ch 2, s 17 of the Swedish Instrument of Government (*Regeringsformen*, the Swedish Constitution): 'A trade union or an employer or employers' association shall be entitled to take industrial action unless otherwise provided in an Act of law or under an agreement'; the Swedish Labour Court (*Arbetsdomstolen*), which acts as a court of last instance for industrial disputes (except where the Court sentence is alleged to be a grave violations of fundamental rights and where recourse to the Constitutional Court might be allowed), has interpreted this provision to apply horizontally and to entail civil liability (*civilrättslig verkan*). Constitutional protection for trade union rights, including the right to industrial action, is not unusual in Europe, as noted by AG Mengozzi in his *Laval* Opinion, paras 31-33.

[27] Britannia Case AD 1989, No 120.

Britannia, flying a flag of convenience and employing a Filipino crew covered by a collective agreement under Filipino law; the interpretation became known as the *Britannia Principle*. As a consequence of this judgment, the Swedish government immediately approved a legislative amendment to the MBL[28], adding three paragraphs, including a third paragraph to Section 42, to the effect that: 'The provisions of the first two sentences of the first paragraph shall apply only if an association takes collective action by reason of the terms and conditions of employment falling directly within the scope of the present law'[29]. The amendment excluded industrial action against employers having concluded agreements out with Swedish law, effectively allowing industrial action against foreign employers and employers of posted workers covered by collective agreements under the home state law. The amendments became in turn known as *Lex Britannia,* devised by the Swedish parliament as a way to counter the risk of social dumping[30]. This is the same stated purpose of the Posted Workers Directive[31]: the crucial difference is that, in the Swedish model, the social partners, and specifically, the trade unions, are *entrusted* with the tools necessary to avoid social dumping and maintain fair competition in the Swedish labour market.

IV. THE LAVAL CASE

[28] Official Gazette 1991:681, Government Bill 1990/91: 162.

[29] The other two amendments stipulate that a foreign collective agreement that is invalid under foreign law is valid under Swedish law if it complies with the MBL (s 25a) and that later collective agreements will trump an earlier collective agreement that does not comply with the MBL (s 31a).

[30] As stated in the Government Bill, 5ff.; of particular relevance, in light of the proportionality analysis performed by the ECJ to the detriment of the collective right of industrial action against the individual right of provision of services, the report stated that: 'This regulation [*lex Britannia*] is based on the idea that employment relationships which in no way fall within the scope of the MBL, cannot, reciprocally, be given the special protection it provides. *The starting point must be, rather, the constitutional rules on the freedom and the right to take industrial action*' [emphasis added]. See also Ronnie Eklund, 'A Swedish Perspective on *Laval*', (2008) 29 Comparative Labour L and Policy J 551, 554. Social dumping can be defined as '[the] practice involving the export of goods from a country with weak or poorly enforced labour standards, where the exporter's costs are artificially lower than its competitors in countries with higher standards, hence representing an unfair advantage in international trade.' (Eurofound website <www.eurofound.europa.eu> accessed 12 October 2012).

[31] See Preamble of Directive 96/71/EC of the European Parliament and of the Council of 16 December 1996 concerning the posting of workers in the framework of the provision of services.

1. *The Facts*

The facts of the case are well known. Laval un Partneri Ltd (Laval) won a contract for the renovation of a school in Vaxholm, Sweden, through its fully owned Swedish subsidiary L&P Baltic Bygg AB (Baltic). Between May and December 2004, Laval posted 35 Latvian workers to work on the project. In June 2004, Byggettan[32] started negotiations with Laval and Baltic with the intention of entering into a collective agreement for the posted workers. Following the beginning of negotiations, Laval entered into an agreement with all its posted workers[33]. In November 2004, with the negotiations stalling, Byggettan started industrial action against Laval, by blockading the construction site. In December 2004 a conciliation hearing was held at the *Arbetsdomstolen*, in which Laval refused a final offer by Byggettan and requested an interim injunction to stop the industrial action, claiming that it was in violation of arts 12 and 49 EC. The request was refused by the *Arbetsdomstolen* on 22 December 2004. The hearing on the merits took place on 11 March 2005; in it, Laval petitioned the *Arbetsdomstolen* to request a preliminary ruling from the European Court of Justice (the "ECJ" or "the Court") under art 234 EC, in addition to demanding damages from Byggnads and Elektrikerna for a total of SEK 600,000[34].

2. *The Law*

The Swedish legislative framework has been reviewed in Part 3; this section contains a review of the EU and international law applicable to the decision by the Court. The necessary historical background to the applicability of EU legislation to Swedish labour disputes is certainly the position that Sweden took with respect to its own model of social relations when negotiating its accession to the EC in 1994. At the time, Sweden appended a declaration to its accession protocol, to the effect that 'In an exchange of letters between the Kingdom of Sweden and the Commission, [...] the Kingdom of Sweden received assurances with regard to Swedish practice in labour market matters and notably the system of determining

[32] The three unions involved in the case were Svenska Byggnadsarbetareförbundet (the Swedish building and public works trade union, "Byggnads"); Svenska Byggnadsarbetareförbundet avdelning1, Byggettan (the local branch of Byggnads, "Byggettan"); and Svenska Elektrikerförbundet (the Swedish electricians' trade union, "Elektrikerna")

[33] Byggnads claimed that this agreement was no more than 'a device for the Company to try to avoid signing a Swedish collective agreement' (*Arbetsdomstolen* Judgment no. 49/05, Case no A268/04, Byggnads' submission to the Court).

[34] 55,000 GBP at July 2012 exchange rate.

condition of work in collective agreements between the social partners'[35]. Equally, the *Lex Britannia* already discussed in pt 3 engendered a reaction at the international level, following the Swedish employers' organisation's claim that this law breached ILO's Conventions C.87, C.98 and P.147[36]. This claim was rejected both by the Swedish government at the time and eventually by the ILO Committee of Experts (CEACR)[37]; consequently, the ILO reaffirmed the compliance of Swedish labour legislation with internationally-agreed standards.

The 'Posted Workers Directive' was adopted by the Social Affairs Council on 24 September 1996 with the contrary vote of only Portugal and the UK [38]; the Swedish parliament adopted the relevant implementing legislation in May 1999, to the exclusion of collective agreements on pay, as per domestic labour policy[39]. Specifically, Section 5 of the Act contains the provisions on the conditions of employment, as per art 3(1) of the Directive, which covers the minimum rates of pay at 3(1)(c).

In a way, the Directive departs from the international private law rules on the applicability of employment contracts for temporarily deployed workers as stipulated in art 6 of the 1980 Rome Convention on the law applicable to contractual obligations[40], which states that the laws of the

---

[35] Declaration No. 46 by the Kingdom of Sweden on social policy, annexed to the Accession Act of Austria, Finland, Norway and Sweden, OJ C241, 29.8.1994.

[36] Freedom of Association and Protection of the Right to Organise (adopted 9 July 1948, entered into force 4 July 1950) 68 UNTS 18; Right to Organise and Collective Bargaining (adopted 1 July 1949, entered into force 18 July 1951) 96 UNTS 258; and Protocol of 1996 to the Minimum Standards of Merchant Shipping Convention (adopted 22 October 1996, entered into force 10 January 2003) 2206 UNTS 106.

[37] The complaint was initiated by Swedish representative for employers Johan von Holten at the ILO conference in 1991; the complaint was rejected both by the Swedish government, which distanced itself from it, and by the CEACR; the information is taken from the LO website, <www.lo.se/home/lo/home.nsf/unidView/.../$file/waxholm.pdf> accessed 12 October 2012.

[38] (n 32) Since the Directive was adopted according to Qualified Majority Procedure under art189b EC, there was no power of veto available to the UK and Portugal. The choice of legal base, current arts 53 and 62 TFEU, was made precisely to avoid the necessity of a unanimous vote in the Council; see Paul Davies, 'Posted Workers: Single Market or Protection of National Labour Law Systems' (1997) 34 Common Market L Rev 571.

[39] Act on the Posting of Workers, *Official Gazette* 1999:678, Government Bill 1998/99:90.

[40] Council 80/934/EEC: Convention on the law applicable to contractual obligations opened for signature in Rome on 19 June 1980 / Consolidated version CF 498Y0126(03), OJ L266/1.

country of origin (home country) apply to the employment relationship[41]. Instead, in order to avoid social dumping and guarantee fair competition in the labour market, the Directive adopts the device of a 'core' of labour guarantees ('a nucleus of mandatory rules for minimum protection'), listed at art 3, as 'laid down by law, regulation or administrative provision, and/or by collective agreement or arbitration awards which have been declared universally applicable'. According to para 8 of the article, this refers to 'collective agreements or arbitration awards which must be observed by all undertakings in the geographical area and in the profession of industry concerned'. Where there is no system for collective agreements of universal application (as is the case in Sweden) the Directive allows for 'collective agreements or arbitration awards which are generally applicable to all similar undertakings in the geographical area and in the profession or industry concerned, and/or collective agreements which have been concluded by the most representative employers' and labour organizations at national level and which are applied throughout national territory', with the proviso that their application will guarantee equal treatment to the undertakings involved.

The Directive was not applicable to the dispute between Laval and the three trade unions, as a consequence of the fact that Directives do not have horizontal direct effect and so cannot be relied upon in a dispute between private parties or create rights and obligations directly enforceable by national courts or by the ECJ[42]. However, this does not prevent the Court from taking directives into consideration when examining a case, and this the Court did do extensively in its Judgment, nor does it exempt national courts from interpreting their national laws in conformity with EU law, including Directives, therefore ensuring their

[41] As noted also by Advocate General Mengozzi in his *Laval* Opinion, para 132[which case????]. Conversely, the ECJ had already established in Case C-113/89 *Rush Portuguesa Limitada v Office National d'Immigration* [1990] ECR I-1417, that 'Community law does not preclude Member States from extending their legislation, or collective labour agreements entered into by both sides of industry, to any person who is employed, even temporarily, within their territory [...]'.

[42] The applicability of dirs is not as clear cut as the general rule seems to imply; the ECJ has pronounced numerous times on their effect; see mainly Case 41/74 *Van Duyn v Home Office* [1974] *ECR* 1337; Case 148/78 *Pubblico Ministero v Tullio Ratti* [1979] ECR 1629; Case 14/83 *Von Colson and Kamann v Land Nordrhein-Westfalen* [1984] ECR 1891; Case 152/84 *Marshall v Southampton and South-West Hampshire Area Health Authority (Teaching)* [1986] ECR 723; Case C–188/89 *Foster and Others v British Gas plc* [1990] ECR1-3313; Case C–106/89 *Marleasing SA v La Comercial Internacionale de Alimentacion SA* [1990] ECR 1-4135; Case C–201/02 *The Queen, on the application of Delena Wells v Secretary of State for Transport, Local Government and the Regions* [2004] ECR 1- 723; Cases C–397–403/01 *Pfeiffer and Others v Deutsches Rotes Kreuz, Kreisverband Waldshut eV* [2004] ECR 1-8835.

indirect effect.

Laval argued that the MBL, and specifically Section 42(3), was in breach of EU law by discriminating against foreign undertakings and by unlawfully violating the freedom to provide services protected under art 49 EC. Both in the case of the Directive and in the case of art 49, restrictions are allowed either for public policy reasons, or for the protection of a legitimate interest. In both cases, the Court did not accept that the right to engage in collective bargaining between private parties could be affected by a public policy exception, because of the lack of involvement of the State in the Swedish model of industrial relations[43]; nor did they accept that the protection of legitimate interests justified the restrictions imposed by the MBL on the freedom to provide services, judging it disproportionate to attain its scope. The Court set the bar extremely low in its standard of review of the proportionality of the action, by stating that,

> '[...] the right of trade unions [...] to take collective action [....] is liable to make it *less attractive,* or more difficult, for such undertakings to carry out construction work in Sweden, and therefore constitutes a restriction on the freedom to provide services within the meaning of Article 49 EC' [emphasis added][44].

V. THE DECISION OF THE COURT

1. *The Background*

The background to the Judgment of the Court is crucial to understanding its outcome. Three elements are worth mentioning: the decision of the *Arbetsdomstolen* to request a preliminary ruling; the opinion given by Advocate General Mengozzi[45]; and the judgment issued by the ECJ only one week previously in *Viking*[46].

Laval had claimed in its submission to the *Arbetsdomstolen* that the industrial action was unlawful under Section 42(1) of the MBL; additionally, it had claimed that Section 42(3) (the *Lex Britannia* amendment) constituted a violation of the principle of non-discrimination

[43] *Laval*, para 84.
[44] On the application of proportionality in the context of collective bargaining, see Brian Bercusson, 'The Trade Union Movement and the European Union: Judgment Day' (2007) 13(3) Eur L J 279, 304.
[45] Delivered on 23 May 2007.
[46] The two cases were joined and the judgment on *Viking* was issued on 11 December 2007.

on grounds of nationality, protected under art 12 EC, and of the freedom to provide services under art 49 EC and to post workers under the Posted Workers Directive. For its part, Byggnads claimed that, since the right to take industrial action is not regulated at Community level[47], national governments retained competence in this area, as reiterated in Recital 22 of the Posted Workers Directive. Without prejudice to this, they also claimed that restrictions of art 49 can be justified if undertaken in the public interest (such as measures taken for the protection of employees and to avoid social dumping)[48]. The union claimed that Laval workers had been paid SEK 20-35 per hour and made to work 56 hours per week, in contrast with the union's request of an hourly wage of SEK 145, with a fall-back rate of SEK 109 in case of lack of agreement by the parties[49].

The *Arbetsdomstolen* accepted that the industrial action undertaken by Byggnads was unlawful under Section 42(1) of the MBL; it held however that, Section 42(3) of the same Act being applicable, the industrial action was therefore lawful under Swedish law. On the question of Community law, it accepted the request of a preliminary ruling from the ECJ advanced by the Company in order to clarify the lawfulness of the industrial action under arts 12 and 49 EC and under the Posted Workers Directive. The Company had argued that the Court had competence, notwithstanding art 137(5) EC, to decide the dispute insofar as, first, the industrial action constituted a disproportionate and unlawful restriction of a fundamental freedom and, second, when national law is in conflict with Community law, the latter one takes precedence.

The *Arbetsdomstolen* therefore referred the dispute to the ECJ, seeking clarification on the following two points: 'the issue of the compatibility of the industrial action with the rules on free movement of services and the prohibition against discrimination on the ground of nationality; and, 'the conditions under which legal rules which in practice discriminate against foreign companies carrying out activities temporarily in Sweden with labour from their own country [*lex Britannia*], are compatible with the rules on free movement of services and prohibition against discrimination on grounds of nationality.'

Subsequent to the request for the preliminary ruling, Advocate General

[47] As expressly stated in art 137(5) EC.
[48] As established by the ECJ in Case C-164/99 *R. v. Portugaia Construções Limitada* [2002] ECR I-787, para 19.
[49] According to interviews granted under condition of anonymity by Laval workers and published by *Byggnadsarbetaren* magazine <www.byggnadsarbetaren.se/> accessed 12 October 2012.

(AG) Paolo Mengozzi delivered his Opinion on 23 May 2007. This is not analysed in detail in this article; it is worth noting however that markedly different approach taken by the AG in his analysis of the Swedish model of industrial relations and the weight that this is attributed in drafting the Opinion. To this effect, it will suffice to provide two quotes: the first one is from para 61 of the Opinion, in the Preliminary Observations (Legal Analysis section), where AG Mengozzi states that:

... if the application of the freedoms of movement provided for by the Treaty, in this case the freedom to provide services, were to undermine the very substance of the right to resort to collective action, which is protected as a fundamental right, such application might be regarded as unlawful, even if it pursued an objective in the general interest.

It is clear from this quote that AG Mengozzi does not take as his starting point the presumption that the right to collective action is to be intended as a possible restriction to a fundamental freedom, and therefore has to be proportionate in order to be lawful, which is the approach taken by the Court. Rather, he opines that the fundamental right against which possible restrictions have to be assessed for proportionality is the collective right to industrial action. The approach of the Court is of course dictated by the case as presented, since the Court is asked by the claimant to decide on a breach of the freedom protected in art 49; however it is noticeable that AG Mengozzi seems to at least entertain a possible *categorical approach* to the question posed, where the right to industrial action is found to fall 'outside the scope of the freedoms of movement'[50], rather than the *balancing approach* used by the ECJ, where inevitably one of the two rights is seen to *cut into* the other one and the role of the Court is to assess the proportionality of this infringement.

It is well known that AG Mengozzi concluded that art 49 did not preclude industrial action to force a foreign employer to accept a collective agreement guaranteeing better conditions for the posted workers, provided the collective action was motivated by public interest goals (inclusive of the prevention of social dumping). I would like to point out another aspect of his Opinion, and specifically his more sophisticated and nuanced understanding of the Swedish social model, as exemplified by para 260:

However, those circumstances [unforeseeable results when entering the negotiations, or excessive wage claims] are inherent in a system of

[50] Para 60. This categorical approach is ultimately rejected in favour of a balancing exercise, paras 78ff of the Opinion.

collective employment relations *which is based on and favours negotiation between both sides of industry* and, therefore, contractual freedom, rather than intervention by the national legislature [emphasis added]. I do not think that, at its present stage of development, Community law can encroach upon that approach to employment relationships through the application of one of the fundamental freedoms of movement provided for in the Treaty.

It is my argument, and I am not alone in this, that this encroachment was precisely the strategic decision undertaken by the Court and further, that in order to execute this strategy, the Court had to wilfully disregard that very system of social relations even while upholding the rhetoric of social dialogue[51].

Finally, the *Laval* Judgment has to be read in the context of the developing jurisprudence of the Court on the right of collective action, and specifically, *Viking*. As stated in the Introduction, it is not my intention to compare the two cases[52], and even less, to use them as symbols. But it is nonetheless important to note that the Court did overstep its own mark in delivering the *Laval* Judgment, by arrogating to itself the task of establishing the proportionality of the interference with the fundamental freedom involved, a task that it had left to the national court in *Viking*[53].

2. *The Judgment of the Court*

Many elements of the *Laval* Judgment have created a considerable amount of debate. To start from where we ended in the previous section, the proportionality analysis performed by the Court has been criticised, as downgrading the fundamental right of collective action and representation

[51] See for example para 105 of the Judgment.

[52] Both cases have been analysed extensively, including by way of comparison; see Joerges and Rödl (n12); Alain Supiot, 'L'Europe gagnée par « l'économie communiste de marché' Revue du MAUSS permanente (30 janvier 2008), <www.journaldumauss.net/spip.php?article283> accessed 12 October 2012; Norbert Reich, 'Free Movement v. Social Rights in an Enlarged Union – the *Laval* and *Viking* Cases before the ECJ' (2008) 9(2) German Law Journal 125; Rebecca Zahn, 'The *Viking* and *Laval* Cases in the Context of European Enlargement' (2008) 3 Web Journal of Current Legal Issues <http://webjcli.ncl.ac.uk/2008/issue3/zahn3.html> accessed 12 October 2012; Roger Blanpain and Andrzej Świątkowski (eds.) *The Laval and Viking Cases: Freedom of Services and Establishment v. Industrial Conflict in the European Economic Area and Russia* (Kluwer Law International 2009).

[53] At para 87 of its Judgment, the Court stated: 'As regards the question of whether or not the collective action at issue in the main proceedings goes beyond what is necessary to achieve the object pursued, it is for the national court to examine...'

to the exercise of the freedom of movement protected by the Treaty[54]. Equally, the horizontal application of art 49, and the consequent liability of the trade unions for breaches of EU law, has been investigated extensively. *Laval, Viking, Rüffert* and *Luxembourg* have been taken as an authoritative view of the Court on the status of fundamental rights, at least pre-Lisbon, against the four freedoms, and the exemplification of the economic model defended by the Court against social policies, at the European and national level.

This article is investigating the significance of this Judgment through the prism of the discourse of social dialogue at the European level, and how this interacts and cuts across national policies on industrial relations, taking the example of the Swedish system as the one that the Court itself adjudicates upon. To this effect, particular attention will be paid to the language adopted by the Court in explaining the rationale for its decision. A couple of preliminary points need to be made: the first one is the determination of the Court to focus its analysis on the interpretation of the Directive on Posted Workers, which could not be relied upon by Laval in its claim in the Swedish courts. This approach has been 'puzzling' for many authors, but explained by the wish of the Court to 'express its views on the role and interpretation of the Directive'[55]; arguably, more is at play here, because effectively, the Directive is used to give substance to the general principle protected by art 49 (freedom to provide services). Second, it has been suggested that the Court transformed the 'floor' provided by the Directive in its nucleus of minimum requirements to a 'ceiling' by making them into the maximum standards instead[56]; to this, it is important to add that this is accomplished by effectively tying the principle of freedom of establishment to the specific criteria listed in the Directive, even while defending in principle the sovereign right of States to apply more generous criteria.

The very framing of the Court's decision to the exclusion of any meaningful engagement with the particularity of the Swedish system of industrial relations is evident by the way in which the Court rearticulates

[54] See for example Bercusson (n 43).
[55] ACL Davies, 'One Step Forward, Two Steps Back? The *Viking* and *Laval* Cases in the ECJ', (2008) 37(2) Industrial Law Journal 126, 127.
[56] Study produced by the EU Parliament, Employment and Social Affairs Department, DG Internal Policies, 'The Impact of the ECJ Judgments on *Viking, Laval, Rüffert* and *Luxembourg* on the Practice of Collective Bargaining and the Effectiveness of Social Action', IP/A/EMPL/ST/2009-11 (May 2010) 7.

the first question posed by the *Arbetsdomstolen* wilfully changing its scope[57]. The *Arbetsdomstolen* had posed the question in these terms:

> Is it compatible with EU rules [...] for trade unions to attempt, by means of collective action, to force a foreign provider of services to sign a collective agreement in the host country [...] if the legislation to implement Directive 96/71 has no express provisions concerning the application of terms and conditions of employment in collective agreements?'[58]

The Court rephrased the question as follows:

> The national court's first question must be understood as asking [...] whether Articles 12 EC and 49 EC, and Directive 96/71, are to be interpreted as precluding a trade union [...] from attempting, by means of collective action in the form of blockading sites [...] to force a provider of services established in another Member State to enter into negotiations with it on the rates of pay for posted workers, and to sign a collective agreement, the terms of which lay down [....] *more favourable conditions* than those resulting from the relevant legislative provisions....'[59].

With this reframing, the Court shifts the focus of the question from the issue of non-discrimination, to that of the freedom to provide services unencumbered by national legislation protective of social rights.

When adopting the implementing legislation for the Posted Workers Directive, the Swedish government had dealt with the requirement of art 3(1) of the Directive not by means of collective agreements applicable *erga omnes* but through the possibility provided by art 3(8)(2), justifying its approach in the following terms:

Legislating to require posting employers to comply with the applicable collective agreement without creating discrimination against them as compared to Swedish employers who are not required by law so to do would mean that there is actually only way to avoid a declaration of the universal applicability of collective agreements. That is for the legal text to have approximately the same wording as the Directive, namely that posting employers must comply with collective agreements to the same

[57] As Joergens and Rödl (n12), put it at 16, n 61, 'The Court simply ignores Swedish policy'.
[58] *Laval,* para 41.
[59] *Laval,* para 53.

extent that Swedish companies in a similar situation do. This would entail always needing to make a comparison of each individual case. Such a solution would obviously seem alien to the Swedish tradition[60].

In other words, the Posted Workers Directive and its implementation could not be used to determine two different categories of collective agreements under Swedish law, and this was the same rationale underpinning the *Lex Britannia*. However, the intention of the Court is to internationalise collective agreements, and at the same time to deprive the unions of their power to use industrial action as a negotiating tool for anything above the minimum level guaranteed by the Directive. And to do so, it rephrases the question so as to make its focus the *more favourable conditions*, rather than the technical issue raised by the *Arbetsdomstolen* with respect to the applicability of art 3(8) when the law of the host state does not allow for the applicability of collective agreements *erga omnes*.

Once rephrased in the above fashion, it is not difficult for the Court to further its argument on the basis that forcing more favourable conditions is not allowed by the Directive, which only protects the *voluntary* decision by the social partners to enter into more favourable conditions of employment with respect to posted workers. This is a typical move, where the diversity of the speaking positions is masked by the apparent equality of choice. So the Court can state both that Recital 17 of the Directive holds, which states that '[...] the mandatory rules for minimum protection in force in the host country must not prevent the application of terms and conditions of employment which are more favourable to workers' – as well as Recital 22, '[...] this Directive is without prejudice to the law of the Member States concerning collective action to defend the interests of trades and professions' – and contextually decide that 'Article 3(7) of Directive 96/71[61] cannot be interpreted as allowing the host Member State to make the provision of services in its territory conditional on the observance of terms and conditions of employment which go beyond the mandatory rules for minimum protection'. It seems irrelevant to the Court that the Swedish system does nothing of the sort in its legislation, leaving the matter to the social partners[62]. This, it seems, is a freedom too far for the Court. What then remains of the right of industrial action if it can be exercised only to obtain the observance of the minimum standards already

[60] As cited by the *Arbetsdomstolen* in its Judgment (n 33) 31.
[61] Which states that '[p]aragraphs 1 to 6 [on minimum requirements] shall not prevent application of terms and conditions of employment which are more favourable to workers.'
[62] With the proviso that the horizontal application of the *acquis* might require the social partners to be subject to the same rules tying Member States.

guaranteed by the Directive or through its mechanisms?

After effectively depriving the right of collective action of its main function, the Court moves on to question more widely the Swedish system of social dialogue. Let us remind ourselves that, in Habermasian terms, the equality of speaking positions is crucial, and the Court has already dispensed with that. Equally essential is the openness of outcome. AG Mengozzi had remarked, as we had seen, that this is a structural, physiological and unavoidable element of collective bargaining[63]. If the outcome is predetermined, what is the value of the dialogue? This framing condition depends organically on the 'equality of arms', in the sense that both social partners are equally exposed to the openness of the outcome: the worker as well as the employer enter into pay negotiations in Sweden without certainty of outcome, except two, very important provisos: the rate of pay is supposed to reflect the general rate of pay applicable for a similar job in the same geographical area, and, if an agreement is not reached, the fall-back rate will be applicable (which is probably lower than the employees wish to get and higher than the employers want to pay)[64]. In another blow to meaningful social dialogue, the Court asserts that:

> '[...] collective action [...] cannot be justified in the light of the public interest objective [...] where the negotiations on pay [...] form part of a national context characterised by a lack of provisions, of any kind, which are *sufficiently precise and accessible* that they do not render it impossible or excessively difficult in practice for such an undertaking to determine the obligations with which it is required to comply as regards minimum pay[emphasis added]'.

Apart from being a misrepresentation of the Swedish model, because of the two conditions outlined above on pay negotiations, one cannot help but despair for the complete and wilful misunderstanding of bargaining and dialogue in conditions of democracy. The openness *is* the virtue of the system, not its vice[65].

---

[63] Para 260 of his Opinion.

[64] Eklund (n 30) 552.

[65] ibid 551. By doing so the Court is exposing, maybe unwittingly, the hypocrisy of the rhetoric of 'flexicurity' at EU level; see for example, *Towards Common Principles of Flexicurity: more and better jobs through flexibility and security* - Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, adopted on 27 June 2007, COM (2007) 359. Flexicurity, for the Commission, has to include 'flexible and reliable contractual arrangements' (at 20); without irony, the documents notes that: 'Active involvement of social partners is key to ensure that flexicurity delivers benefits for all. It is also essential that all stakeholders involved are prepared to

And so on to the third element of an effective social dialogue in Habermasian terms, the framework for the dialogue to take place under conditions of equality (procedural equality, as opposed to the substantive equality of speaking positions discussed above). With this, in its answer to the second question posed by the *Arbetsdomstolen*, the Court returns to the issue of discrimination. The framing for social dialogue is in Sweden guaranteed by the MBL, in its post-*lex Britannia* incarnation, designed to guarantee an equal framework for domestic and foreign undertakings with respect to the right to engage in industrial action.

The Court however interprets this rule to the effect that 'collective action is authorised against undertakings bound by a collective agreement subject to the law of another Member State in the same way as such action is authorised against undertakings which are not bound by any collective agreement'[66] and finds consequently the rule to be unjustly discriminatory (by equating domestic undertakings that have not entered into a collective agreement with foreign undertakings covered by a foreign collective agreement). This is only half the story; as we know from amended Sections 42(3), 25(a) and 31(a) MBL, the rule only applies to collective agreements that violate the MBL. In any case, the *Arbetsdomstolen* clearly stated that 'the industrial action would have been lawful if the Company had been a Swedish company' so that it is neither a question of 'circumvention' nor of 'special treatment'. Furthermore, the *Arbetsdomstolen* clarifies the scope of the MBL amendment to the effect that, since the MBL guarantees a 'social truce' under conditions of respect of the legislation, this privilege cannot be extended to foreign undertakings that do not otherwise respect its provisions. In other words, the Court subverts the very rationale of Section 42 MBL, to guarantee social peace provided negotiations are entered in good faith and within the umbrella (the procedural framing) of the MBL, into a prohibition to engage in industrial action. Stripped of the crucial framing, all that remains, for the Court, is the prohibition to strike once a collective agreement (any collective agreement) is entered into[67]. So set adrift from its supporting legislation, the prohibition stands in for the opposite of what it was intended to be, i.e., a consequence of the collective agreement binding in compliance with Swedish law, not a free-standing right to be protected from industrial action and from any duty to engage in

---

accept and take responsibility for change. Integrated flexicurity policies are often found in countries where the dialogue – and above all the trust - between social partners, and between social partners and public authorities, has played an important role' (at 18).

[66] Para 113.

[67] This is apparently the rational underlying the new Lex Laval, see below.

social dialogue.

VI. The Aftermath

The *Laval* Judgment's repercussions were felt at the political level in Sweden, with new legislation being passed; at the domestic legal level, with the Judgment by the *Arbetsdomstolen*[68]; and finally at the EU level, with negotiations on an amended directive on the posting of workers and the initiation of a complaint procedure by the LO and the Swedish Confederation of Professional Employees ("TCO") against Sweden to the European Committee of Social Rights[69].

As a consequence of the Judgment of the Court, a Committee was appointed by the Swedish government in 2008 [at the time, a centre-right coalition] in order to ascertain what legislative action should be taken, in the form of amendments to the *Lex Britannia* and the Posting of Workers Act; in its report, the Committee clarified the provisions of art 3(8)(2) indent one, whereby, in the absence of *erga omnes* application of collective agreements or arbitration awards, 'Member States may [...] base themselves on collective agreements or arbitration awards which are generally applicable to all similar undertakings in the geographical area and in the profession or industry concerned [...]'. As for the necessary compliance with art 49, the report suggested that the right to strike in order to determine the employment conditions of posted workers could be retained under the following conditions: the disputed terms and conditions of employment must correspond to the conditions contained in a collective agreement which complies with art 3(8)(2); the terms and conditions must '[fall] within the "hard core" of the Posting of Workers Directive' (with the proviso that, as concerns minimum rates of pay, it should be left to the trade unions to determine what constitutes said rate, to the inclusion of overtime etc.); the burden of proof that the condition of employment of the posted workers are equivalent to the conditions demanded by the trade unions rests with the posting employer. Other

[68] For the domestic repercussions, see in general Mia Rönnmar, '*Laval* returns to Sweden: The Final Judgment of the Swedish Labour Court and Swedish Legislative Reforms' (2010) 39(3) Industrial Law Journal 280; for a very good , and critical, review of the Labour Court's judgment in light of EU law, see Elisa Saccà, 'Nuovi scenari nazionali del caso *Laval*. L'ordinamento svedese tra responsabilità per danno "da sciopero" e innovazioni legislative (indotte)', *Working Papers* Centro Studi di Diritto del Lavoro Europeo "Massimo D'Antona" 86/2010.

[69] Complaint No. 85/2012, registered on 27 June 2012, available on the Committee's website,
<www.coe.int/t/dghl/monitoring/socialcharter/Complaints/Complaints_en.asp>
accessed 12 October 2012.

procedural amendments were also proposed to improve transparency and communication, always respectful of the principle, at least on paper, that 'the social partners will assume responsibility for [the] proposed regulations satisfying the requirements of Community law'[70]. The new *Lex Laval*, adopted on 15 April 2010 by the Swedish Parliament, in addition to accepting the proposals of the Commission, qualifies the right to resort to industrial action accordingly, by stating that 'An employees' organisation may not use industrial action to achieve a Swedish collective agreement if an employer can show that the employees are already included in terms and conditions (regardless if stipulated by collective agreement, employment contract or managerial decision) that are at least as good as those in a Swedish central branch agreement.'[71] The short paragraph reveals a subtle but fundamental shift from a dialogic model of industrial relations to a situation in which all the partners have to do (and in this case, crucially, the employer) is to show that the working conditions are comparable to the terms agreed at a local level. Not surprisingly, the amendment was immediately criticised by the LO and a request was made for the ILO to examine its compliance with the conventions on the right to union membership and collective negotiations[72]. The Committee in its 2010 Report refers to the case in the following terms: '[...] the omnipresent threat of an action for damages that could bankrupt the union, possible now in the light of the *Viking* and *Laval* judgements, creates a situation where the rights under the Convention cannot be exercised[73]. This of course raises the question of a possible normative conflict between the obligations arising under the *acquis communitaire* and Sweden's (and the other EU countries) international obligations under the ILO Conventions. Equally, the spectre of fragmentation and normative dissonance has been raised with respect of the jurisprudence of the

[70] Report of the Swedish Government, note 25, 35.

[71] Prop 2009/10:48. The new bill amends the Posting of Workers Act by adding Section 5a. Additionally, the *Lex Britannia* could not be applied to any undertaking posting workers to Sweden, including those from outside the EU. An exception of constitutionality was argued for the proposal by the opposition parties, but rejected by the Supreme Court (see Rönnmar (n 67) 286).

[72] For a summary of the LO's objections to the Committee, see the text of the Complaint submitted to the European Committee of Social Rights, note 68, 20). The Committee pronounced on the cases in its Report of the Committee of Experts on the Application of Conventions and Recommendations (2010) ilolex nr 062010GBR087; the Report is available on the ILO website, <www.ilo.org/> accessed 12 October 2012; see also Kerstin Ahlberg, 'Will the Lex Laval work?' Nordic Labour Journal, 2 November 2010.

[73] At 209; this statement was in response to a request raised by the British Airline Pilots' Association (BALPA).

European Court of Human Rights on the right of association, which is going in an opposite direction to the ECJ's stance in the *Laval* quartet[74].
The repercussions extended at the domestic legal level, with a new judgment by the *Arbetsdomstolen*. As a consequence of the preliminary ruling by the ECJ, Laval raised its demand for damages to three million SEK[75], while the trade unions argued that there should not be liability for damages resting on the trade unions, as the breach of EU law was attributable to the Swedish State[76], and in any case the trade unions' action was legal in Swedish law at the time it was taken, questioning the retroactive application of the ECJ's ruling to a dispute between private parties in order to establish civil liability. The *Arbetsdomstolen* disagreed on both grounds (liability for damages under EU law, for violation of art 49[77], and under Swedish law, for breach of the MBL[78]), and with the minimum majority required (four judges out of seven) established that the unions were liable, establishing the amount at 700,000 SEK in punitive damages and two million SEK in litigations costs[79]. As noted previously, there is no right of appeal from the *Arbetsdomstolen*, safe for miscarriage of justice resulting from an 'obvious' and 'grave' mistake in law[80]. This the trade unions have done, requesting a ruling from the Supreme Court[81]; the

[74] *Demir and Baykara v Turkey,* App no 34503/97 (ECtHR 12 November 2008); and *Enerji Yapi-Yol Sen v Turkey*, App no 68959/01 (ECtHR 21 April 2009) reaffirming that the right to strike and collective bargaining is protected under art 11 of the Convention. See Keith Ewing and John Hendy, 'The Dramatic Implications of Demir and Baykara' (2010) 39(1) Industrial Law Journal 2.
[75] 283,000 GBP.
[76] Ss 54 and 55 MBL provide the rules on liability for breaches of the MBL; the rules were applied by analogy by the Court to assess the damages for the breach of EU law.
[77] Following the case law of the ECJ on horizontal direct effect in the area of competition law, eg Case C-453/99 *Courage Ltd v Bernard Crehan and Bernard Crehan v Courage Ltd and Others* [2001] ECR I-6297 and applying the criteria for Member State liability established in *Brasserie du Pêcheur SA v Bundesrepublik Deutschland and The Queen v Secretary of State for Transport, ex parte: Factortame Ltd and others,* joined Cases C-46/93 and C-48/93 [1996] ECR I-1029. Even in the absence of any precedent on liability for damages for breaches of art 49 by private parties (horizontal direct effect), the *Arbetsdomstolen* did not think it necessary to request a preliminary ruling from the ECJ on this point, even if the parties had request so at the time of the first request for a preliminary ruling in 2005.
[78] As the ECJ had established in its ruling the *Lex Britannia* to be unlawful under art 49 EC, the *Arbetsdomstolen* was bound to apply the *Britannia Principle* instead, under which the industrial action was found to be unlawful, with again ss 54 and 55 of the MBL applicable for establishing liability and punitive damages.
[79] Niklaas Bruun, 'The Laval case, act III – Sweden's Labour Court rules union must pay high damages, Nordic Labour Journal, 12 January 2010. The *Arbetsdomstolen* delivered its judgment on 2 December 2009 (*Arbetsdomstolen* AD 2009:89).
[80] Note 26.
[81] Kerstin Ahlbeg, 'Swedish unions want annulment of Laval judgment', Nordic Labour Journal, 18 May 2010.

Supreme Court, maybe predictably, refused their request[82]. Calls were made for the Swedish State to pay the damages[83], but they were ultimately paid by the Swedish trade unions to the administrator of the company, Laval having declared bankruptcy. It may be superfluous, in this context, to remark on the chilling effect of the *Arbetsdomstolen* judgment on the right of trade unions to resort to industrial action, given the extension of liability for action deemed legal by the *Arbetsdomstolen* itself at the time it was taken. Suffice to notice that the *Arbetsdomstolen* imposed punitive damages on the trade unions for having failed to predict that their action would have fallen foul of EU law, when the *Arbetsdomstolen* itself was not certain that this was the case, so much so that it refused the demand for an injunction by Laval and it requested a preliminary ruling from the ECJ on that very question. In fact, the *Arbetsdomstolen* was able to impose damages under Swedish law only by disapplying the *Lex Britannia*, which was the object of the second question posed to the ECJ.

The repercussion at the European level include the joint report produced by the European Social Partners at the invitation of the European Commission[84], which highlights the chasm between the partners on the assessment of the consequences of the ECJ rulings, with the employers' representative favourably commenting on the interpretation of the ECJ being 'helpful to avoid uncertainty [...] and to assure a ground of fair competition; on their part, ETUC remarked that 'the argument of "legal certainty" cannot be used as an excuse to interfere with the essential

[82] Kerstin Ahlbeg, 'Curtain fall for the Laval case', Nordic Labour Journal, 31 August 2010.

[83] The decision to assign liability for punitive damages to the unions for having exercised their right to resort to industrial action disregards the primary characteristic of this right, which is the immunity from civil liability (taking into account that the action was legal under Swedish law, as recognised by the same court in its 2005 judgment). See Tonia Novitz, 'Labour Rights as Human Rights: Implications for Employers' Free Movement in an Enlarged European Union', in Catherine Barnard (ed), (2007) 9 Cambridge Yearbook of European Law 357; Filip Dorssemont, 'The Right to take Collective Action Versus Fundamental Economic Freedoms in the Aftermath of Laval and Viking: Foes are Forever!', in Marc De Vos and Catherine Barnard (eds.), *European Union Internal Market and Labour Law: Friends or Foes?,* (Intersentia 2009) 45. The *Arbetsdomstolen* could have interpreted EU law so as to exclude or limit liability for individuals because the unlawfulness of the action is only the first step to establish liability, and it is not quite clear that damages would have been granted as a matter of EU law; additionally, the Court could have applied s 60 of the MBL, which allows to reduce or waive damages if deemed reasonable under the circumstances.

[84] *Report on joint work of the European social partners on the ECJ rulings in the Viking, Laval, Rüffert and Luxembourg cases*, of 19 March 2010, available on the European Trade Union Confederation ("ETUC") website <www.etuc.org/> accessed 12 October 2012.

features of national labour law and industrial relations systems' and concluded that 'The sustainability of industrial relations has been threatened.'

On a legislative level, negotiations have been ongoing on a Directive on the enforcement of the Posted Workers Directive, first suggested by President Barroso in 2009; an amendment proposed by the Employment and Social Affairs Committee of the European Parliament under the Ordinary Legislative Procedure is currently awaiting its first reading. A parallel proposal for a Council Regulation under the Consent Procedure on the exercise of the right to take collective action within the context of the freedom of establishment and the freedom to provide services is at preparatory phase in Parliament[85].

Finally, as mentioned above, the Swedish LO and TOC have submitted a complaint to the European Committee of Social Rights against Sweden, requesting that the Committee pronounce on their claim that the *Lex Laval* violates Sweden's obligations with respect to arts 4, 6 and 19.4 of the European Social Charter[86] and ILO C.98 (art 4) and C.154[87]. The complaint has just been lodged and there is no way of knowing how it will be assessed. It is interesting to note how the trade unions turn the ECJ's argument about the lack of clarity of the pay agreements on its head, by remarking that the new legislative framework makes it impossible for trade unions to predict if their industrial action will be deemed lawful, or if it will attract punitive damages and they conclude with the following gloomy prediction, worth quoting in full:

---

[85] See information on the website <www.europarl.europa.eu/committees/en/empl/work-in-progress.html#menuzone> accessed 12 October 2012. The ETUC has already issued a position on the proposed Directive, which requests major revisions to address its shortcomings, outlined in the position paper, see <www.etuc.org/a/10037> accessed 12 October 2012. The proposed Council Regulation, on its part, has resulted, in July 2012, in the first 'yellow card' from national parliaments (including the Swedish one), under art 7.2 of Protocol 2 to the Lisbon Treaty; as per procedure, the draft regulation will now have to be reviewed by the Commission, but there is no legal obligation of amendment or withdrawal (<http://extranet.cor.europa.eu/subsidiarity/news/Pages/Early-Warning-System-First---yellow-card--.aspx> accessed 12 October 2012).

[86] European Social Charter (opened for signature 18 October 1961, entered into force 26 February 1965) CETS No. 035; the articles concern the right to a fair remuneration, the right to bargain collectively, and the right of migrant workers and their families to protection and assistance (non-discrimination in remuneration, working conditions and employment matters, including trade union participation).

[87] For C.98 see n 36; ILO Convention 154, Promoting Collective Bargaining (adopted 3 June 1981, entered into force 11 August 1983) 1331 UNTS 268.

The combination of the new rules on industrial peace and full financial tort liability without a negligence requirement has led to great wariness on the part of the trade union organisations as regards signing collective agreements with foreign employers. The fear felt by the trade union organisations of doing the wrong thing by mistake and putting the organisation at risk of being forced to pay high levels of damages has meant that there has been a severe fall in the number of collective agreements signed as regards foreign companies carrying on business in Sweden. This means that foreign workers are entirely without protection as regards reasonable terms and conditions of pay and employment when they are working in the Swedish labour market and that Swedish workers are exposed to competition from workers with very low pay and wretched employment conditions. In the long term there is a risk that this will have negative repercussions for the entire Swedish labour market model'[88].

VII. CONCLUDING REMARKS

It is important to improve working conditions and wages in competing countries, in order to raise the floor[89].

The above quote is from an interview with a Swedish trade unionist on the subject of the ESD at sectoral level; when uttered, the Court had not delivered its *Laval* Judgment, with its well-known transformation of the floor provided by the Posted Workers Directive into a ceiling of what is obtainable through industrial action. One is left to wonder what this trade unionist would make of the sleight of hand by the Court.

The reverberations of this Judgment go well beyond the low numbers involved, as is often the case: 35 Latvian workers involved in the actual dispute, and a total of posted workers in Sweden at the time of the dispute estimated to be at about 2,200, inclusive of about 1,050 in the building sector[90]. Charles Woolfson has rightly noted that Latvia might have used the dispute, and Laval instrumentally, in order to 'prise open new markets in the EU' and certainly the facts of the dispute, and especially its political

[88] Complaint to the European Committee of Social Rights (n 68) 28.
[89] Interview with a Swedish trade unionist on the topic of European Sectoral Social Dialogue, as reported by Sofia Murhem, 'Implementation of the Sectoral Social Dialogue in Sweden', in Anne Dufresne, Christophe Degryse and Philippe Pochet (eds.), *The European Sectoral Social Dialogue. Actors, Developments and Challenges* (Peter Lang 2006) 281, 292.
[90] The numbers are taken from the Eurofound website, <www.eurofound.europa.eu/> accessed 12 October 2012; the website offers a clear disclaimer on the accuracy of the figure, as no official data is collected. Similar statistics, collected by the unions, are provided in the Swedish Government Official Report, note 25, 10.

background, give support to this hypothesis[91]. To this, further strategic considerations can be added, that have to do more with *prising open* the legal structure of industrial relations that Swedish workers have developed in cooperation with capital over almost a century. This strategy of legal disruption has domestic and European elements. It is public knowledge that Laval was supported financially, in bringing its case in the *Arbetsdomstolen*, by the Confederation of Swedish Enterprise, who had, as once again Woolfson noted, 'long argued in favour of reducing the scope of trade union industrial action, especially with regard to sympathy action affecting so-called "third parties"' [92]. The legislative changes at the domestic level have been investigated in depth throughout this article and bear out the impression that it is more than the destiny of a limited number of foreign workers posted in Sweden to be the concern and the real target of the statutory intervention, as pointedly noted by the trade unions in the closing paragraph of their complaint reported above.

At the European level, it will be useful to remind the reader that Swedish law already contained, in Sections 54 and 55 MBL, the imposition of economic *and* punitive damages for breach of the social peace after the conclusion of a collective agreement. In the strictly private law relationship established between the parties, the Swedish State intervenes to punish the unions for unlawful resort to strike action under limited conditions. In a double move of *deracination* or de- and re-localisation, the ECJ and the Swedish Labour Court (applying EU law, or maybe misapplying it) have localised to Sweden EU law by embedding the restrictions of the Posted Workers Directive and giving it the force of hard law in a model of industrial relations predicated upon 'soft' methods of dialogic exchange, and Europeanised the Swedish imposition of punitive damages for unlawful industrial action, which is inconsistent with any other model of labour law that does not include the substantive guarantees under a meaningful social dialogue that the Swedish model provides[93]. In

[91] Charles Woolfson and Jeff Sommers, 'Labour Mobility in Construction: European Implications of the Laval un Partneri Dispute with Swedish Labour' (2006) 12(1) European Journal of Industrial Relations 49, 56. Some interesting background on the political dimension of the dispute in Alban Davesne, 'The Laval Case and the Future of Labour Relations in Europe', Les Cahiers Européens de Sciences Po. No 01/2009, Paris: Centre d'études Européennes.

[92] Woolfson and others (n 26) 15. The quote refers to information provided in Jan Peter Duker, 'Ett arbetsgivarperspektiv på medling', in Anne-Marie Egerö and Birgitta Nyström (eds.), *Hundra år av medling I Sverige: Jubileumsskrift: Historik, analys och framtidsvisioner* (Medlinginstitutet 2006) 184.

[93] Europeanisation is not intended here in the traditional sense of 'an incremental process of re-orienting the direction and shape of politics to the extent that EC [EU] political and economic dynamics become part of the organisational logic of national

doing so they set adrift the Swedish prohibition to strike under penalty of punitive damages and allowed it 'to strike' in quite a different way, as the members of the British Association of Airline Pilots ("BALPA") found out when their employer British Airways decided to seek an injunction in the English courts on the basis of the unlawfulness of a proposed strike action, and to seek punitive damages to the order of 100 million GBP per day were the strike to take place[94].

In its complaint to the European Committee of Social rights[95] the LO puts it succinctly but clearly: '[....] in these cases[96] *collective agreement free zones* are created in the Swedish labour market, where it is only possible to conclude a collective agreement if the employer accepts it voluntarily' [emphasis added][97]. Similarly to the Export Processing Zones ("EPZ") that are a common feature in developing countries[98] these zones signify the transformation of the Western labour market in the direction of a *de-westernisation* and they do so by depriving the trade unions of their speaking position and reducing them to passive listeners.

---

politics and policy making' (see Robert Ladrech, 'Europeanisation of Domestic Politics and Institutions: The Case of France' (1994) 32 Journal of Common Market Studies 69, 69 as quoted in Zahn, note 51); rather it is intended to convey the concept of a policy or law deracinated, that is, taken out of its local context, and re-embedded both in its national context, but having been decontextualised, and in a new international or, in this case, European context.

[94] *BALPA called off the strike but submitted a complaint to the ILO Committee of Experts, see supra note 71 and appealed against the interim injunction granted by McCombe J dated 17 May 2010. For the appeal in the courts see British Airways PLC v Unite The Union [2010] EWCA Civ 669.*

[95] Note 68, 22.

[96] Where, by application of the new Section 5(a)(2) of the Posting of Workers Act, collective action is not lawful if the employer has shown that the minimum conditions of employment are respected, even without a binding agreement.

[97] This voluntariness constitutes the Habermasian element of the Swedish system only insofar as it does not extend to the meta-level of the framework; in other words, it has to be accepted by the social partners that the voluntariness does not include the possibility *not* to enter into a dialogue at all, and to impose labour conditions derived from exogenous sources (such as the hard statutory provisions that form the object of the LO criticism).

[98] Information on the EPZs at <www.ilo.org/public/english/support/lib/resource/subject/epz.htm> accessed 12 October 2012.

THE APPLICATION OF ARTICLE 103 OF THE UNITED NATIONS CHARTER IN THE EUROPEAN COURTS: THE QUEST FOR REGIME COMPATIBILITY ON FUNDAMENTAL RIGHTS

Kushtrim Istrefi*

This contribution identifies and examines three approaches of the European courts to application of certain United Nations Security Council (SC) resolutions that trigger concerns for violations of fundamental rights in their respective legal orders. This analysis argues that a balance between the United Nations (UN) superior norm and preservation of fundamental rights should be aimed outside the monist and dualist constraints of interpretation, where the courts either obey art 103 of the UN Charter trumping fundamental rights (subordination approach), or detach from the UN system in order to safeguard fundamental rights of their autonomous regimes (detachment approach). This submission suggests that further exploration of norms and techniques of treaty interpretation found in the Al-Jedda and Nada cases of the European Court of Human Rights (ECtHR), coupled with constructive contribution of scholars provide tools allowing regime compatibility and harmonization that disturb neither coherence nor autonomy of the respective regimes (harmonization approach) in the world of legal plurality.

TABLE OF CONTENTS

I. OPENING REMARKS: SETTING THE SCENE 100
II. SUBORDINATION APPROACH: WHEN COHERENCE BECOMES THE ANTONYM OF FUNDAMENTAL RIGHTS 104
III. DETACHMENT FROM THE UN SYSTEM: THE EU'S STRONG PLURALISM REFLECTED IN THE KADI CASE 107
IV. HARMONIZATION APPROACH: THE *AL-JEDDA* AND *NADA* MODEL 110
V. FURTHER REFLECTIONS ON TECHNIQUES OF TREATY INTERPRETATION IN 'HARD CASES' 114
VI. CONCLUSION 116

I. OPENING REMARKS: SETTING THE SCENE

Art 103 of the UN Charter (or the 'Charter'),[1] in a rather unambiguous

* Visiting Scholar at Lauterpacht Centre for International Law, University of Cambridge and Assistant Professor of International and European Law at the University of Prishtina.

articulation, provides that, '[i]n the event of a conflict between the obligations of the Members of the United Nations under the present Charter and their obligations under any other international agreement, their obligations under the present Charter shall prevail'. The reference to *any other international agreement* reveals the external character of this clause, which presents an exception to the horizontal nature of international law.[2] Furthermore, the legal force of art 103 covers not only its member states, but also 'international and regional organizations ... private contracts, licences and permits'.[3]

In view of the purposes of the UN Charter and its operation, the legitimacy of art 103 emanates from the wide acceptance of this principle by the UN member states,[4] international courts and tribunals,[5] other international treaties,[6] the ILC Report on Fragmentation[7] and opinions of

---

[1] Art 103 of the UN Charter is inherited from art 20 of the Covenant of the League of Nations (the 'Covenant'). For more on art 20 of the Covenant, see especially Hersch Lauterpacht, 'The Covenant as the Higher Law' (1936) 17 British YB Intl L 54-65.

[2] Also *jus cogens* and *erga omnes* obligations belong to the vocabulary of 'informal hierarchy in international law', see in ILC, 'Fragmentation of International Law: Difficulties Arising from the Diversification and Expansion of International Law, Report of the Study Group of the International Law Commission' (13 April 2006) UN Doc A/CN.4/L.682 (ILC Report on Fragmentation) 327. Other authors argue that art 103 of the Charter should be considered as a mere conflict, rather than a hierarchy rule. See eg Antonios Tzanakopulos, 'Collective Security and Human Rights' in Erika de Wet and Jure Vidmar (eds), *Hierarchy in International Law: The Place of Human Rights* (OUP 2012) 66.

[3] Arnold Pronto and Michael Wood, *The International Law Commission 1999-2009,* IV (OUP 2010) 756; See eg UNSC Res 748 (31 March 1992) UN Doc S/RES/748; UNSC Res 1267 (15 October 1999) UN Doc S/RES/1267.

[4] Rudolf Bernhard, 'Article 103' in Bruno Simma *et al* (eds), *The Charter of the United Nations: A Commentary* (2nd edn, OUP 2002) 1293.

[5] See eg *Questions of Interpretation and Application of the 1971 Montreal Convention arising from the Aerial Incident at Lockerbie (Libya v United Kingdom) and (Libya v United States of America), [1993]* ICJ Rep 39–41; *Military and Paramilitary Activities in and against Nicaragua (Nicaragua v United States of America), [1985]* ICJ Rep para 107; Case T-315/01, *Kadi v Council and Commission* [2005] ECR II-3649, paras 183-204; Case T-306/01, *Yusuf and Al Barakaat v Council and Commission* [2005] ECR II-3533, paras 233-254; *Behrami and Behrami v France, Saramati v France, Germany and Norway,* ECHR, Applic. No. 71412/01 and applic no 78166/01, Decision on Admissibility of 2 May 2007, para 61, respectively at para 141; *Berić and Others v Bosnia and Herzegovina*, ECHR, applic no 36257/04, Decision on Admissibility of 16 October 2007, para 29.

[6] See eg art 30(1) Vienna Convention on the Law of Treaties (adopted 23 May 1969, entered into force 27 January 1980) UNTS 1155 (VCLT); art 131 Charter of the Organization of American States (30 April 1948); Seventh Principle of the Declaration on Principles of International Law concerning Friendly Relations and

academics.[8] Following the wide acceptance of the UN superior norm, Benedetto Conforti asserts that 'the principle contained in Article 103 is considered by the whole international community as a principle going beyond the law of treaties and it has come to be regarded as a customary rule'.[9]

Against this background, one would expect that when an issue arises on the basis of art 103 of the UN Charter, no legal system would intend to redefine its scope and effects, since, as Anthony Aust has put it, 'no wise judge (international or national) wants to reinvent the wheel'.[10]

Nevertheless, in recent judicial and doctrinal dialogues the character of art 103 of the Charter has not been accepted without resistance when the claim of universality had to trump obligations of other legal orders related to individual fundamental rights.[11]

Recalling that the SC in its Resolution 1530 (2004) could erroneously attribute the Madrid bombings of 2004 to the ETA organization,[12] a concern that individuals could be victims of similar 'sorry tales of Security Council'[13] resolutions without any procedural guarantee or the right for judicial review triggered scholars and courts to seek for proper responses to some 'arbitrary' SC resolutions.

With regard to the Security Council resolutions blacklisting suspects of (supporting) terrorism, Justice Zinn sitting in the Canadian Federal Court,

---

Cooperation among States in accordance with the Charter of the United Nations (24 October 1970).

[7] ILC Report on Fragmentation (n 2) 324-409.

[8] See eg Bernhard (n 4) 1292–1302; Benedetto Conforti, 'Consistency among Treaty Obligations' in Enzo Cannizzaro (ed), *The Law of Treaties Beyond the Vienna Convention* (OUP 2011) 189-190; Michael Wood, 'The Law of Treaties and the UN Security Council' in Enzo Cannizzaro (ed), *The Law of Treaties Beyond the Vienna Convention* (OUP 2011) 253-254.

[9] See Benedetto Conforti (n 8) 189.

[10] Anthony Aust, 'Peaceful Settlement of Disputes: A Proliferation Problem?' in Tafsir Malick Ndiaye and Rüdiger Wolfrum (eds), *Law of the Sea, Environmental Law and Settlement of Disputes, Liber Amicorum Judge Thomas A. Mensah* (MNP 2007) 137.

[11] Even the (then) ICJ Judge Bruno Simma, a universalist proponent, stated that '[i]f ... universal institutions like the UN cannot maintain a system of adequate protection of human rights, considerations of human rights deserve trump arguments of universality' in Bruno Simma, 'Universality of International Law from the Perspective of a Practitioner', (2009) 20 *EJIL* 294.

[12] The arrest of several radicals associated with Al-Qaeda organization by Spanish authorities for the Madrid bombings of 2004 is an indication of erroneous attribution to the ETA organization.

[13] Therese O'Donnell, 'Naming and Shaming: The Sorry Tale of Security Council Resolution 1530 (2004)', (2007) 17 *EJIL* 946 ff.

in the *Abdelrazik* case stated that 'there is nothing in the listing or de-listing procedure that recognizes the principles of natural justice or that provides for basic procedural fairness'.[14] Indeed, as already emphasized 'it is the procedural system per se that does not respect such rights'.[15]

While judicial bodies worldwide have been struggling to balance human rights and subordination to the UN supremacy, this has been particularly sensitive for European courts where protection of fundamental rights is of a paramount importance in the hierarchy of norms.
When placed between the UN obligations vested in art 103 and domestic fundamental rights, from the methodological point of view, European courts appear to have adopted three approaches, namely the 'subordination', 'detachment' and 'harmonization'. Oscillation on this varied trinity of approaches, affects not only rights of the individuals concerned but also the legal (un)predictability and coherence of international law.

Majority of European case law follows the 'subordination' approach where constitutional and conventional fundamental rights are trumped when in conflict with the SC resolutions, by considering unattainable to accommodate simultaneously two obligations of different legal orders.[16] The second approach, the 'detachment' from the UN system is a reverse approach of the 'subordination', generated also by a narrow understanding of art 103.[17] This is peculiar particularly for the EU legal order, which in the widely debated *Kadi* case considered the EU law as a 'supreme law of the land', and developed a dualist or strong pluralist approach[18] that led to detachment from the UN supremacy. 'Harmonization' is the third

[14] Case T-727/08 *Abdelrazik v Minister of Foreign Affairs and Attorney General of Canada* [2009] FC 580, 51.
[15] Salvatore Zappalà, 'Reviewing Security Council Measures in the Light of International Human Rights Principles' in Bardo Fassbender (ed), *Securing Human Rights? Achievements and Challenges of the UN Security Council* (OUP 2011) 182.
[16] See eg *Kadi* (n 5); *Behrami and Saramati* (n 5); *Berić and Others* case (n 5); *R (on the applicant of Al-Jedda) (FC) v Secretary of State for Defence* [2007] UKHL 58 (2008); *Youssef Nada v State Secretariat for Economic Affairs and Federal Department of Economic Affairs*, Administrative Appeal Judgment of 14 November 2007, BGE 133 II 450, 1A 45/2007.
[17] See Joined Cases C-402/05 P and C-415/05 P *Kadi and Yusuf v Council and Commission* [2008] ECR I-6351; T-85/09 *Kadi v Commission* [2010] ECR II-05177.
[18] 'Strong pluralist approaches deny the possibility of a shared, universally-oriented system of values and question the meaningfulness of the idea of an international community' in Gráinne de Búrca, 'The European Court of Justice and the International Legal Order After *Kadi*' 51 *HILJ* (2010) 4 (the description contained in fn 10).

approach, developed by the ECtHR in *Al-Jedda* and *Nada* cases.[19] This methodology provides that by utilizing techniques and norms of treaty interpretation, courts can balance concerns of the UN supremacy with human rights.

In the same structural order, this contribution examines the aforementioned responses by adding a section with theoretical considerations that provide pertinent tools for reaching harmonious accommodation of different treaty obligations in domestic cases. The conclusion highlights the key features of these approaches and considers their relevance in the ongoing discussion on regime interaction in international law.

II. SUBORDINATION APPROACH: WHEN COHERENCE BECOMES THE ANTONYM OF FUNDAMENTAL RIGHTS

The wide and cross-regime acceptance of the UN superior norm was reflected in most of the European case law, even when conformity with art 103 resulted in jeopardizing domestic fundamental rights. The General Court of the European Union (GCEU)[20] in *Yusuf and Kadi* chose the subordination approach. It considered that decisions of the SC overruled the European Union (EU) law even in light of fundamental rights and the GCEU had no mandate to review obligations originating from the SC.[21] While the GCEU regarded that a review could be done exceptionally based on *jus cogens* violations,[22] nevertheless it did not find any *jus cogens* violation, and firmly applied art 103 of the Charter as a conflict and superior rule.

The ECtHR in *Behrami and Saramati* and *Berić and Others* did not engage in the alleged violations of the fundamental rights, including right to life,[23] liberty and security,[24] and an effective remedy,[25] as such an examination was precluded due to court's finding on the attribution of the conduct in question. The court observed that in authorizing the military mission in Kosovo the UN SC retained the 'ultimate authority and control' over it.

---

[19] *Al-Jedda v The United Kingdom*, ECHR, applic no 27021/08, Judgment of 7 July 2011; *Nada v Switzerland*, ECHR, applic no 10593/08, Judgment of 12 September 2012.

[20] Before the entry into force of the Consolidated Version of the Treaty on European Union [2008] OJ C115/13 (TFEU), the GCEU was known as the Court of First Instance (CFI).

[21] *Kadi* [2005] (n 5) 221-231. *Yusuf* [2005] (n 5) 277-282.

[22] ibid 176.

[23] See eg *Behrami* (n 5) 61.

[24] See eg *Saramati* (n 5) 141.

[25] See eg *Berić* (n 5) 29-30.

Therefore, the ECtHR was not faced with complexities of art 103 and its possible relationship and effects on the regional convention. At the same time, the ECtHR's reasoning in Behrami and Saramati attracted a significant debate. When observing criticism in writings of jurists on the 'ultimate control' test employed by the ECtHR and the manner of 'attribution of conduct' to states and international organizations,[26] the underlying significance of art 103 of the UN Charter is apparent.

The House of Lords of the United Kingdom in the Al-Jedda case[27] and the Swiss Federal Supreme Court in the Nada case[28] enriched the case law providing for the gentle subordination approach, whereby both European domestic higher courts unanimously held that art 103 gave primacy to resolutions of the SC, even in relation to human rights agreements.

For the purposes of this contribution, the common denominator of the aforementioned cases in the European jurisdictions is that in implementation of certain SC obligations, courts chose to obey rules of the SC and set aside their respective fundamental conventional or constitutional rights. This approach where courts are unable to reach synergy between, what appear to be, conflicting norms provides for a narrow interpretation of norm conflict 'where giving effect to one international obligation unavoidably leads to the breach of another obligation or right'.[29] It may be said that a court rather looks at the terms of conflicting obligations in clinical isolation and omits the quest for 'regime compatibility'[30] or harmonization between the SC obligation to maintain peace and security on the one hand, and protection of fundamental rights in the European or domestic legal order on the other.

Following a broader interpretation of art 103, the European courts could

[26] See eg Giorgio Gaja, Seventh Report on Responsibility of International Organizations, UN Doc A/CN.4/610, (27 March 2009), at 10-12; Andrew Clapham, 'The Subject of Subjects and the Attribution of Attribution' in Laurence Boisson de Chazournes and Marcelo Kohen (eds), *International Law and the Quest for its Implementation Le droit International et la quête de samise en oeuvre: liber amicorum Vera Gowlland-Debbas* (MNP 2010) 53-56; Marko Milanovic, 'Al-Skeini and Al-Jedda in Strasbourg', (2012) 23 *EJIL* 1, 134-135; Kjetil Mujezinović Larsen, 'Attribution of Conduct in Peace Operations: The "Ultimate Authority and Control" Test', (2008) 19 *EJIL* 509.

[27] *Al-Jedda* (n 16).

[28] *Nada* (n 16).

[29] Erika de Wet and Jure Vidmar (n 2) 1. See also C Wilfred Jenks, 'Conflict of Law-Making Treaties' (1953) 30 British YB Intl L 401.

[30] Dirk Pulkowski, 'Universal International Law's Grammar' in Ulrich Fastenrath *et al* (eds), *From Bilateralism to Community Interest: Essays in Honour of Judge Bruno Simma* (OUP 2011) 154.

power" committed to effective multilateralism under international law'[47] should be reflected upon when facing application of international law. Following the position of the EU foreign policy, one may observe that the EU is cognizant of the importance of pluralism and international law.[48]

Accordingly, it seems unsurprising that the approach of the EU courts in *Kadi* caused dissent within its legal order. In December 2010, the Commission, the Council and majority of the EU member states filed another appeal and intervened in *Kadi*,[49] arguing:

[t]he United Nations Charter requires compliance with its obligations by its Member States. Such obligations prevail over the obligations which may arise under any other international agreement. Such obligations include those imposed under Security Council resolutions intended to combat international terrorism. Having regard, in particular, to Articles 3(5) and 21 TEU and Article 351 TFEU, the obligation upon EU Member States to comply with the decisions of the Security Council prevails over any obligations which may arise under the EU Treaties. The EU must consider itself bound by the terms of the UN Charter and the UN Security Council decisions made under it.[50]

If the two EU legislating institutions and the majority of its member states claim different understanding of the relationship between the EU law and the UN law, the EU courts should pay due regard to the claimed impression that the EU law is solely within discretion of its judges. In fact, even in cases when judges make law, Ronald Dworkin suggests that 'they should act as deputy to the appropriate legislature, enacting the law that they suppose the legislature would enact if seized of the problem'.[51] Immanuel Kant also argues that the "law establishes the omnilateral or

[47] ibid 3.
[48] The President of the European Commission José Manuel Durão Barroso has outlined a vision of the EU's foreign policy in the following terms: 'We certainly welcome pluralism in international relations ... In international relations, partnerships and a multilateral approach can achieve so much more ... we need a renewed politics of global engagement, particularly with international institutions ... because that is the only way we can consolidate and strengthen a stable, multilateral world, governed by internationally-agreed rules'. See José Manuel Durão Barroso, 'A Letter from Brussels to the Next President of the United States of America' (2008), Lecture at Harvard University.
[49] Joined Cases C-584/10 P, C-593/10 P, C-595/10 P.
[50] ibid C-595/10 P. Appeal brought by the United Kingdom of Great Britain and Northern Ireland against *Kadi* judgment [2010] (n 16).
[51] Ronald Dworkin, 'Hard Cases' (1975) 88 *HLR 6,* 1058.

'general united will' of a community".[52] 'This will is understood as an "all-sided will" or, ... "all the Wills of a Community together".'[53] This does not suggest that the EU courts should be influenced by political pressures of EU institutions and member states, but rather to quest for a more balanced interpretation when dealing with the UN obligations that might affect responsibility of the EU member states.

If the CJEU decides not to revisit its memorable *Kadi* reasoning, the case might enhance political and legal tensions. From a political perspective, the EU member states being obliged to respect the SC resolutions and the supremacy of the EU law will be left to respond with the principle of political decision. This principle developed by Manfred Zuleeg provides that 'the state concerned simply has to make a political decision which commitment to prefer'.[54]

From the international law's perspective, the challenge of art 103 based on the EU strong pluralism, where legal orders escape the UN obligations by rules of domestic law, 'may obliterate boundaries of legality'.[55] As a consequence, the detachment approach 'might reinforce perceptions of international law as non-law (or quasi-law) – i.e., a loose system of non-enforceable principles, containing little, if any real constraints on state power'.[56] Indeed, if the backbone UN principle - art 103 of the Charter is challenged by strong pluralist views of legal orders, not applying the higher UN law, the claim of the CJEU Judge Allan Rosas that '[i]nternational [l]aw is dead' would not be an exaggeration.[57]

IV. HARMONIZATION APPROACH: THE *AL-JEDDA* AND *NADA* MODEL

The foregoing observations indicate that by means of treaty interpretation the courts of the same legal order in the *Kadi* case could reach diametrically different outcomes. While both approaches present

[52] Patrick Capps and Julian Rivers, 'Kant's Concept Of International Law' (2010) 16 *Legal Theory* 233.
[53] ibid. See also Immanuel Kant, *The Philosophy of Law: An Exposition of the Fundamental Principles of Jurisprudence as a Science of Right,* (William Hastie trs, T&T Clark 1887) 84.
[54] Manfred Zuleeg, 'Vertragskonkurrenz im Völkerrecht. Teil I: Verträge zwischen souveränen Staaten' (1977) 20 German YB Intl L 246-76, cited in Jan Klabbers, *Treaty Conflict and the European Union* (CUP 2009) 88.
[55] Adré Nollkaemper, 'Rethinking the Supremacy of International Law', (2010) 65 *ZaöR,* 74.
[56] Yuval Shany, 'Toward a General Margin of Appreciation Doctrine in International Law?' (2005) 16 *EJIL* 912.
[57] Allan Rosas, 'The Death of International Law?' (2011) 20 Finish YB Intl L 227.

examples of a narrow interpretation of norm conflict, the reasoning of the ECtHR in *Al-Jedda* provides an addendum as to how pertinent tools of treaty interpretation allow reaching a more harmonious and constructive outcome than the one of the House of Lords, which decided to set aside the application of the European Convention on Human Rights (ECHR) for the sake of the UN supremacy.[58]

Al-Jedda triggered a question of continual internment in light of art 5(1) of the ECHR and the SC Resolution 1546 (2004). The ECtHR's response began with a survey of commonalities on the issue of human rights as reflected in the principles and purposes of the UN Charter. By means of harmonious interpretation, the ECtHR considered that 'in interpreting ... [the SC] resolutions, there must be a presumption that the Security Council does not intend to impose any obligation on Member States to breach fundamental principles of human rights'.[59] Indeed, the SC itself appealed in its resolutions that while 'promoting the maintenance of security and stability ... to act in accordance with international law'.[60] This implies also acting in accordance with human rights treaties, because '[h]uman rights are part of international law'.[61]

After setting the scene on 'regime compatibility', the ECtHR argued that the Resolution 1546 did not provide for 'clear and explicit language... [requiring] States to take particular measures which would conflict with their obligations under international human rights'.[62] Consequently, the ECtHR asserted that 'it must therefore choose the interpretation which is most in harmony with the requirements of the Convention and which avoids any conflict of obligations'.[63] In this manner, the ECtHR came to its eloquent findings by considering that Resolution 1546 was ambiguous with regard to the issue of 'obligation to intern'.[64] While there could be grounds to argue in support of internment in the context of the Resolution 1546,[65] the ECtHR by relying on the relevant UN reports[66] considered that internment was not *intended* as an obligation that set aside human rights

[58] *Al-Jedda* (n 16).
[59] ECtHR, *Al-Jedda* (n 19) para 102 [emphasis added].
[60] UNSC Res 1546 (8 June 2004) UN Doc S/RES/1546, Preamble.
[61] See eg UNAMI, Human Rights Report 1 April – 30 June 2007, para 77.
[62] ECtHR, *Al-Jedda,* (n 19) para 76 [emphasis added].
[63] ibid.
[64] ibid 105.
[65] ibid 34 and 108. The annexed letter to the SC Resolution 1546 (2004) provides that, 'will include combat operations against members of these groups, internment where this is necessary for imperative reasons of security ...'. Para 10 of the UNSC Res 1546 (8 June 2004) considered the annexed letters as integral part of the Resolution.
[66] ibid 40-41, the ECtHR made reference to the UN Assistance Mission for Iraq (UNAMI), Human Rights Reports.

obligations.[67]

In sum, the ECtHR, after scrutinizing every angle of the UN superior norm and clarifying that art 103 did not create obligations to intern in the present case, gave effect to its conventional rights without eroding the vertical norm of international law.

In view of the recent jurisprudence of the ECtHR, it appears that the *Al-Jedda* model of harmonization does not reveal an accidental response to situations when the ECHR is juxtaposed with SC obligations. On 12 September 2012, the Strasbourg Court in its landmark *Nada* judgment asserted:

[w]here a number of apparently contradictory instruments are simultaneously applicable, international case-law and academic opinion endeavour to construe them in such a way as to coordinate their effects and avoid any opposition between them. Two diverging commitments must therefore be harmonised as far as possible so that they produce effects that are fully in accordance with existing law.[68]

While the ECtHR in the *Nada* case reiterated its readiness to pursue simultaneous accommodation of fundamental rights and SC obligations, the facts surrounding Mr. Nada's case should be distinguished from the facts in *Al-Jedda*. In particular, the situation in *Nada* emanated from more explicit terms of the SC Resolution 1390 (2002), requiring, *inter alia*, freeze assets, and apply the entry and transit ban against Mr. Nada.[69] The listing of Mr Nada in the 'Taliban Ordinance', associating him with serious terrorist activities, created a situation that differed from that in *Al-Jedda*, and reflected similarities to *Kadi*. In this regard, the ECtHR recognized:

contrary to the situation in Al-Jedda, ... Resolution 1390 (2002) expressly required States to prevent the individuals on the United Nations list from entering or transiting through their territory. As a result, the ... [*Al-Jedda*] presumption is rebutted in the present case, having regard to the clear and explicit language, imposing an obligation to take measures capable of breaching human rights, that was used in that resolution (... in paragraph 7 of Resolution 1267 (1999) ... the Security Council was even more explicit in setting aside any other international obligations that might be incompatible with the resolution).[70]

---

[67] See also Marko Milanovic (n 26) 137.
[68] ECtHR, *Nada* (n 19) 170.
[69] UNSC Res 1390 (28 January 2002) UN Doc S/RES/1390, part 2 (a) and 2 (b).
[70] ECtHR, *Nada,* (n 19) para 172 [emphasis added].

Although the ECtHR acknowledged that the clear and explicit terms of the Resolution 1390 obliged Switzerland to take measures that may breach human rights, it also found that 'the Charter in principle leaves to UN member States a free choice among the various possible models of transposition of those resolutions into their domestic legal order'.[71]

After giving attention to the CJEU's reasoning in *Kadi*, particularly on domestic courts' latitude in choosing the means for implementation of the SC resolutions, the ECtHR again shifted to its 'linguistic ambiguity exercise', by identifying spaces in the terms of the SC Resolution 1390 in which to accommodate fundamental rights.[72] In this manner, the ECtHR held that the wording employed in the Resolution 1390 '*where appropriate*' and '*necessary*' comprised on the part of the national authorities 'certain flexibility in the mode of implementation of the resolution'.[73] In support of this argument, the ECtHR took into account the Swiss Parliament's statement to the UN SC 'that it would no longer unconditionally be applying the sanctions prescribed against individuals under the counter-terrorism resolutions'.[74]

In this light, the ECtHR in *Nada* concluded that 'Switzerland enjoyed some latitude, which was admittedly limited but nevertheless real, in implementing the relevant binding resolutions of the UN Security Council'.[75] In this way, the ECtHR neither challenged the supremacy of art 103 of the UN Charter, nor sacrificed fundamental rights. In this 'hard case' of a rather straightforward language of the SC Resolution 1390, the ECtHR maneuvered to preserve both, the coherence of international law and fundamental rights, by recourse to a linguistic test which, as one may observe, centers on spaces in SC resolutions.

The substantial ECtHR's findings flowing from the less clear SC resolutions in *Al-Jedda*, and, more recently, in *Nada*, constitute an alarm for the SC that more precision may be expected in the language of its resolutions, particularly, when triggering human rights issues.

Overall, the ECtHR holdings in *Al-Jedda* and *Nada* are a result of a broader interpretation of norm conflict and provide invaluable examples of how

[71] ibid 176.
[72] Marko Milanovic, 'European Court Decides Nada v Switzerland' (*EJIL Talk*, 14 September 2012) <http://www.ejiltalk.org/european-court-decides-nada-v-switzerland/> accessed 30 September 2012.
[73] ECtHR, *Nada* (n 19) 178.
[74] ibid 179.
[75] ibid 180.

coherence of international law could be preserved outside the 'either, or' constraints when dealing with two prima facie conflicting obligations. By relying on a systemic integration, the ECtHR read different treaty rule systems in a 'mutually supportive light'.[76] The customary nature of the principle of systemic integration,[77] enshrined in art 31(3)(c) VCLT urges that in cases of treaty interpretation, together with the context there should be taken into account 'any relevant rules of international law applicable in the relations between the parties'.

In search of what Cicero called *topoi* or common places, the ECtHR applied topical jurisprudence[78] and referred not only to SC resolutions as the applicable law in the present cases, but also utilized other inspirational sources from the national case law, the ILC Reports, and other non-binding documents.[79] Even though the outcome in *Al-Jedda* and *Nada* was solely based on the applicable law, the reference to other non-binding sources and the expanded interpretation of art 103 provide that the ECtHR reasoning in *Al-Jedda* and *Nada* is not only convincing to the litigants and the community of the court's regime, but also persuasive in the context of international community interest.

V. FURTHER REFLECTIONS ON TECHNIQUES OF TREATY INTERPRETATION IN 'HARD CASES'

While one may regard that the SC resolutions discussed in *Al-Jedda* and *Nada* contained rather unambiguous terms, the ECtHR still attained to find a room for its invaluable harmonization approach. The court embarked on a rule, which entails that it is only when the relevant SC resolution employs clearly defined terms the court will consider whether the SC resolution trumps the fundamental rights in question. Until such

[76] ILC Report on Fragmentation (n 2) 417. See also para 271.

[77] PCA in *Iron Rhine* case stipulated that: 'Articles 31 ... of the Convention reflect pre-existing customary international law'. See *Arbitration regarding the Iron Rhine (Ijzeren Rijn) Railway (Belgium v the Netherlands) (2005)* PCA Reports (2005) para 45.

[78] For the relevance of *topoi* and topical jurisprudence for international law, see Kushtrim Istrefi and Zane Ratniece, 'Think Globally, Act Locally: Al-Jedda's Oscillation Between the Coherence of International Law and Autonomy of the European Legal Order' (2012) 24 Hague YB Intl L 249-253. See also Vassilis P Tzevelekos, 'The Use of Article 31(3)(C) of the VCLT in the Case Law of the ECtHR: An Effective Anti-Fragmentation Tool or a Selective Loophole for the Reinforcement of Human Rights Teleology?' (2010) 31 *MJIL* 620-690; Anita Soboleva, 'Topical Jurisprudence: Reconciliation of Law and Rhetoric' in Anne Wagner, Wouter Werner and Deborah Cao (eds), *Interpretation, Law and the Construction of Meaning: Collected Papers on Legal Interpretation in Theory, Adjudication and Political Practice* (Springer 2006) 49.

[79] ECtHR, *Al-Jedda* (n 19) paras 42-58. ECtHR, *Nada* (n19) paras 65-101.

time, the ECtHR's linguistic exercise in identifying spaces in the language of SC resolutions, appears to achieve the preservation of the conventional rights when, at first sight, juxtaposed with arbitrary SC resolutions.

While recognizing the value of *Al-Jedda* and *Nada* in developing the treaty interpretation techniques, one may put forward that the ECtHR's harmonization approach is not a one-size-fits-all methodology capable to resolve all cases involving SC resolutions, particularly those with clear and firm terms. Although this approach may be applicable to cases involving international and regional courts, the ECtHR itself held in the *Nada* case that domestic courts 'are to choose the means by which they give effect to the [SC] resolutions'.[80] In this regard, observations of Joost Pauwelyn may be recalled that 'if the reconciliation between the two norms is not feasible, that is where the presumption [against normative conflict] ends'.[81] This entails limitations in possibilities to avoid conflicting obligations of different rule systems, particularly for domestic courts. However, even in hard cases, techniques of treaty interpretation remain useful to alleviate the degree of contradiction and confrontation between legal orders. This may still be useful when discussing the coherence of international law.

With regard to hard cases that emanate from clear terms of the SC resolutions and opposing fundamental rights of other legal orders, André Nollkaemper suggests that a balance with human rights could be achieved by 'identifying a criterion for qualifying the principle of supremacy that may lead to synergies between the international and domestic legal orders'.[82] This criterion is 'the conformity of a rule of fundamental rights under domestic law with international rights'.[83] Considering that most domestic fundamental rights originate or are similar to those of international law, domestic courts by means of treaty parallelism and harmonization could reach a similar conclusion for the protection of domestic fundamental rights as it would, if a dualistic approach were taken. However, this argument echoes the necessity to apply an international law approach[84] and thus leaves no flexibility for domestic legal orders to consider balancing obligations of the UN Charter by means of 'domestic choices'. A domestic fundamental right may not necessarily

[80] ibid *Nada*, (n 19) at para 176 [emphasis added].
[81] Joost Pauwelyn, *Conflict of Norms in Public International Law: How WTO Law Relates to other Rules of International Law* (CUP 2003) 242 [emphasis added].
[82] André Nollkaemper (n 55) 76.
[83] ibid.
[84] Eg Universal Declaration of Human Rights (adopted 10 December 1948 UNGA Res 217 A(III) (UDHR), International Covenant on Civil and Political Rights (adopted 16 December1966, entered into force 23 March 1976) 999 UNTS 171 (ICCPR) etc.

always have universal relevance (eg Sharia law)[85] and the dualist approach of legal orders does not correspond to the external and vertical character of art 103 of the Charter.

VI. CONCLUSION

In the 2008 ESIL Biennial Conference, Judge Bruno Simma argued that 'heterogeneity does not exclude universality of international law'.[86] Four years later, this argument has become even more pertinent considering the increased vivid regime interaction in judicial fora. In the world of plurality and co-existence of legal orders, when dealing with treaty conflicts, the 'either, or' approach as observed in 'subordination' or 'detachment' approaches, seems not to follow the trend of international law development. While art 103 of the Charter must be obeyed as a rule of 'last resort' (save when in conflict with *jus cogens*), in many instances the best application of art 103 may be no application at all.[87] Legal orders juxtaposed with the UN superior norm should by means of treaty interpretation search for common places and harmonization instead of confrontation.

This contribution thus argues that norms and techniques of interpretation, led by systemic integration are not dogmatic tools for theoretical entertainment. Instead, they present concrete and useful techniques in mitigating treaty conflicts in the new reality of international law.

Koskenniemi suggests that any legal concept must have its normativity and concreteness.[88] 'The normativity ... has to do with its "oughtness", the way it does not merely describe some aspects of reality but poses requirements for it ... [and the concreteness] must reflect what actually takes place in the political and economic world'.[89] In the author's view, the request for broader and systemic interpretation has the element of normativity as it presents the customary norm and the treaty obligation.[90] Considering that heterogeneity and fragmentation have become an integral part and the parcel of current international law, its concreteness could have never been

[85] See also André Nollkaemper (n 55) 67.
[86] Bruno Simma (n 11) 264.
[87] Andreas Paulus and Johann Leiß, 'Article 103' in Bruno Simma *et al* (eds), *The Charter of the United Nations: A Commentary* (3rd edn, OUP 2012) 2114.
[88] Martti Koskenniemi, 'International law in the world of ideas' in James Crawford and Martti Koskenniemi (eds), *The Cambridge Companion to International Law* (CUP 2012) 60.
[89] ibid [emphasis added].
[90] For customary norm see PCA, *Iron Rhine* case (n 77). For treaty obligation see art 31(3)(c) VCLT (n 6).

more apparent. Moreover, 'the demand to relate interpretations to the system of law is part of positive law and of the prevailing legal ethos'.[91]

Consequently, when confronted with issues of fragmentation, judges should seek for judicial comity in applying the virtues of the legal techniques of interpretation discussed above. This would also meet the double and simultaneous function of the judges to contribute to the coherence of international law and safeguard the fundamental rights of their respective legal orders. In addition, domestic and regional courts should revisit harsh legislating powers on the issues of global concern and rather resolve disputes on a case-by-case basis, thus making it unnecessary to engage in on the morality and values of other legal orders.[92]

The evidence that European courts have employed three diverse approaches when applying art 103 of the Charter reflects that there is no single understanding of how 'arbitrariness' associated with the SC obligations should be diminished. Nevertheless, the ECtHR recent case law, led by *Al-Jedda* and *Nada*, and the constructive engagement of scholars indicate that European courts are developing solutions that allow harmonization and regime compatibility that disturb neither coherence of international law nor autonomy of the respective regimes.

[91] Armin von Bogdandy and Ingo Venzke, 'In Whose Name? An Investigation of International Courts' Public Authority and Its Democratic Justification' (2012) 23 *EJIL* 1, 37.

[92] Jan Klabbers suggests that 'treaty conflicts are unsolvable as a matter of law as soon as they emanate from clashes of values' [emphasis added]. Klabbers (n 54) 35.

THE SCOPE AND THE LEGAL LIMITS OF THE 'IMMIGRATION FEDERALISM': SOME COMPARATIVE REMARKS FROM THE AMERICAN, BELGIAN AND THE ITALIAN EXPERIENCES

Davide Strazzari*

Traditionally, immigration has generally being conceived of as a matter to be dealt with by the national legislator. However, immigration federalism - that is, the regulatory role that sub-national territorial units, enjoying legislative powers, experience with regard to issues related to immigration policy - has become a very sensitive issue in many countries.

By focussing on the comparison of three legal systems (the USA, Belgium and Italy), this article highlights three main issues challenged by the emergence of immigration federalism: the division of powers, access to welfare and cultural-linguistic integration in the context of multinational states.

The comparative analysis reveals one important difference between these countries. While in the US immigration is interpreted as a federal reserved power - allowing the federal authorities to regulate, not only, the entry and stay of aliens, but also their rights and duties, to the point of encroaching on State matters - this does not occur in both Italy and Belgium.

As a consequence, in these two countries sub-national units have had more chances to freely develop immigration-related policies. In the Italian case this has occurred especially in the field of welfare, while in Belgium it has emerged in the linguistic integration policy. At the same time, however, the judiciary has used the principle of equality and the protection of fundamental rights to ensure a certain level of territorial harmonisation, and contrast discriminatory approaches by the sub-national units. Both models present some inconsistencies.

In the final part of the essay, we suggest the development of cooperative federalism as an alternative means of structuring territorial relations within the immigration field. This solution seems more consistent with the idea that immigration is not in itself a jurisdiction, but constitutes a policy, composed of measures falling under various constitutional jurisdictions, which are vested in both the national and the sub-national tiers of government.

---

* Researcher in Constitutional Comparative Law, Department of Sociology, Trento University. This paper draws upon researches conducted as visiting researcher at the Law Faculty of the Université Catholique de Louvain la Neuve (B). I would like to thank Professor Y. Lejeune for his precious comments and suggestions on an earlier draft of this paper. Any inaccuracy is my own.

TABLE OF CONTENTS

I. INTRODUCTION 119
II. THE DIVISION OF POWERS ISSUE 124
1. *USA* 125
2. *Belgium* 132
a. Explicit Powers Regarding Immigration Policy 134
b. Implicit Powers Regarding Immigration Policy: The Case of Welfare 136
c. The Powers Reserved to the Federal Authorities: The Case of the Legal Status of Aliens as an Autonomous Standing Power Clause 140
3. *Italy* 142
III. THE 'SOCIAL REGIONAL CITIZENSHIP' DIMENSION 145
1. *USA* 150
2. *Belgium* 152
3. *Italy* 157
IV. THE "CULTURAL REGIONAL CITIZENSHIP" DIMENSION 162
1. *The Flemish Inburgering Decree* 166
V. CONCLUSIONS 169

I. INTRODUCTION

It is commonly argued that globalisation has led to the weakening of the regulatory role of the State in favour of supranational or international organisations, without simultaneously preventing sub-national territorial units from strengthening their role as promoters of territorial specificities through their regulatory functions. A new label has even been coined in order to describe this, namely "glocalisation".

In this context, it is germane to examine the issue of 'immigration federalism', that is, the regulatory role that sub-national territorial units, enjoying legislative powers, experience with regard to issues related to immigration policy.

Traditionally, immigration has generally being conceived of as a matter to be dealt with by the national legislator. This is consistent with the idea

that the power to decide who may or may not enter the country derives from the sovereignty principle, which pertains to the national authorities.

However, at least in the European context, international and supranational legal orders are increasingly providing limits to the discretion of national States with relation to immigration policy and the legal status of aliens. Take, for example, the increasing measures the EU has taken in relation to third-country nationals in recent years. This is especially evident since the Amsterdam treaty 'communitarised' the relevant policy area.[1]

The European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR), as interpreted and applied by the European Court of Human Rights (ECtHR), also deserves due consideration. Although relatively few provisions of the ECHR are explicitly directed towards aliens, it is important to emphasize that the Convention applies to all individuals who are subject to the jurisdiction of the States that are parties to the Convention.[2]

These short remarks demonstrate how, at least in Europe, international and supranational forces may influence national decisions concerning immigration measures or the legal status of aliens.

Less attention has been paid to the role that sub-national units perform with regard to immigration policy. However, it is increasingly frequent for sub-national units to act in this area.

In the US, the Supreme Court has recently deemed illegitimate an Arizona statute that empowers state and local officials to stop individuals suspected to be illegal immigrants.

Both Flanders and Catalonia have recently passed acts imposing compulsory linguistic and civic courses upon immigrants. Failure to attend such courses is sanctioned either by an administrative fine or by legal constraints to the renewal of the permit of stay.

The Italian Constitutional Court has already settled a variety of cases concerning the constitutionality of regional statutes dealing with immigration issues.

---

[1] This is now regulated in Title V, ch 2, arts 77-80, of the TFEU. See Steve Peers, *EU Justice and Home Affairs Law* (3rd edn, OUP 2011).

[2] See Bruno Nascimbene and Chiara Favilli, 'Straniero (tutela internazionale)', in Sabino Cassese (ed) *Dizionario Diritto Pubblico* (Giuffrè 2006) 5804; Vincent Chetail (ed), *Mondialisation, Migration et Droits de l'Homme: Le Droit International en Question* (Bruylant 2007).

This paper aims at focussing attention on 'immigration federalism' in three relevant States, namely the USA, Italy and Belgium. The choice of these States is due to the fact that issues pertaining to "immigration federalism" are currently very sensitive in each of these jurisdictions. However, meaningful institutional differences characterise the three States. It can thus be recalled that whereas the USA represents an example of a federal State constituted from the aggregation of previously independent States, in the case of Italy and Belgium we face two legal orders where decentralisation occurred in the context of a previously unitary State. At the same time, the Belgian federalisation process is due to the need to preserve cultural and linguistic sub-national territorial identities (of the Walloon and the Flemish nations), a feature that it does not share with the US and to a large extent with Italy as well. Finally, both Italy and Belgium are part of the EU and are signatory parties of the ECHR, a feature that is evidently not shared by the USA.

All of these aspects – ie the type of federalism, the minority linguistic issue and the legal influence exercised by supranational and international legal orders – constitute grounds that may influence "immigration federalism", rendering it an interesting topic for a comparative analysis.

The three national experiences will thus be evaluated taking into account three dimensions of 'immigration federalism': namely the division of powers, the emergence of a 'regional social citizenship' and the emergence of a 'cultural regional citizenship'.

This paper suggests that immigration as such cannot be considered as a matter in itself but rather as a policy composed of measures falling under several different constitutional jurisdictions. Immigration can be assessed as if it were a shared policy, where the role of sub-national units varies in relation to the specific area that can be connected to the immigration policy label.

The development of 'immigration federalism' does not only involve the division of powers.

The massive influx of immigrant newcomers alters the relationship between the sub-national units' authorities and the individuals falling under their jurisdiction. Whereas the idea of the relationship between the public authorities of a national State and the relevant people is based upon the notion of citizenship, as a rule, there is no formal instrument to describe the relationship between the authorities of sub-national units and

the individual.[3]
However, in order to determine to which persons a regional act is applied, sub-national units must utilise some criteria that link those persons to the territory. Amongst these, residence is becoming increasingly important, especially in areas, such as welfare, that are strictly concerned with individuals.

Thus, the more the sub-national units are called to deal with policies related to the individuals, the more that residence becomes a means of stressing the sense of belonging to the given sub-national unit. This sense of regional territorial belonging becomes even stronger in cases of multinational states where the *raison d'être* of the decentralised form of government is the preservation of the cultural identity of the persons inhabiting that territory.

In such cases, we may say that residence is the functional equivalent of a sort of 'regional' citizenship, to the limited extent of expressing the idea of a relationship of 'belonging' between the person and his territorial unit.[4]

There are at least two areas where this can be easily noted.

The first of these is welfare. Since welfare rights, especially when they are financed via general taxation, are dependent upon budgetary resources, the sub-national units have an interest in limiting the regional welfare

[3] Some federal states have federal citizenship status, as well as state citizenship status. However, whereas at the beginning of federations, federal citizenship is usually derivative of the existing state citizenship, such systems tend to be abandoned at a later juncture: citizenship in a sub-national unit follows from the fact of being a national citizen residing in the relevant sub-national unit. This is the case of the US, with the insertion in the Constitution of the XIV Amendment after the Civil War. Thus, since currently the status of state citizenship, where it formally exists, derives from being a *national* citizen residing in the relevant state unit, it is not a suitable instrument to describe the ties of *non-national citizens* with a sub-national unit, according to the perspective considered in this contribution. On the issue, see Christoph Schönberger, 'European Citizenship as Federal Citizenship – Some Citizenship Lessons of Comparative Federalism', in *Citizenship in the European Union/Citoyenneté dans l'Union Européenne,* (Esperia Publications LTD, European Public Law Series 2007) 61; Olivier Beaud, 'The Question of Nationality within A Federation: a Neglected Issue in Nationality Law' in Randall Hansen and Patrick Weil (eds), *Dual Nationality, Social Rights and Federal citizenship in the U.S. and Europe - The Reinvention of Citizenship* (Berghahn Books 2002) 314; Peter H Schuck, 'Citizenship in Federal Systems' (2000) 48 A J Comp L 195.

[4] See, in the Italian legal scholarship, Federico Dinelli, 'La Stagione della Residenza: Analisi di un Istituto Giuridico in Espansione' [2010] Diritto Amministrativo 639, Stefano Sicardi, 'Essere di quel Luogo. Brevi Considerazioni sul Significato di Territorio e di Appartenenza Territoriale' [2003] Politica del Diritto 115.

eligibility to persons showing a genuine link with the territory. Thus, residence in the given sub-national territory as a precondition for being awarded social benefits is not enough, and tends to be coupled with a long-term residence requirement that discriminates against newcomers, and particularly new immigrants.

The second area where the sense of belonging to a regional territorial area may be significant is language. We refer here to those countries characterised by a multinational structure, or where linguistic-national minorities are established. Those sub-national units where a national component or minority is mainly settled (thus constituting a 'majority' with respect to that territory) may feel the need to preserve their cultural homogeneity and distinctiveness with respect to new immigrants, especially when the latter are deeply motivated to learn the 'majoritarian' language.[5]

Both 'social' and 'cultural regional citizenships' may be developed through instruments aimed at protecting the autochthonous communities, and at discouraging immigrants from settling in the relevant sub-national unit. Questions may thus arise in relation to the respect of fundamental rights and the principle of equality. Since the enforcement of these principles is conferred to the judiciary, both at the national and at the international/supranational level, attention must be paid to their role in this regard.

The comparative analysis performed herein will show an important difference emerging from the comparison of the US case with the two European countries. This is essentially based on a different conception of what immigration, as an issue strictly related to the national sovereignty principle, should mean for the purpose of the division of powers.

In US, immigration is interpreted as a federal reserved power, allowing the federal authorities to intervene not only in the area of the regulation of the

[5] It is certainly disputable to apply the findings made above in the text to the case of Dutch, since this is currently the language spoken by the majority of the people in Belgium. However, Flemish finds itself in a somewhat unfavourable position in respect to the French language, spoken by the minority in the State, and in any case, Flemish is a minority language in the Brussels region.
With regard to the concept of national minority as applied to the Belgian case, it is very useful to read the *Opinion of the Venice Commission on possible groups of persons to which the Framework Convention for the protection of national minorities could be applied in Belgium* (Venice Commission, 8-9 march 2002) where Dutch speakers are considered as a national minority in the sense of the Framework Convention in the French-language region, but not at the State level.

entry and the stay of aliens, but also in the area of the rights and duties that the aliens enjoy, even if this could lead to encroachment upon matters normally reserved to sub-national units. The justifications for this wide federal power are based upon the fact that the legal treatment of aliens may become an issue involving international liability, or, at least, potentially affecting international relations, as well as the fact that integrating people in a common nation is essentially a national interest. As a consequence, it is manly up to the national legislator to shape the 'immigration federalism'. Sub-national units are allowed to enact measures in this field as long as federal law does not intervene by extensively regulating the matter or as long as sub-national actions do not stand as an obstacle to objectives laid down by federal law.

On the contrary, both the Italian and the Belgian Constitutional Courts have refused to conceive of the legal status of aliens as an autonomous standing power clause that enables the national legislator to intervene in areas otherwise reserved to the sub-national units. As a consequence, sub-national units have had more chances to freely develop immigration-related policies, potentially undermining coherence at national level. However, judicial enforcement of the equality principle and the protection of fundamental rights emerged as a means for the judiciary to guarantee a certain territorial harmonisation, as well as representing a means of combating discriminatory approaches by the sub-national units.

Both models present some inconsistencies.

In the final part of the essay, we suggest the development of cooperative federalism as an alternative means of structuring territorial relations within immigration. This solution seems more consistent with the idea that immigration is not, in itself, a jurisdiction, but rather a policy composed of measures falling under several different constitutional jurisdictions, which are vested in both the national and the sub-national tiers of government.

II. THE DIVISION OF POWERS ISSUE

Traditionally, the division of powers between national and sub-national units in relation to the immigration field has been essentially based on the idea that, whereas conditions of entry and residence of immigrants should be dealt exclusively by the national authorities, both the treatment of aliens, once legally admitted, and their integration are issues to be dealt

with by the sub-national units.[6]

This dividing line has shown some inadequacies. On the one hand, it prevents sub-national units from expressing their views on issues related to immigration that are crucial for their public interests (for instance the admission of immigrants according to the labour needs of the relevant regional territory). On the other hand, it does not consider that the legal status of aliens with regards to the rights that they enjoy (especially welfare rights) constitutes part of any integration project that the national level has an interest in shaping.

The aim of this paragraph is to provide the reader with some references to the division of powers issue and its relation to immigration policy in the relevant legal orders.

1. *USA*

The US Constitution does not explicitly grant the federal legislator any powers in relation to immigration as such.[7] The only constitutional clauses that may be considered as generally referring to the field are the naturalization clause of art 1 (8) sec 4, and the migration clause contained in art 1 (9) sec 1. This provision, however, has a specific historical reason, namely the compromise reached at the beginning of the XIX century to allow the southern states to maintain slavery. Thus, the reference to migration included in this clause does not have any current legal effects.

Lacking a federal statute dealing with immigration, and with no specific constitutional clause reserving immigration to the federal level, many member states in the XIX century enacted measures in order to regulate

[6] See Tomas Hammar, *Democracy and the Nation State* (Averbury 1990) regarding the distinction between 'policies of immigration' – to be reserved to the national level – and 'policies for immigrants' – to be reserved to the sub-national level.

[7] See, generally, Thomas Alexander A Aleinikoff, David A Martin, Hiroshi Motomura and Maryellen Fullerton, *Immigration and Citizenship* (Thomson/West 2008). On the immigration federalism see also Peter J Spiro, 'The States and Immigration in an Era of Demi-Sovereignities' (1994) 35 Virginia J Intern L 121; Cristina M Rodriguez, 'The Significance of Local in Immigration Regulation' (2008) 106 Michigan L Rev 567; Clare Huntington, 'The Constitutional Dimension of Immigration Federalism' (2008) 61 Vanderbilt L Rev 787, who deems immigration as a *de facto* shared power; Peter H Schuck, 'Taking Immigration Federalism Seriously' [2007] U Chicago L Forum 57. For critical views on the role of the member states in the immigration policy, see Michael J Wishnie, 'Laboratories of Bigotry? Devolution of the Immigration Power, Equal Protection, and Federalism' in (2001) 76 NY U L Rev 493; Linda S Bosniak, 'Immigrants, Preemption and Equality' (1994) 35Virginia J Intl L 179; Hiroshi Motomura, 'Immigration Outside the Law' (2008) 108 Columbia L Rev 2027.

immigration. [8] Usually, these were restrictive measures with a discriminatory effect towards newcomers.

Despite the lack of a specific provision granting the federal legislator the relevant power, in 1849 the Supreme Court struck down two statutes from New York and Massachusetts, which imposed a levy upon foreigner passengers landing in their ports. The Supreme Court based its decision on the commerce clause, thus considering these measures as a restriction to interstate commerce.[9]

In a subsequent decision, the Supreme Court explicitly stated that the regulation of the entry and the stay of aliens in the national territory is 'an incident of sovereignty belonging to the government of the US'. As a consequence, the Court granted the federal legislator a broad power to deal with immigration issues, and exercised the greatest judicial deference with regard to its political decisions in the field (so-called *plenary power doctrine*).[10]

Finally, in *Chy Lung v Freeman,* the Supreme Court derived the federal power to deal with the field of immigration from the federal powers related to foreign affairs. According to the Supreme Court, immigration is a matter to be vested in the federal legislator, since the legal treatment of aliens may become a reason of concern for the international legal order and thus the federal government may be called upon to answer for it.[11]

The aforementioned US Supreme Court decisions clarified that immigration is, in principle, a matter reserved to the federal level. However, it was not clear whether the federal immigration power was limited to the core meaning of immigration policy – ie the conditions for entry and stay and the removal of aliens – or whether it extended to cover other immigration-related areas, such as the legal status of aliens once legally admitted (the so called *alienage law*). Moreover, even admitting that the federal level has the power to deal with immigration, would it follow that any state is precluded from enacting provisions involving the classification of aliens? Or, would a state statute be vitiated insofar as it is effectively incompatible with the federal law?

These questions were answered years later, at a time when the federal

[8] See Gerald L Neuman, 'The Lost Century of American Integration Law (1776-1875)' (1993) 93 Columbia L Rev 1833; Hiroshi Motomura, *Americans in Waiting. The Lost Story of Immigration and Citizenship in the United States,* (OUP 2006).
[9] See *The Passenger cases*, 48 US (7 How) 283, 512-513, 12 L.Ed, 702 (1849).
[10] *Chae Chan Ping v US* 130 US 581, 604 (1889)
[11] *Chy Lung v Freeman*, 92 US 275, 23 L.Ed. 550 (1875).

legislator had already approved a comprehensive regulatory scheme on immigration law, namely the *Immigration Nationality Act* (INA).[12]

In 1971, in *Graham v Richardson*, the Supreme Court struck down two state statutes by means of which Pennsylvania and Arizona had purported to limit the award of state welfare entitlements, respectively only to American citizens resident in Pennsylvania, and to American citizens and aliens, providing that the latter had at least a 15-year legal residence in Arizona.[13]

The US Supreme Court decision relied on both an equal protection and a pre-emption analysis.
As to the pre-emption analysis, the Supreme Court held that the two state acts were pre-empted by the INA regulatory scheme insofar as 'state laws that restrict the eligibility of aliens for welfare benefits merely because of their alienage conflict with these overriding national policies in an area constitutionally entrusted to the federal government'. According to the Court, once aliens are admitted into the United States, they have a right to enter and reside in any state. Thus, any state measure that may place a burden on them, so to discourage them from settling in a given state, is ergo unconstitutional.

This decision seems to suggest that the federal power in the field of immigration covers immigration in its narrow sense as well as the legal status of aliens. State law is admissible only insofar as it does not hamper relevant federal law.

This issue was later reconsidered by the Supreme Court, which refined the pre-emption test to be used.

In *De Canas v Bica*, the Supreme Court upheld a Californian statute, which made it a civil offence for an employer to hire illegal aliens.[14] The Supreme Court based its reasoning on the distinction to be made based on whether the measure challenged falls under immigration law or under alienage law.

The facts that aliens are the subject of a state statute does not make it a

[12] See the Immigration Nationality Act 1952, now codified in Title 8 of the United States Code.
[13] *Graham v Richardson*, 403 U.S. 365, (1971).
[14] *De Canas v Bica*, 424 US 351, (1976). At that time, no sanctions were imposed by the federal legislator in cases where an employer hired illegal immigrants. As a consequence of the Supreme Court decision, in 1986, an amendment to the relevant federal legislation was inserted so to provide for such sanctions. This had the effect of precluding the measures of state legislators in the area.

regulation of immigration which is essentially a determination of who should or should not be admitted into the Country and the condition under which a legal entrant may remain', the Court stated. It also added 'the Court has never held that every state enactment which in any way deals with aliens is a regulation of immigration and thus *per se* pre-empted by the constitutional power, whether latent or exercised.

The federal legislator may, in principle, deal with immigration and alienage as well. What seems to change is the pre-emption test applied in order to determine whether the state measure is legitimate or whether it is interfering with the federal powers.

According to the Court, the federal law may pre-empt state acts in the immigration field in three different ways.

The first of these occurs when the state statute tries to regulate the conditions for entry and residence in the country. Any state regulation falling into this area is *per se* unconstitutional, whether the federal legislator has legislated or not.

The second instance occurs when the Congress intended to 'occupy the field'. This may occur when there is intent to ouster the state intervention. This intent can be expressed or inferred by the pervasive nature of the federal regulation, which does not leave any room for concurring state interventions.

Finally, the third instance (the so-called *conflict or implied pre-emption*) occurs when the state law 'stands as an obstacle to the accomplishment and execution of the full proposed and objective of Congress'.

It may be said that *De Canas* excluded that the INA had 'occup[ied] the field' in relation to both immigration and alienage laws. This means that, apart from the few cases where a state regulation attempts to substitute federal regulation concerning aliens' entry and stay in the country, the pre-emption test to be applied should be the conflict/implied one: the conflict between a state statute and a federal statute should be evaluated in concrete terms, favouring the best interpretation for the safeguarding of both statutes.

In *De Canas*, the Court applied the conflict pre-emption test. The Court recognised that California intended to pursue a legitimate aim, namely to deter irregular immigrants from entering its territory. The Supreme Court deemed that the federal law did not prevent a state from imposing a sanction upon an employer who hires illegal immigrants, thus upholding

the Californian statute.

In the light of *De Canas* decision and the above three pre-emption tests, it is useful to evaluate the legitimacy of several statutes, recently passed by some US states, which aim at deterring the presence of illegal immigrants.

To this extent, we may note that in the debate concerning immigration federalism, there has been a policy focus shift: whereas, until 1996, member states have concentrated their efforts upon the issue of immigrants as beneficiaries of state welfare benefits, in the last few years, they have been more inclined to take measures against illegal immigration.[15]

This goal has been pursued through an increase in limitations to possibilities for people to enter into a contractual agreement with illegal aliens, and via measures allowing state and local police to stop and detain illegal immigrants in order to facilitate their removal, which is a federal procedure.[16]

The *Arizona Support Our Law Enforcement and Safe Neighbourhoods Act* fits this scheme and has represented a model for other state measures in the field.[17] On the one hand, this statute makes it a misdemeanour for a person

[15] See Michael J Wishnie, 'Welfare Reform after a Decade: Integration, Exclusion, and Immigration Federalism', in Michael E Fix (ed) *Immigrants and Welfare – The impact on Welfare Reform on America's Newcomers* (Russel Sage Foundation - Migration Policy Institute 2009) 69.

[16] See Rodriguez, 'The Significance of Local in Immigration' (n 7) 591-592, who makes a distinction based on whether the State or the local unit measure directly or indirectly enforces the federal removal procedure of illegal aliens.
It must be noted, however, that cases of states or local authorities maintaining opposite policies occur, thus favouring illegal immigrant communities settled in their territories. This is the case with the so-called *sanctuary law*. As a reaction to states and local units enacting measures which forbade state or local public officers to pass information to federal immigration officers on the aliens they contacted by reason of their office, the federal legislator introduced a provision in the *Illegal Immigration Reform and Immigrant Responsibility Act* (IIRIRA) – now codified in Title 8 USC 1373 a) – forbidding state and local authorities from passing such measures. In spite of this, many states and local units maintained their policies. For an overview of the relevant measures still in force, see *Laws, Resolutions and Policies Instituted across the US Limiting Enforcement of Immigration Laws by State and Local Authorities*, (updated until 2008), available at <www.nilc.org>accessed 02 January 2013.

[17] Among others, see in Alabama the *Beason-Hammon Alabama Taxpayer and Citizen Protection Act*, Ala. Laws. Act 2011-535, House Bill (H.B. 56); in *Utah, the Illegal Immigration Enforcement Act*, House Bill 497 and the *Utah Immigration Accountability and Enforcement Act*, House Bill 116; in Georgia the *Illegal Immigration Reform and Enforcement Act* del 2011, House Bill 87.

to be unlawfully present in the US and knowingly apply for work, solicit work in a public space or perform work as an employee or independent contractor in the state. On the other hand, it provides state and local authorities with the duty of determining the immigration status of the person they stop, detain or arrest, whenever a reasonable suspicion exists that the person is an unlawful immigrant in the US. It also enables police officers to arrest a person when they have probable cause to believe the arrested person has committed any public offense that renders that person removable from the US.

The INA presents some inconsistencies as to if and when states can perform any action in enforcing the federal removal procedure. It explicitly provides for the possibility of deputizing state or local authorities' officers to perform the functions of federal immigration officers, provided that states or local units enter into written agreements with the Federal Attorney General. The agreements are intended to set the legal framework according to which functions are to be performed.[18] Moreover, the 8 USC sec 1252 lett c) provision expressly authorises state and local officials to arrest unlawfully present aliens only after confirmation of their illegal status by the competent federal authorities and in the limited cases the alien has re-entered the Country after leaving it or being deported following the commission of a felony.

A systematic reading of these provisions seems to suggest that states have no authority to enforce federal immigration law unless they act in pursuance with the 8 USC sec 1252 lett c) provision or within the conditions provided by a written agreement passed with the federal executive.

However, the INA also provides a saving clause according to which the lack of a previous written agreement between a state and the Attorney General does not prevent that state '(from) co-operat[ing] with the Attorney General in the identification, apprehension, detention or removal of aliens not lawfully present in the US'.[19]

Within this statutory framework, we shall now consider the case of the Arizona statute.

Following a suit brought by the US federal government to enjoin the

[18] This legal framework was introduced by the 1996 *Illegal Immigration Reform and Immigration Responsibility Act*. It is now codified in the Title 8 of the US Code sec 1357 lett g), 1-9.
[19] See art 8 U.S.C. para 1357 g), (10).

Arizona statute before it took effect, the Ninth Circuit Court of Appeal considered the Arizona provisions unconstitutional.[20]

The petition for a *writ of certiorari* submitted by the Arizona governor against the Ninth Circuit Court of Appeal decision was accepted by the Supreme Court, which delivered its decision by way of a majority.[21]

The opinion of the majority begins by recalling that the national legislator undoubtedly has a broad power over immigration and alien status, based on the Constitutional naturalisation clause (art I (8) sec 4 US Constitution) and on the inherent sovereign powers to control and conduct foreign relations. At the same time, the Court admits that states, too, have an interest in deterring illegal immigration, especially when this causes effective public order concerns, as it is the case in Arizona. State measures in this field can be justified on the grounds of state policing powers, which are reminiscent of their original sovereign status.

However, the fact that the police powers are an inherent component of state sovereignty does not alter the pre-emption test usually applied when determining whether or not federal law must supersede state law. This is the main point of contrast between the majority opinion and some of the dissenters' opinions. According to Justice Scalia, the fact that the Arizona statute is an exercise of the still-existing inherent state sovereign powers implies that Federal immigration law may pre-empt state law only if it is expressly declared to do so.[22] Explicit pre-emption is then required. On the contrary, the majority of the Court deemed that even in the area of illegal immigration, state measures may implicitly be pre-empted by federal law whenever they stand as an obstacle to the accomplishment and execution of the full purposes and objectives of Congress.

Hence, the Court held that most of the Arizona statutory provisions conflicted with the relevant federal law. In relation to the provision allowing state officers to detain illegal immigrants, the Court highlighted that, as a general rule, it is not a crime for a removable alien to remain in the United States. Removal is a civil matter where a broad discretion is accorded to the federal administration, which may decide whether to pursue it or not. According to the Court, the Arizona statute is

[20] *U.S. v Arizona*, 641 F.3d 339 (9th Circuit, 2011).
[21] See *Arizona et Al. v United States*, 567 U.S. (2012) 25 June 2012.
[22] See Justice Scalia's dissenting opinion, para 8 'We are talking about a federal law going to the core of state sovereignty: the power to exclude. Like elimination of the States' other inherent sovereign power, immunity from suit, elimination of the States' sovereign power to exclude requires that "Congress ...unequivocally expres[s] its intent to abrogate" ... Implicit "field pre-emption" will not do'.

illegitimate, since it encroaches upon the wide discretion entrusted to the federal government insofar as it authorizes state officers to decide whether or not an alien should be detained for removal.

The decision seems to be in line with the previous well-settled line of cases where the Supreme Court accorded a broad margin of discretion to the federal legislator when dealing with both immigration and the legality of aliens. Although the Court admits that member states may intervene in the field of immigration in its narrow meaning, as a consequence of their police power, and although it states that 'the historic police powers of the states' are not superseded «unless that was the clear and manifest purpose of Congress»', it finally held that federal law pre-empts state law, even where it is not explicitly stated that it should do so. Consequently, a state autonomous power in dealing with the detention of illegal immigrants for the federal removal procedure is inadmissible. States may act in this area only within the limits prescribed by the written agreements that they conclude with the federal government, in pursuance of the relevant provisions established by the INA.

Thus, only mechanisms of cooperative federalism may allow States to have a role in the immigration policy, as it is narrowly considered, notwithstanding the fact these mechanisms have not proved to be effective.

2. *Belgium*

Before taking into consideration the division of power issue concerning immigration policy in Belgium, it is germane to briefly outline some of the salient features of this federal system. According to art 1 of its Constitution, Belgium is a federal State made up of Communities and Regions. The former are the Flemish Community, the French Community and the German-speaking Community, the latter are the Walloon Region, the Flemish Region and the Brussels Capital Region, which are superimposed upon the three Communities. Both Communities and Regions are granted legislative powers.

The Constitution also recognises the existence of four linguistic regions, the territorial boundaries of which are set in a legislative act that can be amended only according to a special legislative procedure requiring a qualified majority.[23] The linguistic regions, to which any local Belgian municipality belongs to, are: the French-speaking region, the Dutch-speaking region, the German-speaking region and the bilingual region of

[23] See art 4 of the Constitution, to be read in conjunction with arts 2-8, *Lois coordonnées le 18 juillet 1966 sur l'emploi des langues en matière administrative*.

Brussels Capital.

Linguistic regions are not public entities but merely territorial delimitations serving to determine the use of the official language in the relevant territory and to delimitate, to a certain extent, the territorial scope of application of the Communities' measures, as we shall soon specify. The rule of monolingualism applies in the three single language regions while the rule of bilingualism applies in the Brussels Capital Region.[24]

Currently, the powers vested in the Regions and Communities are enumerated and listed partly in the Constitution, and partly in the 1980 Special Act on Institutional Reform (hereinafter the 1980 Special Act).[25] Accordingly, the federal legislator retains legislative powers in all matters that have not been expressly conferred to the Regions and the Communities, as well as in those other areas expressly reserved to the federal legislator by both the Constitution text and the 1980 Special Act.[26]

The three Regions exert their competences according to a territorial principle. This means their measures apply only within the relevant regional territory as defined in the Constitution and in the 1980 Special Act.

As far as the three Communities are concerned, their competences are prescribed by the Constitution at arts 127, 128, 129 and 130, and are further defined by the 1980 Special Act. The exercise of these powers is not exclusively based upon a territorial principle. In fact, all three Communities exert their powers in the territory of the three unilingual linguistic regions in the field of linguistic, cultural, educational and personal-related matters. However, in the Brussels bilingual region, both the Flemish and the French Communities are competent in relation to the

[24] A few municipalities (so-called *communes à facilité ou à statut (linguistique) special*) have special regulations with a view to protecting their minorities, enabling them to use an official language other than that of the language zone in which the municipality stands.

[25] In relation to the Région de Bruxelles capitale and to the German-speaking Community, the powers of these sub-national units are established, respectively, by the *Loi spéciale du 12 janvier 1989 sur les institutions bruxelloises,* and by the *loi de réformes institutionnelles du 31 décembre 1983*. In both cases the two acts make substantially reference to the division of powers established in the 1980 Special Act.

[26] This applies as long as art 35 of the Belgian Constitution would be effectively implemented.

above-mentioned matters.[27] Since in Brussels no sub-nationality exists in order to determine who is Dutch-speaking and who is French-speaking, the two communities are competent to act with regard to any institutions – and not with regard to the individuals – which, either by virtue of their activities or of their organisation, are deemed to belong to one of the two communities.

Returning to the matter of powers of the Belgian sub-national units in the immigration field, we should consider now the provisions of the 1980 Special Act.

Accordingly, art 5 (II) 3 grants the Communities the legislative powers in relation to the reception and integration of immigrants, and art 6 (IX) 3 grants the Regions the executive powers vis-à-vis the issuing of work permits to immigrants.

These are the only provisions explicitly giving Regions and Communities powers in the field of immigration policy, but other areas may also be relevant for our study. In this regard, it should be recalled that Communities are granted powers in relation to matters related to the person, and that Regions are competent in the field of housing, including social housing. When dealing with these issues, Regions and Communities may also include aliens in the personal scope of their measures.

We will now consider the explicit and the implicit powers of the Belgian sub-national units regarding immigration policy.

a. Explicit Powers Regarding Immigration Policy

As previously noted, the 1980 Special Act explicitly confers upon the Communities the power to deal with the reception and integration of immigrants.[28] In 1993, the French Community shifted these powers, respectively, to the Walloon Region and to the *Commission Communautaire Française* that exerts the relevant powers in relation to the Brussels bilingual region.

The attribution and the exercise of such competences were not a source of problems until 2003, when the so-called '*inburgering*' decree was approved

[27] Art 129 of the Belgian constitution grants the federal legislator the power to deal with language matters in the Brussels Capital region and in some municipalities with special linguistic status.
[28] See Art 5 (II) (3°) 1980 Special Act.

by the Flemish community.[29]

The other power within immigration policy that the Belgian federal legal system grants to its sub-national units relates to the issuing of work-permits by Regions.

Indeed, regional intervention in this field has represented a source of problems.

According to a federal 1999 statute, the regional authorities are responsible for the delivery of authorisation for employment. Before delivering such an authorisation, a thorough examination of the regional employment market must be carried out, in order to verify that no Belgian national is suitable for employment in the position in question. Moreover, requests for an employment authorisation and work permit are examined only if a third-country national comes from one of the countries having a bilateral agreement with Belgium. However, such conditions do not apply to certain categories, such as highly skilled workers.

It is then up to the regional authorities to verify whether the different conditions, required prior to the issuing of the working permit have been fulfilled. Moreover, the competent regional Ministry of Labour can deliver the work permit even in cases where the above-mentioned nationality and labour force conditions are lacking for certain special social reasons.

The enforcement of this power brought about significant regional differentiation. As a consequence, in 1992/1993, on the occasion of a wide reform of the federal Constitution and of the 1980 Special Act, the idea of re-federalising the relevant power emerged, so as to have a more uniform application. This did not occur, but the 1980 Special Act was amended in order to oblige Regions to conclude a cooperation agreement with the federal authorities before they exercise their competence.[30]

The parliamentary *travaux préparatoires* concerning the drafting of the

[29] We will take the *inburgering* decree into further consideration in the final part of this paper, when addressing the cultural regional citizenship issue.

[30] According to some scholars, the requirement of a binding cooperation agreement can be seen as an alternative way to achieve re-federalisation of the power in those cases where the exercise of power by the federated units, has resulted in a markedly uneven application. See Johanne Poirier, 'Le droit public survivra-t-il à sa contractualisation ? - Les cas des accords de coopération dans le système fédéral belge' [2006] Rev Dr ULB, 261, 278-279. See also Hugues Dumont, 'L'Etat Belge résistera-t-il à sa contractualisation – Considérations critiques sur le mode belge des accords de coopération' [2006] Rev Dr ULB, 315.

(inserted) art *92 bis* lett c) provision are particularly meaningful. The need for a coordinating framework among the different territorial levels is justified by the necessity of having a more effective implementation. The agreement should define a common socio-economic pattern to be considered by the Regions when delivering the work-permits, taking into consideration the establishment of an annual quota of work permits, if so required.[31]

Such an agreement was concluded between the regional and the federal authorities in 1995, but never officially published. However, scholars highlight that the agreement fell shorter the requirements demanded by the special legislation. This agreement provides only a common coordinating framework in order to develop common inspecting instruments. However, it does not supply any common reference for the definition of the socio-economic framework and of the yearly quota of migrant workers. Moreover, the agreement does not define provisions guaranteeing a uniform application of federal law. In short, the agreement is considered as unsuitable for the purposes of reaching any form of substantial harmonisation in the delivery of the work permits. Thus, an uneven application in this area continues, notwithstanding the approval of the agreement.[32]

b. Implicit Powers Regarding Immigration Policy: The Case of Welfare.

According to art 128 of the Belgian Constitution, Communities are granted powers in matters related to the person. In the parliamentary debates, these were intended to cover matters closely connected with the lives of

[31] See *proposition de loi spéciale visant à achever la structure fédérale de l'Etat*, Doc. Sen, 1992-1993, n. 558/5, pp. 453-454 : 'La coordination entre, d'une part, l'autorité fédérale qui est compétente pour délivrer des permis de séjour et pour déterminer les normes relatives à l'emploi de travailleurs étrangers et, d'autre part, les Régions qui sont compétentes pour délivrer des permis de travail peut être considérablement améliorée en imposant l'obligation de conclure un accord de coopération, reprenant, entre autre, les éléments suivants : a) le cadre socio-économique dans lequel les permis de travail peuvent être délivrés, avec fixation éventuelle d'un contingent ; b) dispositions assurant une application uniforme de la réglementation en matière de permis de travail sur l'ensemble du territoire, c) mesures visant à réaliser une application cohérente de la réglementation en matière de cartes de travail par rapport à la réglementation des permis de séjour ; d) la mise au point d'un système d'échange d'information, e) l'élaboration d'un système de control adéquat, entre autres, en vue de limiter l'application de la technique dite de rotation'.

[32] See Mieke Van de Putte and Jan Clement, 'Het migrantenbeleid', in Geert Van Haegendoren and Bruno Seutin (eds) *De bevoegdheidsverdeling in het federale België* (die Keure, 2000) 75, quoted by Jean-Thierry Debry, 'Les accords de coopération obligatoires' in [2003] Chroniques de Droit Public 209.

individuals in their communities.

A clear definition and listing of matters related to the person are not found in the Belgian Constitution, but are rather described in the 1980 Special Act. With regard to social matters, the 1980 Special Act accords competence relating to social security to the federal level, and that relating to social assistance to the Communities. However, even within the social assistance field, important branches have been arrogated to the federal level. According to art 5 of the 1980 Special Act, this is the case for some social assistance allowances, which apply generally to needy persons - such as the rights to subsistence income (originally *minimum de moyens d'existence*, in 2002 repealed and replaced with the *droit au revenue d'intégration*[33]) and the so-called *aide sociale*[34] - and of other social assistance benefits, which are tailored to specific categories of persons - such as the guaranteed income for the elderly[35] and the allowances for disabled persons.[36]

The federal attribution of legislative power, not only in the social security field, but also in large branches of the social assistance domain, was probably considered necessary to preserve a sense of common national belonging.

Due to the aforementioned division of competences, Communities have not played a relevant role in the field of welfare thus far. As a consequence, the issue of immigrants' eligibility for social entitlements has been primarily addressed by the federal legislator.

However, in 1999, the issue of welfare federalism was accorded new emphasis following the adoption of the Flemish Community decree concerning a care insurance scheme.[37] This is a universal insurance scheme for care dependency funded with lump-sum contributions. It is compulsory for anyone over the age of twenty-five who lives in Flanders.

---

[33] See *loi du 2 août 1974 instituant le droit à un minimum de moyens d'existence* repealed by *loi du 26 mai 2002 concernant le droit à l'intégration sociale*

[34] See *loi du 8 juillet 1976 organique des centres public d'aide sociale*. The "*aide sociale*" is a social assistance benefit aiming at enabling each person to conduct a life compatible with the principle of human dignity. It is not necessarily provided in form of stable financial aid. It may comprise medical or a material aid. Nonetheless, the practice was to provide it mostly in the form of a monetary sum equivalent to that of a subsistence income.

[35] See *loi du 22 mars 2001 instituant la garantie de revenus aux personnes âgées*

[36] See *loi 27 février 1987 relative aux allocations aux personnes handicapées*.

[37] *Décret du 30 Mars 1999* in *Moniteur belge* 28 mai 1999.

For residents in the Brussels area, this scheme is optional.[38]

This was the first time that a Belgian sub-national entity had supplemented the federal social security system with an entirely autonomous branch of social protection.[39] Because of this, the Flemish insurance scheme was a cause of serious concern for the French Community and the Walloon Region that saw the Flemish care insurance scheme as a threat to the federal social security system.

The Flemish decree raised objections on two different grounds.

The first concerned the issue as to whether the Flemish Community had the competence to deal with the matter. This relates to the fact that, as previously noted, in the Belgian legal system, the division of competences in relation to social matters is not clear: on the one hand, the federal level retains powers in the field of social security and in large branches of social assistance; on the other hand, Communities are vested with a general power to deal with social assistance.[40]

The second issue raised by the Flemish decree concerned its *ratione personae* scope of application. The Flemish decree originally applied only to persons residing in Flanders or in the Brussels Capital Region, in the latter case on a voluntarily basis. The choice of the place of residence criterion, instead of the place of employment, as a means of determining the persons to whom this decree applied, was deemed to better fit the principle of territorial exclusivity that underpins the Belgian federal system of the allocation of territorial powers. In line with the case law of the Constitutional Court, this principle implies that the object of any Community measure must be located within the territory for which that legislator is competent, in such a way as to exclude any potential extraterritorial effect. Because of this, residence was thought to represent the most suitable criterion. The residence criterion did not cause problems vis-à-vis the internal territorial distribution of powers. However, it did it in relation to European Union law, since the Flemish care insurance scheme

[38] See the volume Bea Cantillon, Patricia Popelier and Ninke Mussche (eds), *Social Federalism: the Creation of a Layered Welfare State – The Belgian Case* (Intersentia 2011).
[39] The point is highlighted by Bea Cantillon, 'On the Possibilities and Limitations of a Layered Social Security System in Belgium – Considerations from a Social Efficacy perspective', in Cantillon, Popelier, Mussche (eds), Social federalism (n 38) 72.
[40] See Jan Velaers, 'Social Federalism and the Distribution of Competences in Belgium', in Cantillon, Popelier, Mussche (eds), Social federalism (n 38) 137. See also Marc Joassart, Pierre Joassart, 'La Répartition des Compétences en Matière de Sécurité Sociale: L'Érosion Progressive de la Compétence Fédérale' [2006] Revue Belge de Droit cCnstitutionnelle, 167.

was considered to constitute a social security benefit in the sense of Regulation 1408/71 [41] which determines the applicable social security system to European Union workers exerting freedom of movement within the EU according to the *lex labor loci* criterion.[42]

Both issues were addressed by the Constitutional Court.

As to the first, in 2001, the Constitutional Court recognised the Flemish region's power to adopt the decree.[43] The Court took the view that in welfare matters, federal and Communities' powers are to be seen as parallel. Thus, a social scheme can fall either under the social security competence pertaining to the federal authority or under the social assistance competence of the Communities. The two can co-exist, no matter how they are financed.[44]

As to the second issue, following an ECJ Decision[45] on a preliminary ruling request issued by the Belgian Constitutional Court, this latter admitted that the Flemish care insurance scheme discriminated against those persons working in Flanders but living in the Walloon Region, insofar as they previously exerted their EU freedom of movement. The criterion of the *lex labor loci* could apply only to these cases.[46] However the Court recognised that the use of this criterion represented a derogation from the territoriality principle. It could be constitutionally accepted only insofar as

[41] Regulation (EEC) 1408/71/EEC of the Council of 14 June 1971 on the application of social security schemes to employed persons and their families moving within the Community [1971] OJ L149/2, repealed by Regulation (EC) no 883/2004 of the European Parliament and of the Council of 29 April 2004 on the coordination of social security systems, [2004] OJ L166/1.

[42] The Flemish authority had already introduced some amendments to the decree with the aim of rendering it consistent with the EEC 1408/71 Regulation, as a consequence of the beginning of an infringement procedure by the European Commission. See for further details, Steven Vansteenkiste, 'La Sécurité Sociale Flamande, Belge et Européenne – Aspects Juridiques de l'Assurance-Dépendance en Droit Belge et Européen', in [2004] Revue Belge de Sécurité Sociale, 35; Herwig Verschueren, 'La Régionalisation de la Sécurité Sociale en Belgique à la Lumière de l'Arrêt de la Cour de Justice Européenne Portant sur l'Assurance-Soins Flamande' [2008] Revue Belge de Sécurité Sociale, 173.

[43] See Belgian Constitutional Court decision n 33/2001. For a critical reading, see Xavier Delgrange, 'La Cour d'Arbitrage Momifie la Compétence Fédérale en Matière de Sécurité Sociale' in [2001] Revue Belge de Droit Constitutionnel, 216.

[44] See Velaers, Social Federalism (n 40) 138.

[45] Case C-212/06, *Government of the French Community and Walloon Region Government v Flemish Government*, [2008] ECR I-1683.

[46] These are EU citizens or the Belgian nationals, the latter, insofar as they returned back to Belgium after exerting their freedom of movement in another EU Member State. See Belgian Constitutional Court, decision 11/2009, para B 14.

it was imposed by EU law, and it applied to a small category of persons. The residence criterion was considered as the sole criterion that could satisfy the constitutional prohibition of extraterritorial effect of regional and Community acts.

Thus, the saga of the Flemish decree concerning the care insurance scheme made two things clear. First, the Communities are free to develop their own social security/social assistance scheme alongside the federal legislator. Second, the criterion the Communities must follow in order to determine the personal scope of the social assistance/security measures is residency. In fact, this is, according to the Constitutional Court, the only means of avoiding any otherwise impermissible extraterritorial effects of the relevant acts.[47]

These findings thus represent the framework within which Belgian welfare federalism may further develop. Within this context, the issue of aliens' eligibility to welfare entitlements from sub-national units may become more relevant. Even now, however, the issue is not being neglected by the federated legislators. The issue will be examined later in connection with the limits stemming from the principle of equality and non-discrimination.

c. The Powers Reserved to the Federal Authorities: The Case of the Legal Status of Aliens as an Autonomous Standing Power Clause

According to the division of powers to be found in the Belgian legal system, all powers not explicitly conferred to sub-national units are vested in the federal legislator. Thus, the federal State retains all powers concerning the conditions for entry, residing, and aliens removing[48].

Yet, we may wonder whether an explicit federal power to deal with the fundamental rights of aliens is provided for in the Belgian legal system.

To this extent, it should be noted that art 191 of the Belgian Constitution – a provision dating back to the 1831 Liberal Constitution – establishes the principle according to which aliens are substantively equal to Belgian nationals in the enjoyment of rights and freedoms, unless otherwise stipulated by the law (*loi*).[49] The explicit reference to the term *loi* gave rise

[47] According to Velaers, *Social federalism* (n 40) 138, the choice of residence can be further justified by considering that residence 'in person related matters is conducive to a system of solidarity between residents of the same region'.

[48] See *loi du 15 décembre 1980, sur l'accès au territoire, le séjour, l'établissement er l'éloignement des étrangers*.

[49] See Pascal Boucquey, 'La Cour d'Arbitrage et la Protection des Droits Fondamentaux de l'Étranger' in [1996] Annales de Droit de Louvain, 289.

to the problem as to whether Regions and Communities had the power to take measures dealing with alien rights and freedoms. The legislative acts passed by Regions and Communities are defined as decrees or ordinances, not *loi*.

The Constitutional Court held that in cases where the Constitution uses the term *loi*, it means that the constituent power wanted to reserve the relevant matter to the federal legislator, provided that the relevant provision was inserted after 1970, when the process of federation began. Otherwise – and this is the case of art 191 of the Belgian Constitution which was inserted in the 1831 Constitution – the term *loi* may refer to sub-national units' legislative measures, provided that the 1980 Special Act *expressly* and *precisely* assigns the relevant power to the federate units[50].

As far as immigration is concerned, the 1980 Special Act on institutional reforms, as noted above, grants the Communities and the Regions powers respectively in the field of the integration of immigrants and in the delivery of work permits, but it does not make any reference to their role in dealing with the legal status of aliens. Some authors have argued then that art 191 of the Belgian Constitution should be considered as a power clause reserving to the federal level the power to deal with the civil and social rights of the aliens.[51]

Such an interpretation has not been adopted by the Constitutional Court

[50] This is known in the Belgian legal system as '*la théorie des matières réservées*'. See Belgian Constitutional Court, decision 35/2003, para B(12)(6). See Mark Uyttendaele, *Précis de Droit Constitutionnel Belge – Regards sur un Système Institutionnel Paradoxal* (Bruylant 2005) 956-965.

[51] See Sébastien Van Drooghenbroeck, 'L'Article 191 de la Constitution' in [2006] Revue Belge de Droit Constitutionnel, 305, 309, who speaks of a 'gommage pur et simple de la dimension répartitrice de compétences de l'article 191', which he nonetheless justifies for systematic motives and reasons of coherence. However, the same author has mainly relied on this argument to deem the Flemish *inburgering* decree illegitimate. The author considers that the 1980 Special Act on institutional reform gave the Communities the power to enact integration measures for immigrants, provided that they were based on a voluntary scheme. Since the Flemish *inburgering* decree provides for compulsory integration measures, it falls outside the scope of the 1980 Special Act provisions, and the Flemish Community lacked the power to enact it. See Sébastien Van Drooghenbroeck, 'Fédéralisme, Droits Fondamentaux et Citoyenneté: Les Certitudes à l'Épreuve de *l'Inburgering*' in Eva Brems and Ruth Stockx (eds), *Recht en minderheden. De ene diversiteit is de andere niet* (Die Keure 2006) 257. See also for a similar view, Matthieu Lys, 'Les Droits Constitutionnels des Étrangers' in Marc Verdussen and Nicolas Bombled (eds), *Les Droits Constitutionnels en Belgique* (Bruylant 2011) 607.

and the Council of State.[52] As a consequence of the 1994 Constitutional amendments, which expressly call upon the sub-national units to deal with fundamental rights issues[53], the Constitutional Court has finally stated that fundamental rights protection does not constitute a jurisdiction, but rather an objective that all the territorial units are called to pursue within their relevant constitutionally entrusted jurisdictions.[54] This finding has certainly undermined the idea that the legal status of aliens with regard to the fundamental rights they enjoy could be considered as a matter that only the federal legislator is entitled to deal with *per se*.

3. *Italy*

The Italian Regions started dealing with issues related to immigration in the 1980s. Lacking a comprehensive immigration national statute, regional measures were especially directed towards the integration of migrants. This was carried out by regulating aliens' access to regional welfare and through the establishment of advisory bodies where migrants were represented. These measures were deemed to fall under the regional powers on social assistance, which at that time was an area of competence that was shared with the national legislator.

The attention regarding immigration issues acquired a new emphasis at the beginning of the new millennium, when the Italian regions were called to adopt their "*statuto di autonomia*". According to art 123 of the Italian Constitution, inserted in 1999, the "*statuto di autonomia*" is a regional act, to be passed with a special majority and procedure, which lays down the form of government and the basic principles for the organisation of each individual Region and the conduct of its business.

Thus, the "*statuto of autonomia*" is intended to have a specific content. Nonetheless, Regions considered their "s*tatuto of autonomia*" as a sort of political manifesto wherein to outline the general aims of the regional authorities. Within this framework, references were often made to regional immigrant integration, and to the extension of the right to vote at the local elections to immigrants.[55] This occurred despite the fact that the regulation of the right to vote at local elections is a matter reserved to the

---

[52] See Belgian Constitutional Court decision 62/98. For further references to the relevant case law see Van Drooghenbroeck, 'L'article 191' (n 50) 309-310.
[53] See for example arts 22 (2), 22 *bis*, 23 of the Belgian Constitution.
[54] See Belgian Constitutional Court decisions 124/99 and 124/2000.
[55] See Andrea Gentilini, 'Statuti e Leggi Regionali in Materia di Migrazioni', in *Osservatorio sulla Legislazione, Rapporto 2010 sulla Legislazione tra Stato, Regione e UE, Tomo II, Tendenze e Problemi della Legislazione Regionale* (Camera dei Deputati 2010) 199.

national level, and that a constitutional amendment is deemed necessary in order to allow immigrants to vote.

A claim was thus brought before the Constitutional Court by the national government, asking the court to declare void these and other similar provisions containing references to general political aims on that the grounds Regions were lacking any competences in these areas. The Constitutional Court rejected the claim, deeming that these provisions did not have any legal force (despite their inclusion in a statute) rendering them merely political commitments.[56]

Meanwhile in 1998, the national legislator passed a comprehensive statute on immigration law (hereafter, the 1998 Immigration Act)[57], dealing not only with the entry and stay of immigrants, but also with their civil and social rights.

According to the statute, the enjoyment of social rights is dependent upon the legal or illegal status of the immigrant and on the length of his permit of stay. More precisely, art 41 of the 1998 Immigration Act prescribes an equal treatment principle between nationals and immigrants in the enjoyment of social assistance entitlements, provided that the immigrant is legally present in the Italian territory with a permit of stay of at least one year. Art 42, which concerns access to social housing, provides for equal treatment between nationals and foreigners, provided that the latter are workers, legally present in the national territory with a permit of stay of at least two years. The Regions are explicitly called to collaborate with the State, and to act in the areas falling under their jurisdiction (mainly in the field of welfare) provided that they respect the provisions set out in the national law.

The 1998 Immigration Act reflected the system of the powers of territorial allocation as it was established, originally, in the 1948 Italian Constitution. According to the original text of the 1948 Italian Constitution, the Regions were granted only legislative powers in specifically enumerated matters. Moreover, the legislative powers that they were accorded were shared competences, meaning that the Regions could act only within the framework of, and with due respect accorded to, the fundamental principles established by the national legislator in relation to each shared

[56] See Italian Constitutional Court decision n 372/2004 in relation to art 3 (6) of the Tuscany *Statuto di Autonomia* which states 'la Regione promuove, nel rispetto dei principi costituzionali, l'estensione del diritto di voto agli immigrati'.

[57] See d lgs n 286/1998, *Testo Unico delle disposizioni concernenti la disciplina dell'immigrazione e norme sulla condizione giuridica dello straniero*, in GU n 191, 18 August 1998.

matter. Thus, as far as immigration policy is concerned, the specific requirements concerning welfare eligibility, laid out in arts 41 and 42 of the 1998 Immigration Act, were considered by the national legislator as fundamental principles, which could not be bypassed by the regional legislator.

The 1998 Immigration Act was more reluctant to provide a role for the regional actors vis-à-vis immigration policy, particularly with regard to the regulation of immigrants' entry and expulsion. However, in 2001, an amendment to the 1998 Immigration Act stated that the decree setting the number of immigrants to be admitted annually into the country could be adopted only following consultation with the Regions concerning the need for migrant workers at the regional level.

In 2001, the Constitution was amended in order to strengthen the Regions' powers. Art 117 of the Italian Constitution deals with the division of powers with regard to the legislative function. It provides two lists. The first enumerates those powers that are granted to the national legislature. The second deals with those matters – defined as concurrent – where regional measures can be enacted within the limits of the fundamental principles laid down by the national legislator. Finally, matters that are not enumerated in either of the two lists are vested in the Regions.

As to immigration policy, the 2001 constitutional amendment has somehow blurred the previously-established dividing line: on the one hand, the new constitutional provisions seem to confer full powers to national State in addressing not only immigration (art 117 (2) lett b Cost) but also asylum and the legal status of non-European citizens (art 117 (2) lett a Cost); on the other hand, Regions now dispose of full powers in matters such as housing and welfare assistance. The national State also retains the exclusive power to determine the basic level of benefits relating to civil and social entitlements to be guaranteed throughout the national territory (art 117 (2) lett m Cost).

Following the 2001 constitutional amendment, some Regions began passing legislative acts dealing with aliens' access to regional welfare, in some cases providing different rules than those established in the 1998 Immigration Act.

In 2005, the national government brought an action before the Constitutional Court, claiming that the Regions were not empowered to deal with aliens' access to regional welfare, since the new art 117 of the Constitution granted the national legislator all powers in relation to immigration policy.

The Constitutional Court nonetheless confirmed the legitimacy of the aforementioned regional measures.[58] It interpreted the two jurisdictions reserved to the national level, namely regarding immigration and the legal status of non-EU citizens, as constituting a single power, allowing the national level to deal only with the conditions of entry and residency. Therefore, the Court excluded that the "legal status of non EU citizens" (art 117, 2, lett a Cost) might represent a legal base allowing the national legislator to take measures in areas that would otherwise fall under the exclusive regional legislative competences, as is the case of social assistance and social housing. This allows Regions to deal with the eligibility of aliens for welfare entitlements, as long as this falls under their current powers.

To this extent, it is remarkable that the choices made by the Italian regional legislators, thus far, have not been always consistent with those undertaken by the national legislator. In some cases (Campania, Toscana), regional statutes provide public benefits to be eligible even to irregular immigrants, thus questioning the very idea that integration should concern only regular migrants. In some other cases, regional statutes provides for requirements that are stricter than those originally foreseen in the 1998 Immigration Act when awarding social assistance benefits or social housing.

These inconsistencies between regional and national law have been assessed by the Constitutional Court as primarily involving respect for the equality principle and fundamental rights.

III. THE 'SOCIAL REGIONAL CITIZENSHIP' DIMENSION

Social assistance is based upon the idea of collective solidarity. Because of this, it constitutes a useful field to be explored in order to verify the attitude that the sub-national units (and the State in general) have towards the newcomers. Are these considered as a part of the relevant regional community and thus beneficiaries of its social solidarity instruments?

'Social welfare citizenship' is, then, primarily a question of equality and non-discrimination, namely to what extent sub-national units may make the award of regional social benefits conditional on the grounds of nationality or because of long-term residence in the relevant sub-national territory.

To this extent, it is frequently submitted that after World War II,

[58] See Italian Constitutional Court decisions 300/2005 and 156/2006.

nationality, as a condition for the enjoyment of fundamental constitutional rights, has progressively lost importance.

Hence, fundamental constitutional rights are to be considered as pertaining to the individual as such, rather than to the citizen.[59]

Whereas this move is evident with regard to civil fundamental rights, things become more nuanced in relation to social assistance rights.[60] Their awarding has traditionally been made conditional upon the fact that the needy person possessed the nationality of the given State.

Within this context, the input emanating from the European supranational and international judicial institutions are significant.

Regarding the EU, the EC Treaty enshrined the principle of non-discrimination on the basis of nationality, to be applied in all areas falling within the Community's competences. This principle has been substantially applied to EU citizen workers (and their family members), exerting their rights to move and to reside in another EU Member State.[61]

It is the ECJ that has progressively recognised the right to move and reside freely within the territory of EU as a genuine independent right, inherent to the status of the Union citizen.[62] According to the ECJ, a Union citizen – even if non-economically active – who is lawfully resident in one of the host Member States, can rely on art 18 TFEU and may claims equal treatment in all situations which fall within the scope *ratione materiae* of EU law. However, asking for social assistance may ultimately result in his removal from the host country insofar as this may be considered as a proof that he does not possess sufficient economic resources and that he has become a burden for the welfare of the relevant State.[63]

---

[59] See in the Italian legal literature, Paolo Carrozza, 'Nazione' in *Digesto Discipline Pubblicistiche* (UTET 1995) vol 10, 787.

[60] Things are different for social security as long as this is based on benefits funded, at least in part, from the contributions of workers' earned income. No nationality requirements usually applies. See A Math, *La Protection Sociale des Ressortissants d'Etats Tiers dans l'Union Européenne. Vers un Citoyenneté Sociale de Résidence*, Institut de Recherche Économiques et Sociales (IRES), 3/2001, 4.

[61] See Regulation (EEC) 1612/68/EEC of the Council of 15 October 1968 on freedom of movement for workers within the Community [1968] JO l257/2, whose art 7 guarantees social security and social assistance to EU workers – and to their relatives - on equal terms with the nationals of the host Member States.

[62] See Case C-413/99 *Baumbast and R v Secretary of State for the Home Department* [2002] ECR I-70912, para 81.

[63] See Case C-85/96 *María Martínez Sala v Freistaat Bayern* [1998] ECR I-2691; Case C-456/02 *Michel Trojani v Centre public d'aide sociale de Bruxelles (CPAS)* [2004] ECR I-

Thus, the principle of equal treatment between EU citizens and the nationals of the host EU Member State has been finally extended to non-economically active EU citizens. This has not occurred in relation to third-country nationals, however. With reference to the latter, a principle of equal treatment with the nationals of the host State is foreseen only in relation to certain categories of qualified aliens, namely the long-term immigrant residents and the refugees or beneficiaries of international protection, and it does not apply generally but only in relation to the matters and according to the specific conditions set out in the relevant EU secondary law.[64]

7573, para 30. The conditions for the exercise of the right of EU citizens and their families to move and reside freely within the territory of the Member States are now set out in Directive 2004/38/EC of the European Parliament and of the Council on the right of citizens of the Union and their family member to move and to reside freely within the territory of the member State amending Regulation 1612/68/EEC and repealing Directives 64/221/EEC, 68/360/EEC, 72/194/EEC, 73/148/EEC, 75/34/EEC, 90/364/EEC, 90/365/EEC and 93/96/EEC [2004] OJ L158/77. According to the Directive, Union citizens are only entitled to reside in the host Member State for more than three months if they are either economically active or have sufficient resources for themselves and their family members not to become a burden on the social assistance system of the host Member State and have comprehensive sickness insurance cover in that State. Although the Directive provides a codification of previous secondary EU law and ECJ case law, there have been signs in the ECJ case law of possible tensions between the Directive and the interpretation of the relevant Treaty provisions (arts 18 and 21 TFEU). In fact, the Court seemed to suggest the need for a case-by-case assessment of whether the Union citizen constitutes an unreasonable burden, according to the proportionality principle. See Michael Dougan 'The Constitutional Dimension to the Case Law on Union Citizenship' (2006) 2 E L Rev, 613. However, some scholars have emphasized a possible change of approach in the latter ECJ case law. See Siofra O 'Leary, 'Free Movement of Persons and Services', in Paul Craig and Gráinne de Burca, *The Evolution of EU Law* (2nd edn, OUP 2011) who, by reference to the *Förster* case (case C-158/07 *Jacqueline Förster v Hoofddirectie van de Informatie Beheer Groep* [2008] ECR I-8507) notes: 'The Court's position in *Förster* is surprising not simply because it suggests that a directive can renege on or restrict the jurisprudential *acquis* established with reference to the Treaty, but also because it contradicts the Court's initial assessment of the relationship between Directive 2004/38 and the existing *acquis*'.

[64] With regard to the social assistance field, see arts 11 (1) lett d) and 11 (4) of the Directive 2003/109/EC of the Council of 25 November 2003 concerning the status of third-country nationals who are long-term residents [2003] OJ L16/44 and Article 28 of the Directive 2004/83/EC of the Council 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. See also Directive 2011/98/EU of the European Parliament and of the Council of 13 December 2011 on a single procedure for a single permit for third-country nationals to reside and work in the territory of a Member State and on a common set of rights for third-country workers legally

One may wonder whether directive 2000/43/CE, introducing the principle of non-discrimination irrespective of race and ethnic origin, could apply in case of discrimination based upon nationality, insofar as a condition based on nationality may turn out to constitute an indirect form of discrimination because of race and ethnic origins.[65]

To this extent, it should be noted that the directive itself excludes from its scope any classification based on nationality.[66] Recently, the ECJ has denied that directive 2000/43/CE could apply in case of legal classifications based on nationality.[67] The Court, however, explicitly considered neither the possibility that a nationality requirement could amount to a form of indirect discrimination because of ethnic origins, nor whether, in the light of the ECtHR case law, a prohibition of discrimination on grounds of nationality can be regarded as a general principle of the EU legal order.

Setting aside the EU law and the ECJ case-law, attention should be drawn to the ECtHR,[68] the European judicial body that pushes to the greatest degree toward considering nationality as an illegitimate criterion to judge eligibility for the receipt of social benefits.

It was in the *Gaygusuz v Austria* case[69] that for the first time the Court gave a broad reading to the notion of pecuniary rights for the purposes of art 1 of the Protocol Number 1 of the ECHR in such a manner so as to apply it to social benefits, whether contributory or not.[70] Once the Court

---

residing in a Member State [2011] OJ L343/1, in which the right to equal treatment, according to art 12, is strictly linked to the third country nationals' legal residence and to their worker status. Moreover, the right to equal treatment applies to specified fields whose scope of application Member States are empowered to limit within certain conditions.

[65] Directive 2000/43/EC of the Council of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origins [2000] OJ L180/22.

[66] Directive 2000/43/EC, art 3(2).

[67] Case C-571/10 *Servet Kamberaj v Istituto per l'Edilizia sociale della Provincia autonoma di Bolzano (IPES) and Others* (ECJ - 24 April 2012).

[68] See generally on the anti-discrimination principle set in the Convention and its recent evolutions: Oddny M Armadóttir, *Equality and Non-Discrimination under the European Convention of Human Rights* (Nijhoff Publishers 2003); Françoise Tulkens, 'L'Évolution du Principe de Non-Discrimination à la Lumière de la Jurisprudence de la Cour Européenne des Droits de l'Homme' in Jean Yves Carlier (ed) *L'Étranger Face au Droit* (Bruylant 2010) 193.

[69] *Gaygusuz v Austria* App no 17371/90 (ECHR, 16 September 1996).

[70] See *Koua Poirrez v France*, App no 40892/98 (ECHR, 30 September 2003); *Stec and others v UK*, App no 65731/01 – 65900/01, (ECHR, 12 April 2006). For comments and further bibliographic indications on these decisions, see Elise Dermine, Mikaël

had included the social benefits category into the realm of pecuniary rights, it could apply art 14 of the ECHR, which provides for non-discrimination in the enjoyment of rights and freedoms set forth in the ECHR, on several grounds, amongst which is included that of nationality. The Court then held that a difference in treatment grounded on nationality may be accepted only in narrow circumstances. According to the Court, 'very weighty reasons would have to put forward before the Court could regard a difference of treatment based exclusively on the ground of nationality as compatible with the Convention'.[71]

The above-mentioned ECtHR case law has exerted a notable influence in the European national Constitutional Courts, leading some of them to consider classifications based on nationality as intrinsically 'suspect' and thus mostly impermissible.[72]

Some questions still remain open to debate. The ECtHR did not expressly consider whether the strict scrutiny review also applies to classifications based on the different legal status of aliens. Thus, it might be submitted that the length of the immigrant stay in the host State could be a ground for the ECtHR to evaluate the legitimacy of a national measure limiting welfare entitlements to those aliens having a genuine link with the territory of the host State, such as long-term immigrants.

Thus, whereas nationality as such is increasingly being considered as a suspect criterion, a less conclusive statement can be made with reference to cases where the award of a social assistance benefit is made conditional upon the length of the legal stay of the immigrant in question.

Clearly, this framework influences the capacity of sub-national units to deal with the issue. Regional social assistance entitlements are generally provided to persons residing in the territory of the given sub-national unit. A social benefits limitation to the nationals residing in the given territorial unit would be inadmissible in the light of the aforementioned ECtHR case law. Thus, due to the tendency to consider nationality as a suspect criterion, those sub-national units wishing to limit social assistance benefits to their autochthonous communities may be pushed towards using a long-term residency requirement. However, if a durational residency requirement condition is generally applied to all individuals, thus

---

Glorieux, Steve Gilson, *Aperçu des Droits Sociaux des Étrangers en Belgique et Questionnements Actuels* in Carlier *L'étranger face au droit* (n 68) 549.

[71] See *Gaygusuz v Austria* (n 69) para 42; *Koua Poirrez v France* (n 70) para 46.

[72] See, for a representative example, decision 187/2010 of the Italian Constitutional Court.

including EU citizens, it may turn out to be an indirect form of discrimination against the latter, thus potentially breaching EU law.[73]

Within this framework, we will now consider the cases of the relevant States taken into consideration in our comparison.

1. *USA*

The case of welfare immigration federalism in the US legal order involves both a division of competences and an equal treatment issues. In the already mentioned *Graham v Richardson* case, the Supreme Court relied on both a pre-emption and an equal protection analysis to strike down two State statutes limiting the eligibility of aliens to welfare entitlements.

Focussing here on the equal protection analysis, the Supreme Court sets the principle according to which any State classification, based on *alienage*, is intrinsically suspected of breaching the XIV Constitutional Amendment, and is thus subject to the rigorous strict scrutiny test. According to the Supreme Court, legal immigrants are to be considered as 'a discrete and insular minority', since they do not enjoy the right to vote. Because of this, any *alienage* classification has to be considered as inherently suspect.[74]

The *Graham v Richardson* case led to the supposition that a strict scrutiny test could be applied even in those cases where an *alienage* classification had been enforced by the federal legislator. However, this was not the solution applied by the Supreme Court. In *Mathews v Diaz* the Supreme Court upheld a federal law restricting the eligibility of immigrants to welfare entitlements.[75] The Court here applied a rationale test and it explicitly stated that, as a consequence of the great discretion that federal legislator enjoys in relation to immigration field (the so-called plenary power doctrine), *alienage* federal classifications are not subject to the strict scrutiny test, while State *alienage* classifications are.

In the 1990s, the issue of welfare immigration federalism became salient

[73] Although traditionally the ECJ has considered residence as a suspect criterion, deeming it a way to indirectly discriminate against EU citizens exerting their right to free movement, in more recent years a line of cases emerge in which the Court held residence requirements to be valid. This has been applied especially in cases of social assistance benefits required by non-economically active EU citizens. See Case C-209/03 *The Queen, on the application of Dany Bidar v London Borough of Ealing and Secretary of State for Education and Skills* [2005] ECR I-2119 and Case C-158/07 *Förster* [2008] ECR I-8507 both referring to social benefit asked by EU students.

[74] *Graham v Richardson*, 403 U.S. 365, (1971).

[75] *Mathews v Diaz*, 426 US 67 (1976).

again following a Californian legislative proposition aiming at reducing welfare entitlements for irregular immigrants. The proposed bill was successfully challenged before a local federal court.[76]
In 1996, the federal legislator enacted the *Personal Responsibility and Work Opportunity Reconciliation Act* (PRWORA). The PRWORA was intended to constitute a comprehensive reform of the federal funded welfare programs, with the aim of providing the States with more discretion in dealing with their implementation.

One of the main changes introduced by the PRWORA concerns immigrants' access to federal and State-funded welfare programs.[77] The PRWORA establishes that legal permanent immigrants, who entered the US after 1996, are no longer eligible for most of the *federally*-funded programs unless they can fulfil a five-year legal residence requirement. However, States are empowered to modify the eligibility conditions in order to cover immigrants otherwise excluded by the federal assistance program. If they decide to do so, they must cover the costs.[78]
PRWORA also deals with the State capacity to define the conditions of immigrants' access to State-funded welfare programs. On the one hand, it explicitly authorises States to determine whether permanent immigrants are eligible to receive State welfare benefits. On the other hand, it prohibits States from providing welfare allowances to illegal migrants, unless the States themselves decide to explicitly derogate from it.[79]

The enactment of the PRWORA made it clear that the federal power on immigration issues is not limited to the entry and the stay of immigrants, but that it may also cover their legal status, even in areas otherwise reserved to States.

Moreover, the PRWORA has had the effect of allowing States to circumvent the limitations that the Supreme Court set in the *Graham* decision with regard to State *alienage* classifications. The PRWORA, then, shifted the political decision and the cost of discriminating against qualified aliens from national level to State level.

Following the PRWORA enactment, the federal courts took different views with regard to the effects produced by the federal statute on the State capacity to introduce *alienage* classifications in awarding State welfare

[76] See *League of United Latin American Citizen v Wilson*, 908 F. Suppl. 755 (C.D. Cal 1995).
[77] See Wishnie 'Welfare Reform after a Decade' (n 15) 69.
[78] The PWORA has been codified in the US Code. The relevant provisions we refer to in the text are now contained in Title 8 of the USC paras 1612-1613
[79] See Title 8 U.S.C. 1622 and 1624.

benefits. According to some decisions, due to the great discretion the federal legislator enjoys in the immigration field, it may decide to structure the immigration policy in such a manner as to leave it up to the States to decide whether to discriminate or not against immigrants when awarding social assistance benefits.[80] On the contrary, others posit the opposite, deeming that the PRWORA cannot bypass the Supreme Court's position in the *Graham* decision.[81]

As a matter of fact, relatively few States passed laws restricting qualified aliens' welfare entitlements, and generally accepted to supplement the associated costs, but a change of policy is likely to take place due to the recent budget crisis some of the States have to cope with.[82]

2. *Belgium*

The Belgian Constitutional Court has scrutinised the constitutionality of federal statutes restricting aliens' access to welfare entitlements several times, mostly for supposed breaches of the equality principle. The fact the relevant legal classifications were federal is consistent with the leading role played by the national legislator in welfare policy, as we noted earlier in the text. Nevertheless, we will shortly refer to these decisions, since they can give important insights concerning the limits stemming from the equality principle in relation to the subnational units' activity.

The Belgian federal legislator provides for a substantial division in relation to the eligibility of aliens to social assistance benefits.[83] Whereas the so called *aide sociale* is granted to the individual as such,[84] thus including immigrants,[85] other social assistance measures (ie the right to subsistence income,[86] guaranteed income for the elderly,[87] allowances for the disabled

[80] cf *Soskin v Reinertson*, 353 F.3d 1242 (10th Cir. 2004)
[81] cf. *Aliessa v Novello*, 96 N.Y.2d 418, (N.Y. 2001).
[82] See Wishnie, 'Welfare Reform after a Decade' (n 15) 69-70.
[83] For a short overview of the main social assistance measures provided for the Belgian legal system, see part II, B of this paper.
[84] See art 1 of *loi du 8 juillet 1976 organique de centres public d'aide sociale*.
[85] Only legal immigrants are entitled to full *aide sociale*, whereas in case of illegal migrants, only medical care is provided. The case of the asylum seekers is different. They are entitled to receive a material form of assistance, insofar as they reside in a federal centre during the period necessary to define their status as refugees or other beneficiaries of international protection.
[86] See art 1 (1) *loi du 2 aout 1974 instituant le droit à un minimum de moyens d'existence*, This statute was repealed by *loi du 26 mai 2002 concernant le droit à l'intégration sociale* whose art 3.3 extended the right to subsistence income to registered long-term immigrant residents.
[87] See art 4 of *loi du 22 mars 2001 instituant la garantie de revenus aux personne* âgées. This social benefit has not been extended by the legislator to the registered long-term

persons[88]) have been primarily reserved to Belgian citizens and to those aliens that can be assimilated to the status of nationals such as EU workers, refugees or stateless persons. More recently, these social benefits – with the exception of the guaranteed income for the elderly – have been extended to long-term immigrants, in part as a consequence of decisions of the Belgian Constitutional Court.

The Belgian Constitutional Court has assessed the constitutionality of these limitations based on the grounds of nationality on a number of occasions.

Generally, the Constitutional Court has been keen on according a more generous reading to the personal scope of the *aide sociale*, in some cases extending it to categories of immigrants not originally included.[89]

As far as the other social assistance benefits are concerned, the court has taken the view that the equal treatment principle between the Belgian citizens and the third-country nationals should apply only in relation to registered long-term immigrant residents,[90] and not in relation to the other categories of immigrants who are legally present in the State.[91]

---

immigrant residents. The Belgian Constitutional Court has considered this exclusion as legitimate (69/2010).

[88] See art 4, *loi 27 février 1987 relative aux allocations aux personnes handicapées*. In 2009, following a Constitutional Court decision - 153/2007 - (see later the text), art 4 has been amended to cover registered long term immigrant residents.

[89] See for example decision 106/2003, concerning the awarding of the *aide sociale* to non-accompanied minors. Further references in Hugo Mormont and Katrin Stangherlin (eds) *Aide Sociale – Intégration Sociale* (La Charte 2011) 117.

[90] These are the so-called *étrangers établis ou les étrangers qui sont inscrits au registre de la population*. According to the relevant provision of the 1980 immigration national statute, an alien may request this status after 5 years of regular and continuous stay in the Belgian Kingdom, provided that public order or national security reasons do not oppose. The aliens, who the status of *étranger établi* has been recognised to, is inserted in the register of the general population (*registre de la population*) whereas the legal immigrant is listed in a different register (*registre de la population étrangère*).

[91] See Belgian Constitutional Court decision n 5/2004, in relation to the right to subsistance income, para B (6) (3): 'Il existe une différence entre les étrangers qui sont autorisés à s'établir dans le Royaume et les étrangers qui sont autorisés à y séjourner pour une durée limitée ou illimitée. [...] Le critère de "l'autorisation d'établissement dans le Royaume", qui ressort de l'inscription au registre de la population, est pertinent par rapport à l'objectif de promouvoir l'intégration sociale des personnes résidant en Belgique. Il n'est pas déraisonnable, en effet, que le législateur réserve les efforts et moyens particuliers qu'il entend mettre en œuvre en vue de réaliser cet objectif à des personnes qui sont supposées, en raison de leur status administratif, être installées en Belgique de manière définitive ou à tout le moins pour une durée significative. Il s'agit d'ailleurs d'étrangers dont la situation de

According to the Court, the following reasons justify this limitation. First, legal migrants who are not long-term residents are nonetheless eligible for the *aide sociale* scheme.[92] Second, the Constitutional Court stresses the fact that the relevant social assistance benefits are paid by general taxation and not by contributions on earned income. Budget concerns may thus justify that the beneficiaries are identified in those having a genuine link with the national territory, i.e. nationals or registered long-term immigrants residents.[93]

The attitude of the Belgian Constitutional Court does not seem perfectly in line with the ECtHR case-law.

The ECtHR applies what we may call "a pure non-discrimination approach". This means that nationality is *per se* a suspect criterion, the use of which can be upheld only in narrow situations. The fact the relevant social benefit is contributory does not have any consequence *vis-à-vis* the application of the relevant ECHR provisions, as well as the fact that the applicant is already the beneficiary of other social assistance benefits.

More questionable is the matter of whether the length of the legal stay of the immigrant or his connections with the host State may influence the standard of the scrutiny. Although the Court speaks of nationality as a suspect criterion as such, no matter whether the immigrant is a long-term resident or not, in the decisions taken so far, the applicant had such meaningful attachments to the host State that made him almost equivalent to a national. However, the ECtHR based the presence of these attachments on factual elements rather than on any formal administrative recognition of the status of the long-term resident.[94]

The Belgian Constitutional Court seems to follow a partially different scheme. As noted, the federal legislator has provided a sort of minimal treatment in the social assistance field, to be applied equally to national and aliens (the *aide sociale*). According to the Constitutional Court, this

séjour est dans une large mesure semblable à celle des Belges qui ont leur résidence effective en Belgique'. The principle has been confirmed in other Constitutional Court decisions in relation to disabled allowances. See decisions 153/2007 and 3/2012.

[92] See Belgian Constitutional Court, decision 5/2004, para B (6) (4); decision 92/2004, para B (11) (1) (second para).

[93] See Belgian Constitutional Court, decision 75/2003, B (9); see Belgian Constitutional Court decision 92/2004, para B (11) (1)

[94] See para 39 of the *Koua Poirrez* decision (n 70), where the ECtHR mentions the fact that the claimant was residing in France and she had previously obtained other public social assistance benefits.

would allow the same legislator, when it provides supplemental welfare entitlements, to reserve them to nationals or at least to those immigrants that can be substantially equated to nationals. Thus, the logic is not that of a pure principle of equal treatment between citizens and aliens. It rather seems more grounded upon the idea that aliens, as human beings, are entitled to a minimal form of social protection. This being the case, the equal treatment principle does not apply to other supplemental social assistance benefits, unless the immigrant has an administrative status that highlights his strong connection with the territory, according to an incremental approach.[95]

The incremental approach followed by the Constitutional Court might influence the capacity of sub-national units to introduce classifications based on nationality when dealing with immigrants' eligibility for social assistance benefits. Since the federal legislator already provides a minimal uniform social assistance benefit in this area (the *aide sociale*), it may be submitted that sub-national units would not be obliged to guarantee a full respect of the equality principle in relation to the access of immigrants to other sub-national unit welfare entitlements, or at least that they would be obliged to do so only with regard to long-term immigrants.

Although not many in number, sub-national units' measures in the field of social assistance use long-term residency as a criterion to select the

[95] The Belgian Constitutional Court explicitly took this position in relation to the relevant ECtHR case-law in two cases involving the constitutionality of the provisions of art 4 of the *loi du 27 février 1987*, which establishes limitations on the grounds of nationality in relation to the beneficiaries of disabled allowances. In decision n 92/2004, para B (11) (2), the Court did not consider in breach of the non-discrimination and of the equality principle (see arts 10 and 11 of the Belgian Constitution) the fact that the disabled allowance benefits did not apply to legally present third-country national immigrants. The Court explicitly considered whether this finding was in line with the *Koua Poirrez* decision of the ECtHR and with the principle there set that only narrow considerations could justify a classification based on nationality. It deemed that the ECtHR decision had to be distinguished from the case it was called to assess, since the applicant could be the beneficiary of the *aide sociale*, the amount of which would be calculated taking into account his disabled status.
In the subsequent decision n 153/2007, the Belgian Constitutional Court considered it illegitimate that long-term resident immigrants, once they are listed in the *registre de la population*, were excluded by the scope *ratione personae* of the act. The Belgian Constitutional Court reads the *Koua Poirrez* (n 70) case-law as limited to prohibiting discrimination between nationals and aliens insofar as the latter possess an administrative long-term immigrant status. However, even assuming that the ECtHR applies a strict scrutiny review only when immigrants having a meaningful attachment with the host State are involved, it does not require that this condition must be fulfilled only by having a legal status of long-term resident.

beneficiaries.

To give an example, the already-mentioned case of the Flemish decree on the care insurance scheme does not set any limitations based on nationality in relation to the compulsory joining of the insurance scheme. However, amongst the conditions that a claimant must satisfy in order to receive the benefit, the decree provides a five year residency requirement in the Flanders Region for all potential claimants irrespective of their nationality. Since a long-term residency requirement tends to favour the autochthonous persons, it can thus be questioned whether this measure amounts to an indirect form of discrimination on grounds of nationality, at least with reference to EU citizens.

Other cases of long term residency requirements can be found elsewhere in Belgian sub-national units' social assistance measures. This is the case of the German speaking Community decree concerning disability allowances[96] Art 18 of which provides that in order to be eligible for the relevant benefit, the claimant must reside in the German speaking community territory and, alternatively, have the Belgian nationality or the nationality of one EU Member State or have been continuously resident in Belgium for 5 years or for 10 years in case of non-continuous residence.

The French speaking Community regulation reserves disability allowances to persons residing within that Community and having either the Belgian nationality or being refugees, or stateless, or EU workers. However, those who do not satisfy the nationality requirements may be eligible, provided that they fulfil a five-year previous residency in the national territory.[97]
In these two latter cases, we may note that the residency requirement applies to third-country nationals but it does not to EU citizens and that it refers to the national territory instead of that of the sub-national unit.

This could avoid the problems with the possible breaching of the EU equal treatment provision with regard to EU citizens. Yet, it may be suggested that making the enjoyment of a welfare benefit conditional upon a residency requirement, which applies only to third-country nationals but not to EU citizens, would amount to discrimination on the grounds of nationality, according to art 14 of the ECHR. To this extent, it should be recalled that in *Moustaquim v Belgium*[98] the ECtHR held that, since the EC

[96] In *Moniteur Belge* 13-11-1990.
[97] See art 275 of the *Arrêté du Gouvernement wallon portant codification de la législation en matière de santé et d'action sociale du 29 septembre 2011*, in *Moniteur Belge* 21.12.2011. These provisions were originally inserted in the *décret du 6 avril relatif à l'intégration des personnes handicapées*.
[98] *Moustaquim v Belgium*, App no 12313/86, s A193, (ECHR, 18 February 1991)

constitutes a special legal order, there is an objective and reasonable justification for the preferential treatment accorded to the nationals of the EU Member States rather than to third-country nationals.

However, such a statement was made for a case involving the expulsion of a third-country national, an area where States traditionally enjoy wide discretion. Once the right to enter or stay in a given Country is not at stake and the difference of treatment applies to the enjoyment of fundamental rights set out in the Convention, it seems unlikely that the special legal order of the EU can still be seen as a justification for such a differentiation.

3. *Italy*

In Italy, immigration federalism has thus far mainly concerned the access of immigrants to regional welfare entitlements, especially in the areas of social assistance and social housing.

In decisions n 300/2005 and n. 156/2006, the Constitutional Court clearly stated that the access of aliens to regional welfare system falls under the jurisdiction of the Regions rather than that of the national State. Afterwards, 'immigration federalism' issues have been primarily assessed in the light of the equality principle and with regard given to the protection of fundamental rights.

According to the Constitutional Court's case-law, a systematic reading of arts 2 and 3 of the Italian Constitution requires that aliens and nationals are treated equally as far as the protection of fundamental rights is concerned. Yet, derogations to the equal treatment principle are admitted insofar as they derive from the inherent difference in status between aliens and nationals. The Court refers to the fact that while the entering and staying in the Country constitute rights for the nationals, they are, on the contrary, subject to an administrative authorisation for the aliens.[99]

However, in some cases the Constitutional Court has been continuing to accord a wide discretion the legislator in establishing derogations to the equal treatment principle, basing this on supposedly existing factual differences between the legal status of the citizen and the alien. Moreover, since the fundamental rights category lacks a clear constitutional definition, it is quite difficult to single out what fundamental rights the

[99] See Italian Constitutional Court, decision 104/1969.

Court is referring to.[100]

More recently, the Constitutional Court delivered a number of decisions[101] in which, relying on the equality and the reasonableness principle, as well as on the ECtHR case-law, it held the unconstitutionality of several national law provisions, which limited social assistance entitlements to those immigrants in possession of the long-term resident's EC residence permit, in pursuance of art 8 of the directive 2003/109/EC.[102] According to the directive, the issuing of this residence permit is made conditional to the fact that the third-country national has stable and regular resources which are sufficient to maintain himself/herself and the members of his/her family, without recourse to the social assistance system of the Member State concerned. Thus, limiting these social assistance entitlements to those immigrants in possession of the long-term resident's EC residence permit meant to exclude those immigrants that, even if long-term residents in Italy, were not in possession of the long-term resident's EC residence permit because of the lack of stable and regular resources. The Constitutional Court noted that the relevant social benefits were meant to be instruments to satisfy the primary needs of the human being. Consequently, the legislator could not prevent needy persons from having access to these measures because of they do not have a long-term residence permit the issuing of which is made conditional to the possession of economic resources. In fact, the lack of them is the very reason that caused these needy persons to ask for welfare protection in the first place.

Although the Constitutional Court has regarded the relevant social benefits as a tool for the safeguard of human beings as such, it has been careful to limit their application only to legal aliens, excluding illegal aliens. Moreover, the Court seemed to suggest that the legislator may subordinate the social entitlements to the possession of a permit of stay, insofar as its length proves a significant link with the State. However it did not provide any precise reference to the length of the permit of stay.

Within this context, we can now consider how the Constitutional Court has approached the case of the eligibility of aliens for regional social assistance benefits.

---

[100] See generally for a critical approach to this Constitutional Court line of cases, Marco Cuniberti, *La Cittadinanza. Libertà dell'Uomo e Libertà del Cittadino nella Costituzione Italiana* (Cedam 1997).

[101] See Italian Constitutional Court, decision 306/2008. See also decisions n 11/2009, n 187/2010 and n 329/2011.

[102] Directive 2003/109/EC of the Council of 25 November 2003 concerning the status of third-country nationals who are long-term residents in [2004] OJ L16/44.

To this extent, two different frameworks have been experienced thus far.

On the one hand, some Regions have provided for restrictions upon the eligibility of aliens for regional social assistance benefits and regional social housing. This was accomplished either by excluding aliens from the personal scope of application of the relevant acts, or by making the social benefit entitlement conditional upon a long-term residency condition in the Region.

On the other hand, other Italian Regions pursued an opposite policy by granting some forms of social welfare to illegal migrants, namely the access to some form of temporary social housing or to health treatment, in addition to those already provided by the national health service.[103]

These two situations were both scrutinised by the Constitutional Court.

As far as the first hypothesis is concerned, the Constitutional Court already in 2005 scrutinised the constitutionality of a regional provision reserving free public transport to disabled nationals residing in the relevant Region.

The Region claimed for the legitimacy of this classification, since the benefit at stake could not be considered as a fundamental right. In fact, according to the above-mentioned Constitutional Court case law, the principle of equal treatment between nationals and aliens applies only in relation to fundamental rights. Regions – it was argued by the Region's legal defence – should be free to introduce classifications based on nationality whenever the enjoyment of a fundamental right is not involved.

The Constitutional Court admitted that the regional social benefit could not be considered as a fundamental right. This implied, then, that its previous case law concerning the application of the equality principle to aliens could not be applied as such. However, the Court added that this did not mean that the discretion of the regional legislator was unlimited because any legal classification may be reviewed according to the reasonableness principle.

The Court considered that the regional social benefit was based on a principle of social solidarity in relation to which classifications on nationality, rather than on the needy status of the individual, are arbitrary.

[103] Art 35 (3) of the 1998 Immigration Act sets a list of health treatments to be provided by the health national system to indigent aliens who are illegally present in the national territory.

And even budget constraints could not be invoked as an excuse for such a classification. The Court recalled that art 41 of the 1998 Immigration Act states that immigrants who are present in Italy with a permit of at least one year should have access to social assistance on equal conditions with nationals. This provision – the Court said – is to be considered as a principle of general relevance in the national legal framework and, as such, it can be taken as a paradigm in order to accordingly shape the reasonableness test. This means that derogations to the principle of equal treatment between aliens and nationals are to be grounded on clear, specific reasons.[104]

As a consequence of the above-mentioned constitutional decision, regional acts explicitly limiting aliens' eligibility to regional social entitlements were replaced by the increased use of durational residency requirement as a precondition for access to the relevant social benefits.[105]

These measures were taken in order to prevent new immigrants from benefiting from regional social welfare entitlements. However, since a long-term residency requirement applies generally to all persons living in the relevant region - thus covering nationals, EU citizens and third-country nationals - problems arose with regard to EU law. Indeed, some judges did not to apply the relevant provisions because of the breaches of EU law.

Finally, the long-term residency requirement issue has been examined by the Constitutional Court. In a recent decision, the Constitutional Court has considered that classifications based on residency, as well as those based on nationality, conflict with the equality and reasonableness principles whenever they are used to determine the eligibility for social benefits. The Court held that the benefits provided for by the challenged provisions are intended to remedy to a needy status. Therefore, limitations based on nationality or residency, are not justifiable, insofar as they exclude vulnerable individuals.[106]

The second hypothesis that the Constitutional Court has evaluated is the case of regional statutes aimed at extending social benefits to illegal immigrants.[107]

---

[104] See Italian Constitutional Court, decision 432/2005.
[105] Further references in Dinelli, 'La stagione della residenza' (n 4) 639.
[106] See Italian Constitutional Court, decision n. 40/2011.
[107] See *Legge Regione Toscana*, n 29/2009, and *Legge Regione Puglia*, n 32/2009, both aimed at enlarging the list of health treatments that the national State already provides to illegal aliens. In the case of *Legge Regione Campania* n 6/2010, the regional statute was aimed at ensuring to irregular immigrants a form of temporary sheltered housing, a measure which is not provided for by the 1998 Immigration Act.

The national government brought actions before the Constitutional Court claiming the illegitimacy of the relevant regional provisions. According to the national government's legal counsel, the regional measures interfered with immigration and public order, which are jurisdictions reserved to the national level. They do so because they provide regional social benefits to illegal immigrants, a condition that is considered a crime in the Italian legal system in pursuance of art 10 *bis* of the 1998 Immigration Act.

The Constitutional Court did not follow this reasoning. In decision 61/2011, it made clear that the challenged regional measures were meant to provide basic human rights that both the Constitution and the 1998 Immigration Act grant even to illegal immigrants. [108] It also recalled that it is the 1998 Immigration Act itself that states at art 3 (5) the principle according to which the Regions, within their powers, and budget allocations, must adopt those actions required to guarantee to aliens their fundamental rights.

In the above decision of the Constitutional Court, certain findings may provoke some critical remarks.

First, as far as the allocation of competences vis-à-vis immigration policy is concerned, the assumption according to which the regional measures at stake by no means interfered with national reserved matters is questionable. In fact, the contrast to illegal immigration may also imply restrictions upon the illegal immigrants' access to social rights. The dividing line between immigration issues – reserved to the national level – and migrant integration issues – reserved to the sub-national units' level – maintains its value as long as both legislators share the view that the only alien to be integrated is the legal immigrant with a concrete perspective of staying in the country.

Second, it can be noted that the Court considered the challenged provisions as if they were meant to satisfy basic human rights. In this way, the Court might have implied that Regions are empowered to provide illegal immigrants only with a minimal level of protection in social assistance. However, once the Constitutional Court considered that no interference occurs with the national reserved matters in the immigration

[108] The Constitutional Court had already deemed legitimate the provisions contained in the above-mentioned Tuscan and Puglian statutes. However, it interpreted them narrowly, as they were meant to provide almost the same health treatments set out in the 1998 Immigration Act. The decisions are annotated by Francesca Biondi dal Monte, 'Regioni, immigrazione e diritti fondamentali' [2011] Le Regioni.

field, then why should the power of the Regions to provide social entitlements to illegal immigrants be limited to the provision of the basic human rights?

In any event, if we give a broad reading to the above mentioned Constitutional Court case-law, it may be argued, then, that the Italian Regions can build up a truly inclusive immigrant integration model that can apply to illegal migrants too, despite the opposite national pattern.

IV. THE "CULTURAL REGIONAL CITIZENSHIP" DIMENSION

Integration is a process where different components have a role: language, religion, culture, social status, et cetera. The integration of migrants has been regarded, for some time, as a sort of natural and voluntary process taking place, as time elapsed. The migrant, through his work, became actively involved in the host State society and thus enjoyed a series of rights that made himself part of the relevant community.

States have undertaken different policies in order to facilitate the integration process: in some cases, they have guaranteed migrants the right to express, in public, ways of life strictly linked to the culture or to the religion of their country of origin, permitting derogations to the general applicable rules; in other cases, the idea of special rights, as a way to allow minority groups to express their identity, has been denied, due to fears that this could lead to disaggregate the civil society of the host country.[109]

However, in recent years a new idea of integration is occurring. Integration is increasingly becoming a sort of precondition, a positive obligation that a migrant must fulfil in order to have or keep the status of legal immigrant. We can call it "integration by law".

This concept implies a clear link between the level of integration of the migrants and their legal status. Many national laws on immigration are introducing provisions dealing with integration tests: the knowledge of the host State's language, history and civic values is increasingly used at all stages of a migrant's stay, as a precondition for entry into, or remaining in,

[109] See, on the debate on multicultural or assimilation integration approaches Adrian Favel, *Philosophies of Integration – Immigration and the Idea of Citizenship in France and Britain* (Palgrave 2001); Ralph Grillo, *Pluralism and the Politics of Difference: State, Culture and Ethnicity in Comparative Perspective* (Oxford Clarendon Press 1998); Julie Ringelheim, *Diversité Culturelle et Droits de l'Homme* (Bruylant 2006).

the national territory.[110]

With regard to this, it is worthy of note that integration remains basically a national issue even within Europe. Although the EU has recently increased its powers in the field of immigration policy, in the area concerning the integration of immigrants, art 79 (4) EUFT foresees EU intervention as merely complementary to that of the EU Member States.

Within this framework, both the Common Basic Principles for Immigrant Integration Policy[111] and the European Pact on Integration and Asylum[112] expressly admit that States may ask immigrants to learn the State's language and to respect the identities of the Member States and of the EU.

Moreover both directive 2003/86/EC on the right to family reunification and directive 2003/109/EC on the status of third-country nationals who are long-term residents allow States to require third-country nationals to comply with integration measures, in accordance with national law, as a requirement, respectively, for the exercise of the right to family reunification (art 7 (2) dir 2003/86/EC) and for acquiring the status of long-term resident (art 5 (2) dir 2003/109/EC).

However, it may be argued that Member States' discretion in this field is not unlimited, at least when these compulsory integration measures may end up in becoming an excessive burden for immigrants, thus affecting the *effet utile* of the above-mentioned EU directives.[113]

The trend towards imposing national linguistic and cultural requirements as a precondition for the immigrant entry or stay in the national territory reveals that integration of migrants is increasingly considered as part of the national immigration policy. While some scholars emphasize that these cultural requirements may be a surreptitious way of selecting immigrants on otherwise forbidden grounds, such as religion, race, ethnicity and that

[110] See Ricky van Oers, Eva Ersbøll and Dora Kostakopoulou (eds), *A Re-definition of Belonging? – Language and Integration Tests in Europe* (Martinus Nijhoff 2010); Elspeth Guild, Kees Groenendijk and Sergio Carrera, *Illiberal Liberal States – Immigration, Citizenship and Integration in the EU* (Ashgate 2009).

[111] Council of the EU, 1461/04, 2004, Justice and Home Affairs, 2618th meeting.

[112] European Pact on Integration and Asylum, (Council of the EU, 24 September 2008).

[113] See case C-508/10, *European Commission v the Netherlands* (ECJ – 26 April 2012).The Commission challenged the legitimacy of a Dutch provision that imposed a fee in order to obtain the status of long-term resident, arguing that it breached Directive 2003/109/EC. The ECJ found that the fee was disproportionate and thus it undermined the *effet utile* of the Directive.

they are expression of a 'repressive liberalism',[114] others suggest the idea that language and civic integration requirements are not necessarily at odds with a civic notion of nation, since they permit effective immigrant integration.[115]

It is also submitted that the tendency to compel immigrants to learn the host country's language is on the rise due to some changes in the traditional pattern of third generation immigrant linguistic assimilation. These are, namely, the fact that immigrants tend to maintain strong connections with the country of origin and the fact that they form compact and homogeneous communities, thus questioning the language acquisition of the host State.[116]

Consequently, linguistic requirements are not seen as inherently illegitimate instruments. Although they can interfere with the right to respect for private life, according to art 8 of the ECHR[117], they may be considered necessary in order to pursue important public interests, such as building national cohesion and integrating immigrants in the host society.[118] It is the way in which they are applied, though, that it is important to verify, in order to assess their compliance with fundamental rights.[119] The legitimacy of the linguistic requirements relies on a proportionality test that takes into consideration several aspects, such as the fact the linguistic requirement is compulsory or not, or that it applies

[114] See Christian Joppke, *Veil: Mirror of Identity* (Polity 2009) 115, which refers to the case when liberalism promotes its liberal goals by illiberal means, forcing the members of its community to identify with liberal norms. See also Liav Orgad, '«Cultural Defence» of Nations: Cultural Citizenship in France, Germany and the Netherlands' (2009) 15 Eur L J, 719; Liav Orgad, 'Illiberal Liberalism Cultural Restrictions on Migration and Access to Citizenship in Europe (2010) 58 Am J Comp L, 53; Rainer Bauböck and Christina Joppke (eds), 'How Liberal Are Citizenship Tests", (2010) EUI Working Papers RSCAS, 2010/41 <http.//Cadmus.eui.eu/btstream/handle/1814/1396/RSCAS_2010_41corr.pdf?sequence =3> accessed 2 January 2013.

[115] Will Kymlicka, 'Immigration, Integration and Minority Nationalism' in Michael Keating and John McGarry (eds), *Minority Nationalism and the Changing International Order* (OUP 2006).

[116] See Alan Patten and Will Kymlicka (eds), *Introduction: Language Rights and Political Theory* (OUP 2003) 8-9.

[117] See Hugues Dumont and Françoise Tulkens, 'Citoyenneté et Responsabilité en Droit Public', in Hugues Dumont, François Ost and Sébastien Van Drooghenbroeck, *La Responsabilité Face Cachée des Droits de l'Homme* (Bruylant 2005) 219-220.

[118] See Ruth Rubio-Marin, *Language Rights: Exploring the Competing Rationales*, in Alan Patten and Will Kimlicka, Language Rights (n 111) 52-79.

[119] See extensively José Woehrling, 'Linguistic requirement for immigrants', in *Mundialitzaciò, Lliure Circulacio i Immigraciò, i l'Esigència d'una Llengua com a Requisite* (Institut d'Estudis Autonomics 2008) 133.

to obtaining citizenship status rather than a permit of stay.

The issue of cultural-linguistic requirements for immigrants has a further problematic dimension in those States characterised by a multinational structure or where national linguistic minorities are settled. Immigrants may find it more useful, or more attractive, or easier, to learn the language of the majority than to learn the local language. In some cases, they tend to create their own community and they do not wish to learn the local language. This led sub-national units to adopt policies aimed at protecting and promoting the status of their local national language among immigrant newcomers.[120]

A national compulsory integration/linguistic measure could not be suitable in order to preserve the sub-national cultural-linguistic distinctiveness. Since the residence permit has legal effect nationwide and since it cannot limit the right of the legally-admitted aliens to move and reside within the Country, problems arise concerning the choice of the language the alien would be required to learn.

The adoption of cooperative mechanisms between the federal and the sub-national levels in the regulation of the entry and the stay of immigrants has represented one means of satisfying the sub-national units' need for linguistic protection. This solution has been enforced in Canada, where following an agreement with the federal authorities, the Province of Quebec is allowed to select immigrants on the grounds of their French knowledge.

Flanders has pursued a different course. In 2003, the Flemish authority introduced linguistic and cultural integration courses which newly-arrived third-country nationals – and even some categories of Belgian citizens – are compulsorily required to attend. The failure to attend these linguistic-cultural integration courses is punished with an administrative fine, and can be a ground for the suspension of the enjoyment of social welfare entitlements. The integration requirements, then, are not linked to the admission of immigrants into the national territory, as it is the case in Canada. This allows Flanders to autonomously pursue its integration policy, with no need for a cooperation agreement with the federal level.

[120] For a political science view of the issue, see Kymlicka, 'Immigration, Integration and Minority Nationalism' (n 117); Ricardo Zapata Barrero, *Immigration and Self-Government of Minority Nation* (Peter Land 2009). In the legal literature, see the contributions in *Mundialitzaciò, lliure circulacio i immigraciò, i l'esigència d'una llengua com a requisite* (n 119).

We shall consider, then, in further detail, the case of Belgium, which is the only State amongst those considered by our comparative analysis where this topic arose. In Italy, linguistic requirements for immigrants have been introduced by the national legislator as a condition for the renewal of the permit of stay.[121] On the occasion of the passing of the statute, the Autonomous Province of Bolzano, where a German-speaking minority is settled, asked for the linguistic requirements be in German. The request has not been taken into consideration by the national legislator.

1. *The Flemish Inburgering Decree*

As we have already noted above, art 5 (II) (3°) of the 1980 Special Act explicitly grants the Communities the power to deal with the integration and reception of immigrants.
An analysis of the *travaux préparatoires* works suggests that the integration measures that the Communities were empowered to pass were intended to be principally based on a voluntarily scheme.[122]

This has been the case until 2003, when the Flemish authorities decided to enact a first decree concerning the so-called *inburgering* policy. This is defined by the Flemish legislator as an interactive process that implies rights and duties, both for the newcomers and for the Flemish government. The *inburgering* is structured in a two-stage process, the first of which includes Dutch language courses, civic orientation (which covers several aspects of Flemish society, such as education, mobility and health) and vocational guidance, which focuses on the access to the labour market.[123]

The courses are mandatory for certain groups of individuals, namely: immigrants, aged at least 18, who are authorised to stay for more than three months in Belgium and who have been registered in a Flemish local municipality for less than twelve months; Belgian nationals, born outside Belgium, who have at least one parent born outside Belgium and have been registered for the first time in a Flemish municipality for less than twelve months; and aliens who are religious ministers.[124]

---

[121] See art 4 bis of the 1998 Immigration Act.
[122] The point is highlighted by Van Drooghenbroeck, 'Fédéralisme, Droits Fondamentaux et Citoyenneté' (n 51)
[123] See Marie-Claire Foblets and Zeynep Yanasmayan, 'Language and Integration Requirements in Belgium: Discordance between the Flemish Policy of "Inburgering" and the Federal Legislator's View(s) on the Integration of Newcomers and Migrants', in van Oers, Ersbøll and Kostakopoulou (eds) (n 110) 271; Dumont and Tulkens, 'Citoyenneté et responsabilité en droit public' (n 117) 219-220.
[124] See art 5 of the *Inburgering* Decree, following the amendment introduced in 2008.

Besides the mandatory target group, the *inburgering* process is offered on a voluntarily basis to other groups. These are further divided into priority and non-priority groups. The former are accorded priority when demand exceeds the places available. Among the individuals included in the priority group, are those immigrants who have been registered in a Flemish municipality for more than 12 months and who are beneficiaries of social assistance or social security revenues.

The original 2003 decree has been amended several times in order to better target the individuals to whom the act applies, and to define the categories of persons exempted from the courses.[125] Due to the lack of coordinating provisions concerning the *ratione temporis* and *ratione personae* scope of application of these acts, problems arise in the identification of the individuals currently required to attend the courses.[126]

As stated above, the failure to attend a compulsory *inburgering* course is punished with an administrative fine. However, even those persons that do not fall into the mandatory target-group may be requested to attend the *inburgering* course when they apply for social security or social assistance benefits.[127]

The passing of the *inburgering* decree by the Flemish authority constituted a turning point with regard to the immigrant integration policy framework thus far adopted in Belgium.[128]

As far as the federal level is concerned, no mandatory integration requirements have been introduced in order to allow immigrants to enter or stay in the country. Moreover, the previous references to integration criteria as a prerequisite for obtaining Belgian citizenship were abolished in 2000. Previously, proof of willingness to integrate was a prerequisite in order to be awarded the status of Belgian citizen. This provision had been implemented in practice by requiring sufficient knowledge of at least one of the three official languages of Belgium. Since this legislative amendment, continuous residency in Belgium for a certain time has

[125] This is the case for EU citizens, or for a non-EU citizen family member of an EU citizen.
[126] See for details Eric Somers, 'Le Parcours d'Intégration Civique en Flandre. Les Personnes Visée et Leurs Obligations', in Julie Ringelheim (ed) *Le Droit et la Diversité Culturelle* (Bruylant 2011) 301-344.
[127] See *Décret relative à la politique flamande d'intégration par le travail*, 4 June 2003.
[128] See Ilke Adam, 'Une Approche Différenciée de la Diversité? Les Politiques d'Intégration des Personnes Issues de l'Immigration en Flandre, en Wallonie et à Bruxelles (1980-2006)' in Ringelheim (n 126) 251-300 ; Foblets and Yanasmayan, 'Language and Integration Requirements in Belgium' (n 110) 271.

become the most important condition to be fulfilled.[129]

As far as the other federate units are concerned, Wallonia has always refrained from the idea of targeting individuals on the grounds of nationality or ethnic origin. No mandatory form of immigrant integration courses has been introduced thus far, though voluntarily measures directed towards immigrants are in force.

The *inburgering* decree itself has not been challenged before the Constitutional Court. However, the Constitutional Court has nonetheless had the opportunity to scrutinise the Flemish policy of compulsory immigrant integration.

In 2006, the Flemish Region introduced some amendments to the *Code Flamand du logement* in relation to social housing provisions. With the aim of facilitating the communication between social housing tenants and the officers who carry out the service, the Flemish authority required that any social housing tenant must show the will to learn Dutch as a prerequisite in order to rent a house under the social scheme. For those individuals falling under the scope of application of the *inburgering* decree, the obligation is satisfied with the attendance of the *inburgering* course. The failure to fulfil the mentioned requirements is considered as a serious contractual breach, which leads to the unilateral termination of the contract.

The above-mentioned requirement applies to all individuals, irrespective of their national status, provided that they live in the Flanders region and ask for social housing. The problem arose especially with regard to Belgian citizens of the French-speaking group living in some bordering Flemish municipalities – so-called *communes à facilités (linguistiques)* – who are granted the right to use French with the local public administrators.

A constitutional claim was brought by both the French Community and two organisations promoting the immigrants' rights.

The Constitutional Court held that the application of the decree cannot affect the guarantees that the French minority linguistic group enjoys in the *communes à facilités (linguistiques)* and thus the decree cannot apply to French-speaking persons living there.[130]

---

[129] See Marie-Claire Foblets, 'Le Parcours Mouvementé du Code de la Nationalité Belge: Rétrospective (1985-2003)' in [2003] Annales de Droit de Louvain, 259.

[130] See Belgian Constitutional Court 101/2008, annotated by Nicolas Bernard, 'L'arrêt Wooncode de la Cour Constitutionnelle du 10 Juillet 2008, quand l'Arbre (Linguistique) Cache la Forêt' [2008] Journal de Tribunaux, 689.

As for the other persons to whom the act applies, the Court considered whether the Dutch requirement amounted to a violation of art 23 of the Belgian Constitution, which deals with the social and economic rights of the individual. It was argued that, since the failure to learn Dutch could entail the termination of the social house lease, this would amount to a violation of art 23 of the Belgian Constitution.

The Court observed that this provision does not prevent the legislator from making the award of social benefits conditional upon certain obligations, which the applicant must fulfil. This is the case of the Dutch language requirement. The Court considered that the aim pursued by the legislator – namely to ease the communication between the social housing tenants and the service providers – was legitimate, and that the Dutch-learning obligation was a proportionate means, provided that the failure to fulfil the requirement did not automatically imply the termination of the contract and that the attendance of the language courses was free of charge.

The Constitutional Court, then, applied a proportionality test. Accordingly, whereas it considered legitimate in principle the idea of conditioning welfare entitlements upon the fulfilment of the linguistic requirement, it deemed that the means chosen to make the obligation effective were disproportionate. It is also important to stress that the Court, quite surprisingly, did not consider the relevant provision to be in breach of the EU law, notwithstanding the fact they potentially constitute an obstacle to the EU citizens' freedom of movement.

V. Conclusions

'Immigration federalism' questions not only the idea that immigration is a matter to be vested in the national legislator but also what immigration exactly means as to the purpose of the division of powers.

'Immigration' certainly includes those measures that concern the entry, the stay and the expulsion of aliens. But it also includes measures defining the rights of third-country nationals, once they are admitted or they find themselves illegally in the national territory (so-called legal status of aliens). Finally, immigration may also cover those measures that are specifically meant to culturally integrate immigrants.

Immigration, then, is to be considered as a policy rather than a matter, ie a political objective that is pursued through actions falling under different competences. More precisely, immigration is a shared-policy: a same

immigration-related topic may fall under either a national competence or under a sub-national units' one. Integration of migrants is a good example. Although this is a matter usually vested in the sub-national units, national states are increasingly making the issuing of the authorisations for entering or for staying in the national territory conditional upon the meeting of cultural and linguistic tests. Thus, the integration of immigrants is increasingly considered as it were an exercise of the national power of regulating immigrants' entry and stay. Similarly, contrasting illegal migration is a goal that may be pursued by both the national legislators, with actions falling under the aliens removal matter of competency, and by sub-national units acting in pursuance of their police powers, at least insofar they are granted such powers.

The definition of the boundaries within which each territorial unit may act within the notion of the shared immigration policy is not easy to draw. The idea that the national legislator would primarily have the task of regulating the entry and the stay of the immigrants, while the sub-national units would have the power to deal with the aliens' legal status and their integration, is weakening. On the one hand, the attempt of some US states to have a role in combating illegal migration, and the Belgian regional competences in issuing work permits, represent examples of the fact that even the regulation of the entry and the stay of immigrants may involve sub-national units' regulations. On the other hand, the social integration of migrants has been always influenced by national decisions concerning welfare.

Because of the lack of a clear-cut material dividing line, 'immigration federalism' is a dynamic phenomenon, the equilibrium of which is being constantly challenged according to the social and political needs of the territories involved. Thus, it is not surprising that in each of the three States taken into consideration, immigration federalism differs in its material scope: in the US, it is currently more focussed on contrasting illegal immigration; in Italy it is the regional welfare eligibility what matters the most; and finally, in Belgium, it is the integration of linguistic immigrants.

Therefore, it is more convenient to look at the way each legal system structures the relevant equilibrium among territorial authorities, rather than trying to define common dividing lines.

To this extent, the US and the two European countries taken into consideration by our analysis, show important differences.

In the US, the federal legislator has been granted a wide power in dealing

both with immigration, intended as the regulation of the entry and the stay of immigrants, and with the legal status of aliens. This implies that it may intervene in areas that are usually up to the states to regulate. This also means that 'immigration federalism' is a phenomenon, the scope of which may increase or decrease as a consequence of the choices made by the federal legislator. The latter may decide to promote state action and thus differentiation among states – as it did with the PRWORA in relation to aliens' welfare eligibility – or, on the contrary, it may decide to restrain it by explicitly pre-empting the field – as it did in 1986 when it federalised the power to sanction employers who hired illegal immigrants.

The traditional deferential attitude adopted by the US Supreme Court in relation to the national decisions regarding immigration and *alienage* law has deeply influenced the division of competences in the field. State measures in the area of immigration are legitimate as long as federal law does not intervene by extensively regulating the matter, or as long as states actions do not stand as an obstacle to objectives laid down by federal law. The US states' scope of intervention is potentially wider than that of the Belgian and Italian sub-national units, since the former may rely on their police powers as a legal base for measures dealing with illegal migration. Although they cannot autonomously enforce the federal removal procedure as such, they may take actions that allow them to indirectly ease it. It remains nonetheless the case that an action of the federal legislator explicitly pre-empting state measures in the area is always possible. Thus, there are not constitutional guarantees of states' powers in the immigration area, at least insofar as the federal legislator, acting in pursuance of what it deems to be the national interest, may always explicitly pre-empt state measures in the field.

The Italian and the Belgian cases show a different framework: they both consider the power of the national state in dealing with immigration as primarily related to the regulation of the entry and the stay in the country. Both the Belgian and the Italian Constitutional Courts have refused to conceive of the legal status of aliens as an autonomous-standing power clause, which would enable the national legislator to intervene in areas otherwise reserved to the sub-national units.

This means that the national legislator can legislate with regard to the rights and duties of immigrants as long as the relevant area falls under a national competence. For example, with regard to social assistance for immigrants, the national legislator can act, provided that it has powers in the field of social assistance, whether the beneficiaries are immigrants or otherwise.

Although sharing this common feature, immigration federalism has developed differently in the two European countries. While in Italy, immigration federalism has been more focussed on the social area, in Belgium, due to the wide powers the federal legislator still retains in relation to welfare, immigration federalism has been more focussed on immigrants' linguistic integration. However, the case of the Flemish insurance scheme and the emerging idea of welfare as a parallel competence may imply for the future a more meaningful role for the sub-national units in the field of social immigrant integration.

The different approach followed by the US, on the one hand, and by Italy and Belgium, on the other hand, as to the role of the national legislator in immigration policy, has had as a consequence that in these latter two States, the constitutional jurisdictions have played a more meaningful role in guaranteeing a territorial harmonisation of the legal status of immigrants through the enforcement of the equality and non discrimination principles.[131]

This can be clearly noted by considering the development of what we call 'regional social citizenship' and 'cultural regional citizenship'.

We emphasized that the emergence of 'immigration federalism' can be also explained by the fact that immigration flows have pushed the sub-national units to remodel their relationship with their new communities according to, alternatively, restrictive rather than inclusive attitudes towards newcomers. However, this development has to take into account the limitations deriving from the equality and non-discrimination principles as they are judicially enforced at the national and international levels.

Such a framework appears unlikely, both in the US, and in the two European states.

In the US, the issue of regional social citizenship is currently more influenced by the federal legislator than by the judicial enforcement of the equality principle. Although the Supreme Court stated that a strict scrutiny standard of review should apply to state - but not to federal - *alienage* classifications, the federal legislator with the enactment of the

[131] This is consistent with the opinion that in Europe, unlike in US, the judiciary has played a pivotal role in defining the constitutional status of the aliens. See Christian Joppke and Elia Marzal, 'Courts, the New Constitutionalism and Immigrant Rights: The Case of the French Conseil Constitutionnel' (2004) 43 Eur J Pol Research 823-844.

PRWORA has allowed the states to circumvent the aforementioned Supreme Court case-law. This has permitted the states to freely adopt restrictive or rather inclusive measures with regard to aliens' eligibility for welfare state benefits.

In the European scenario, constitutional jurisdictions, also as a consequence of the ECtHR case-law and to a lesser extent that of the ECJ, are increasingly considering nationality as an illegitimate criterion when it is applied to determine welfare eligibility. This has pushed the sub-national units to increasingly make use of long-term residency requirements as a precondition to be the beneficiary of regional welfare entitlements.

The use of long-term residency requirements represents proof that the sub-national units will do what they can to strengthen the sense of common belonging of the people living in their territory. It may also be a way to preserve autochthonous communities, and thus to indirectly discriminate against recently arrived foreigners. This is why they are considered as suspect measures with regard to the non-discrimination principle. As long as they are applied generally, thus covering EU citizens, they may amount to a breach of EU law.

The case of cultural-linguistic regional citizenship is different: fewer legal constraints are found in relation to it. An explanation for this may be that the linguistic integration requirements, applied by sub-national units, may effectively serve two contrasting objectives: they may be used either as a surreptitious way to discriminate in the provision of some public benefits, or as measures effectively helping the immigrants to integrate into society. Because of this, they do not appear as inherently discriminatory measures, as nationality and long-term residency requirements applied in the social field do. It is the way they are implemented that is determinant in order to understand whether they are legitimate or not. For example, the Belgian Constitutional Court has admitted the legitimacy in itself of the *inburgering* Flemish policy, but it has reviewed those aspects of it that were more in contrast with the individual's fundamental rights.

Thus, the comparative analysis conducted thus far with reference to the three legal orders shows two different schemes for accommodating, on the one hand, the interests of the national legislator in defining ultimately the narratives of the integration process of immigrants, and, on the other hand, the demands of territorial differentiation put forward by the sub-national units.

The US model certainly guarantees the federal authorities wide power in order to define a common national framework for the integration of

immigrants. However, this may result not only in potentially unlimited restrictions on the self-government rights of the sub-national units, but even in non-application of the constitutional guarantees of the equality and non-discrimination principles, at least insofar as it is the federal legislator that takes action.

On the contrary, both the Belgian and the Italian experiences have somehow undermined the role of the national legislator in defining a common nationwide immigrant integration framework with regard, respectively, to the welfare and to the cultural linguistic integration of immigrants. It is rather for the judiciary, especially the constitutional judiciary, to perform a homogenizing territorial role through the enforcement of the equality and the non-discrimination principles.

However, this scheme may lead to some inconsistencies. This is, for instance, the case of limitations to regional welfare immigrants' eligibility based on long-term residence requirements. Although both Constitutional Courts refer to the case-law of the ECtHR, they have indeed provided different solutions to this common question: the Belgian Constitutional Court admits the legitimacy of the long-term residence requirements, whereas the Italian Constitutional Court does not, deeming them in breach of the principle of equality.

Moreover, in the Italian case, the Constitutional Court decisions seem to suggest that Regions are free to develop integration policies for immigrants even if the latter are irregular. This occurs despite the fact that the regular status of immigrants is considered not only by the national but also by the European authorities as a precondition for any immigrant integration measures.

Given the inconsistencies that both models present, a third solution may thus be suggested: immigration cooperative federalism.

If we accept to consider immigration as a shared policy, in the sense we have outlined above, cooperative federalism instruments would seem the best way to avoid overlapping and conflicting interventions, at the same time guaranteeing a coordinating role to the national state, thus emphasizing that immigration policy as a whole is a national concern[132].

[132] For a similar view, see Eduard Roig, 'Relaciones Intergubernamentales en Material de Inmigración: Desarrollo de un Modelo en Construcción', in Eliseo Aja, José A Montilla and Eduard Roig (eds) *Las Comunidades Autónomas y la inmigración* (Tirant lo Blanch 2006) 76.

The inherent capacity of cooperative federalism to substantially derogate from the formal division of powers and its polymorphic nature represent other reasons why cooperative federalism seems specifically suitable within the field of immigration policy.

Cooperative federalism may in fact consist of both mere participation – where all territorial components take part in the decision-making process but a leading position is reserved to one of them – and of true collaboration – where the several territorial components are at a substantially equal position because they all have competences within the relevant field.

Thus, as far as immigration policy is concerned, cooperative federalism mechanisms may permit the varying of the intensity of the participation of the territorial units according to the specific immigration-related matter at stake. For instance, in relation to the entry and stay of immigrants a more substantial role may be recognised to the national state in the decision-making process. Consequently a 'weak' cooperative federalism, in the form of a participatory role for the sub-national units, may be preferred. On the contrary, in the field of welfare or cultural integration of immigrants, where sub-national units indeed have their own powers, cooperative federalism should be shaped so as to guarantee to the sub-national units the power to participate on an equal footing with the national legislator.

However, within this common legal framework, each legal system can establish its own equilibrium between territorial uniformity and federalism, thus taking into consideration its institutional peculiarities.

This may explain why in some legal systems 'strong' forms of cooperative federalism – where the consensus among all the territorial participants is required – have been enacted even in the area of the conditions of entry and stay of immigrants, despite the fact that this is usually a matter of concern for the national legislator.[133]

[133] This is the case of Canada where an agreement has been concluded between the federal government and the Quebec Province in order to allow the latter to select immigrants on the basis of their knowledge of the French language. Although the Canadian model of territorial allocation of powers relies on the idea of a clear-cut list of matters (so called water-tight compartment), immigration and agriculture are an exception of shared competences. This means that according to art 95 of the *British North America Act* (BNA) each Province may make laws in relation to immigration into the Province. However, the federal legislator may take action in the field as well. In the case of overlapping interventions, art 95 of the BNA states that: 'any Law of the Legislature of a province relative to [...] Immigration shall have the effect in and for the province as long as far only as it is not repugnant to any Act of the Parliament

Co-operative federalism mechanisms have not been disregarded in the countries considered in our analysis. However, when they have been applied, they were ineffective. As far as the US is concerned, the federal statute on immigration – the INA – provides a framework within which to develop co-operative mechanisms between federal and state authorities in the field of the expulsion of aliens. However, this has only been enforced in few cases and it has not prevented states from autonomously pursuing their policies of contrast to illegal migrants.

In Italy, consultations with the Regions concerning the number of immigrants to be admitted annually in the country has not been an instrument that effectively takes into consideration the specific territorial needs for migrant workers at the regional level.

Concerning Belgium, it is because the uneven regional enforcement of the federal legislation in the issuing of work permits that the federal legislator has required a binding cooperative agreement. However, despite its adoption, uniformity has not been attained.

The complexity of reaching a compromise may also explain why the Belgian national legislator has thus far refrained from dealing with immigrant cultural-linguistic integration. Due to the different political approaches the Communities follow in relation to this area, the federal legislator has preferred to set the issue aside, and to allow the Flemish Authority to develop its own linguistic integration policy for immigrants.

Nevertheless, there are signs suggesting that, for the time being, cooperative federalism solutions could be more effectively pursued even in these countries.

Concerning the US, following the above-mentioned decision of the Supreme Court in relation to the Arizona Bill, it seems clear that a more decisive role for the sub-national units in the federal removal procedure will have be consecrated within the cooperative schemes already provided for by the INA.

---

of Canada'. Thus, although the federal legislator would have had the power to pre-empt the Province legislator in the immigration field, it has refrained to do so, preferring to deal with the issue by means of cooperative federalisms instruments. This clearly highlights the potentialities of cooperative federalism as a way to informally derogate to the otherwise applicable division of competences. See Woehrling, 'Linguistic Requirement for Immigrants' (n 120); Matteo Nicolini, 'La Disciplina Canadese sull'Immigrazione tra Multiculturalismo, Secessionismo, e Riforme' in [2003] Diritto Pubblico Comparato ed Europeo, 726.

Regarding Italy, one should recall that the massive influx of people coming from the Libyan coasts, soon before the fall of the Gaddafi regime, pushed the national government to conclude an agreement with the Italian Regions in order to organise these peoples' reception. This was done despite the fact that specific legal provisions assign the relevant powers exclusively to the national authorities.[134]

Another reason suggesting that, at least in Europe, cooperative federalism could be further developed as a means of coordinating the measures of territorial units in the immigration area is the fact that the EU itself seems to adopt this model.

According to art 79 of the TFEU, the Union shall develop a common immigration policy. This consists of measures in the following areas: the condition of entry and residence and standards on the issue by Member States of long-term visas and residence permits (art 79 (2) lett a TFEU); the definition of the rights of third-country nationals residing legally in a Member State (art 79 (2) lett b TFEU); illegal immigration and unauthorised residence, including removal and repatriation of persons residing without authorisation (art 79 (2) lett c TFEU) and combating trafficking in persons (art 79 (2) lett d TFEU). Art 79 (4) TFEU further mentions the area of integration of third-country nationals residing legally in the territories of EU Member States.

Thus, at the European level, immigration is not conceived of as a jurisdiction in itself, but rather as a political objective to be pursued through actions in different fields. The intensity of the intervention of each territorial component – respectively, the EU and the Member States – varies in relation to the specific area taken into consideration. Thus, in the fields of the conditions of entry and residence, the rights of third-country nationals, and illegal immigration, trafficking in persons, the EU may enact "hard-law" instruments. Cooperation with Member States is ensured not only by the voting procedure in the Council but also by the use of directives, which give a certain margin of discretion to the States[135].

[134] See art 129 (1) lett h) and lett l), Decreto Legislativo 31 March 1998 (n 112). The agreement has been concluded the 26 of September 2012. It is available at <www.statoregioni.it/Documenti/DOC_037760_100%20CU%20(P.1BIS%20ODG).pdf> accessed 2 January 2013.

[135] The directives thus far enacted in the immigration area are characterised by a limited degree of harmonisation and by the setting of minimum standards. This explain why many provisions in the directives itself expressly enable Member States to provide higher standards. This feature inevitably maintains a large political discretion by Member States in these fields.

On the contrary, in the field of immigrant integration, the role of the EU is limited to sustaining Member States' autonomous actions and developing soft-law coordination mechanisms according to the open method of the coordination scheme.

The EU case may thus be taken as an example of the potentialities, especially in terms of flexibility, that cooperative federalism could offer, even within immigration policy.

To conclude, since immigration federalism is a dynamic phenomenon the equilibrium of which is constantly challenged, cooperative federalism will allow each legal system to define its own balance between, on the one hand, the national interest in defining immigrant integration process and, on the other hand, the territorial differentiation that is the consequence of any real federalisation process. Cooperative federalism will guarantee that both the national and the sub-national authorities, each within their relevant granted powers, are on an equal footing in order to find the best means of taking actions within immigration policy.

MODELLING THE COASE THEOREM

F E Guerra-Pujol*

More than fifty years ago Ronald Coase published 'The Problem of Social Cost'. In his paper, Professor Coase presents an intriguing idea that has since become known among economists and lawyers as the 'Coase Theorem'. Unlike most modern forms of economic analysis, however, Coase's Theorem is based on a verbal argument and is almost always proved arithmetically. That is to say, the Coase Theorem is not really a theorem in the formal or mathematical sense of the word. Our objective in this paper, then, is to remedy this deficiency by formalizing the logic of the Coase Theorem. In summary, we combine Coase's intuitive insights with the formal methods of game theory.

TABLE OF CONTENTS

I. INTRODUCTION 180
II. BRIEF BACKGROUND: THEORETICAL SIGNIFICANCE OF THE COASE THEOREM 180
III. COASE'S ARITHMETICAL MODELS OF THE COASE THEOREM (STRAY CATTLE AND RAILWAY SPARKS) 182
1. *Stray Cattle* 182
2. *Railway Sparks* 184
3. *Some Non-Arithmetical Models of Coase's Theorem* 185
IV. COASIAN GAMES 189
1. *A Two-Player Coasian Game with Probabilistic Payoffs* 189
2. *An n-Player Coasian Game with Probabilistic Payoffs* 196

* Associate Professor, Barry University Dwayne O. Andreas School of Law, e-mail: eguerra@barry.edu.

This research was supported by a summer grant from Barry University. In addition, I wish to thank two anonymous referees for taking the time to review my paper and for their helpful comments and suggestions. I also thank my wife Sydjia Guerra as well as Erle Robinson for their help in reviewing the mathematical analysis in Section 4 of the paper. On a more personal note, I would like to thank Mrs Bridgette Walpole for making her home 'Lansdowne' in St Ann parish, Jamaica, available to the author, and my mentor, Guido Calabresi, who introduced me to the Coase Theorem over 20 years ago. Last but not least, I must thank my dear friend and colleague, Orlando I Martinez-Garcia. The idea of combining the ideas of Ronald Coase with the methods of John Nash (equilibrium analysis) and George Price and John Maynard Smith (evolutionary game theory) occurred to me during one of our many late-night discussions. In many ways, this paper is just as much as his as it is mine.

3. *An Alternative Coasian Game with High Transaction Costs, Fixed Legal Rules, and Deterministic Payoffs* 201
V. CONCLUSION 203

I. INTRODUCTION

More than fifty years ago Ronald Coase published his seminal paper 'The Problem of Social Cost'.[1] In his paper, Professor Coase presents an intriguing idea that has since become known among economists and lawyers as the 'Coase Theorem'.[2] Unlike most modern forms of economic analysis, however, the Coase Theorem is based on a verbal argument and is almost always proved arithmetically. That is to say, Coase's Theorem is not really a *theorem* in the formal or mathematical sense of the word. Our objective in this paper is to remedy this deficiency by presenting the Coase Theorem as a formal game. In summary, we try to combine Coase's intuitive insights with the formal methods of game theory.

The remainder of this paper is organized as follows. Sections 2 and 3 provide some background regarding the Coase Theorem. Specifically, Section 2 briefly discusses the significance of the Coase Theorem, while Section 3 presents two of the most famous illustrations of the Coase Theorem—Coase's simple model of farmer-rancher interactions and Coase's arithmetical analysis of the problem of railway sparks—as well as some previous attempts to formally model the Coase Theorem. Next, Section 4 presents a general game-theoretic model of the Coase Theorem, one that does not depend on artificial parameter values. Specifically, Section 4.1 presents a simple two-player 'Coasian game' with probabilistic payoffs, Section 4.2 presents a population model of the Coase Theorem with probabilistic payoffs, and Section 4.3 then models an alternative Coasian farmer-rancher population game with high transaction costs and the presence of legal rules, but with fixed instead of probabilistic payoffs. Section 5 concludes and identifies some areas for future research.

II. BRIEF BACKGROUND: THEORETICAL SIGNIFICANCE OF THE COASE THEOREM

Before proceeding, it is worth taking a moment to explain the wider significance of the Coase Theorem in 'law and economics' and legal studies generally. From a theoretical or academic perspective, the Coase Theorem

[1] RH Coase, 'The Problem of Social Cost' (1960) 3 JLE 1.
[2] ibid 2-8. See also RH Coase, 'The Federal Communications Commission' (1959) 2 JLE 1, 25-26.

is crucial to economic analysis of law. According to Richard Posner, for example, 'The most celebrated application of the concept of opportunity cost in the economic analysis of law is the Coase Theorem'.[3] Remove or disprove the Coase Theorem, and the economic approach to law is reduced to intellectual rubble or just another untestable or normative legal theory.[4] But with Coase's logical Theorem as its underlying theoretical foundation, the economic approach not only provides a clear and cogent lens for engaging in descriptive work and for understanding the effect of law on markets; at the same time, it also offers a powerful and forward-looking program for explaining and reforming almost all aspects of the legal system as well as myriad legal institutions, including property rights, tort law, and contracts.[5]

Moreover, the Coase Theorem has major theoretical and even practical implications as well by exposing the 'reciprocal' nature of economic externalities.[6] That is, the Coase Theorem substitutes the conventional 'victim-wrongdoer' paradigm prevalent in legal studies and moral philosophy with an entirely new and non-normative view of reciprocal conflict.[7] Consider a conflict situation between two parties, A and B. Instead of trying to identify the victim and the wrongdoer to the conflict—the traditional and still dominant method for analyzing conflicts and externalities in both the legal and economics literature—the Coasian approach invites one to see the conflict between A and B as a function of both parties' behavior. On this view, the Coase Theorem is nothing less

[3] Richard A Posner, *Economic Analysis of Law* (3rd edn, Little Brown 1986) 7.

[4] For a small but fairly representative sample of stinging criticisms of the Coase Theorem over the years, see Andrew Halpin, 'Coase's World and Coase's Blackboard' (2011) 31 EJLE 91; Dan Usher, 'The Coase Theorem Is Tautological, Incoherent, or Wrong' (1998) 61 Economic Letters 3; Paul Samuelson, 'Some Uneasiness with the Coase Theorem' (1995) 7 Japan World Economy 1; Daniel Q. Posin, 'The Coase Theorem: If Pigs Could Fly' (1990) 37 Wayne LR 89. For a thoughtful critique (in French) that the Coase Theorem is not a 'theorem', see Elodie Bertrand, 'Le théorème de Coase, une réflexion sur les fondements microéconomiques de l'intervention publique' (2002) 41 Perspectives de la Vie Économique 111.

[5] See generally Stephen G Medema, 'Legal Fiction: The Place of the Coase Theorem in Law and Economics' (1999) 15 Economics & Philosophy 209. See also FE Guerra-Pujol, 'Coase's Paradigm' (2010) 1 Indian JLE 1, 27-30.

[6] The word 'reciprocal' appears for the first time in the economics literature in Coase (n 1) 2 and in Coase (n 2) 26. See also Guido Calabresi, 'Neologisms Revisited' (2005) 65 Maryland LR 736, 738.

[7] For a novel application of Coase's reciprocal conflict idea to a science-fiction context, see FE Guerra-Pujol and Orlando I Martinez-Garcia, 'Clones and the Coase Theorem' (2011) 2 JL Social Deviance 43.

than a paradigm shift, a new way of looking at conflict situations.[8] Before Coase, the central question in legal studies used to be: Who is responsible for the harm? After Coase, the interesting and relevant question becomes: Who can mitigate or avoid the harm at the lowest cost to society? And thus one of the most intriguing and counterintuitive insights of the Coase Theorem is that, oftentimes, it is the ostensible victim who can avoid the harm at the lowest cost.

III. COASE'S ARITHMETICAL MODELS OF THE COASE THEOREM (STRAY CATTLE AND RAILWAY SPARKS)

Given the theoretical importance of the Coase Theorem, we present some simple game-theoretic models of Coase's Theorem in Section 4 of the paper. Since our models of the Coase Theorem are based in large part on Coase's analysis of the problem of railway sparks and his model of farmer-rancher interactions,[9] we briefly review the most salient features of Coase's models in subsections 3.1 and 3.2 below.

1. *Stray Cattle*

We begin by discussing Coase's farmer-rancher model, or what one scholar has dubbed 'the Parable of the Farmer and the Rancher'.[10] Coase introduced this model in his classic paper 'The Problem of Social Cost' to provide a vivid and concrete illustration of 'the problem of harmful effects'.[11] Although Coase's social cost paper contains many other examples of the problem of harmful effects—such as railway sparks, airplane noise, and smoking chimneys—it is the farmer-rancher problem that has captured the imagination of many scholars. Here, we describe the essential features of Coase's farmer-rancher model and summarize Coase's results in order to place our models of the Coase Theorem in their proper context.

Coase presents his farmer-rancher model in the opening pages of his social

[8] For an extended discussion of 'paradigms' and 'paradigm shifts', see Thomas S Kuhn, *The Structure of Scientific Revolutions* (3rd edn, University of Chicago Press 1996) 77-91. See also Guerra-Pujol (n 5) 1-7.

[9] Coase (n 1) 2-8 (stray cattle), 29-34 (railway sparks).

[10] Robert C Ellickson, *Order without Law: How Neighbors Settle Disputes* (Harvard University Press 1991) 2. See also Robert C Ellickson, 'Of Coase and Cattle: Dispute Resolution Among Neighbors in Shasta County' (1986) 38 Stanford LR 623, 624-629.

[11] Coase (n 1) 1. Notice that the problem of harmful effects is more often referred to as 'negative externalities' or 'spillover effects' in the economics literature and is an important theoretical and practical problem in legal studies and in economics. For the standard economic analysis of harmful effects or 'negative externalities', see Paul A Samuleson and William D Nordhaus, *Economics* (19th edn, special India edn, McGraw-Hill 2010) 44-45.

cost paper as follows: 'Let us suppose that a farmer and a cattle-rancher are operating on neighboring properties. Let us further suppose that, without any fencing between the properties, an increase in the size of the cattle-rancher's herd increases the total damage to the farmer's crops'.[12] In other words, although the rancher's business is socially useful, his cattle-ranching activities may harm his neighboring farmer because stray cattle may often invade the farmer's land and destroy the farmer's crops. Coase also notes that this harm increases with the size of the rancher's herd, and he illustrates the link between the magnitude of the externality and the size of the rancher's herd with a simple arithmetical table.[13]

Next, having framed the essence of the problem—cattle versus crops—Coase isolates the two most essential features of his model: transaction costs, and institutions or legal rules. Generally speaking, transaction costs refer to the costs of negotiating and enforcing a fencing agreement between the farmer and rancher. Notice that transaction costs are either high or low relative to the costs of the externality to be avoided, that is, the value of the damaged crops when stray cattle invade the farmer's land. In general, transaction costs are low when the private costs of reaching and enforcing a fencing agreement are less than the costs generated by the externality. By contrast, transaction costs are high when the costs of the fencing agreement exceed the harm to be avoided.

Institutions refer to the rules of the game, that is, the rules of legal liability for crop damage caused by stray cattle. In this case, there are two possible institutions or legal rules to deal with the problem of stray cattle: a 'fence-in' rule, or an alternative 'fence out' rule. In summary, the fence-in rule is pro-farmer because it imposes liability for crop damage on the rancher. The rancher must fence-in his cattle or he will be liable for the crop damage caused by his stray cattle. Thus the rancher assumes the cost of fencing under the fence-in rule. The fence-out rule, by contrast, has the opposite effect. It is a pro-rancher rule because it imposes the cost of fencing on the farmer instead of the rancher: it is the farmer who is required to fence-out his neighbor's cattle under a fence-out regime.

In summary, Coase's farmer-rancher model is thus useful for two reasons. First, his model isolates two key variables—transaction costs and legal rules—and asks, what effect, if any, will these variable have on the allocation of resources among crops and cattle? Given these two key variables, there are four possible scenarios in all:

---

[12] Coase (n 1) 2-3.
[13] ibid 3.

Scenario #1	Low Transaction Costs and a Pro-Farmer Rule (fence-in)
Scenario #2	Low Transaction Costs and a Pro-Rancher Rule (fence-out)
Scenario #3	High Transactions Costs and a Pro-Farmer Rule (fence-in)
Scenario #4	High Transaction Costs and a Pro-Rancher Rule (fence-out)

Second, Coase's model is falsifiable, for Coase is, in effect, making a prediction or conjecture regarding what effect these two basic variables will have on the total allocation of resources (ie cattle versus crops). Moreover, the results of Coase's model are startling and surprising: the allocation of resources will depend entirely on the presence or absence of transaction costs and not on the legal rules, and it is this counterintuitive conclusion that is referred to formally as the 'Coase Theorem' in the academic literature.

Nevertheless, although the logic of Coase's model is unassailable, the premises of his model, such as the existence of transaction costs, are not stated formally or expressed mathematically. And although Coase relies on a simple arithmetical table to illustrate the logic of his model, the parameter values in his make-believe arithmetical table are arbitrary and artificial, a problem that plagues most restatements of the Coase Theorem.

2. *Railway Sparks*

Next, we turn to Coase's analysis of railway sparks, for Coase himself devotes considerable space in his social cost paper to the problem of railway sparks.[14] In summary, Coase introduces the problem of railway sparks by reference to 'Pigou's example of uncompensated damage to

[14] As an aside, it is interesting to note that Professor Coase devotes as much space in his social cost paper to railway sparks (about five full pages) as he does to the problem of stray cattle (seven pages). See Coase (n 1) 2-8 (cattle trespass), 29-34 (railway sparks). On a more personal note, the author also fondly recalls that his torts professor, Guido Calabresi, often referred to the problem of railway sparks in his lectures on tort law during the fall semester of the 1990-1991 academic year at Yale Law School.

surrounding woods by sparks from railway engines'.[15] [16] That is, the problem here is that (i) railway lines run through agricultural lands, and (ii) locomotive engines, especially when they run at high speeds, emit dangerous sparks, and (iii) these sparks may, in turn, produce destructive fires.

Coase's analysis of railway sparks—like his analysis of cattle trespass—is insightful, creative, and surprising. In place of a static analysis of the problem, Coase recognizes that the problem of railway sparks is really a strategic one, for the extent of the harm or damages caused by such sparks is the product of a joint interaction.[17] In summary, the harm caused by railway sparks is not only a function of economic decisions made by the railway company, such as whether to install spark preventers or the number of trains to run per day. This harm is also a function of decisions made by the landowners of property adjoining the railway line, such as whether to plant fire-resistant crops or whether to take their lands out of cultivation. Thus, although the problem of railway sparks appears different from the problem of cattle trespass, Coase correctly shows that, from an economic or social cost perspective, both problems are reciprocal and logically the same.

Despite the originality of his analysis, however, Coase does not really present a formal model of the problem of railway sparks, nor does he present a formal mathematical model of harmful effects or externalities generally. Instead, Coase illustrates his analysis of railway sparks with an arithmetical example. Coase himself, however, appears to recognize the limitations of his arithmetical analysis when he states, 'Of course, by altering the figures, it could be shown that there are other cases in which it would be desirable that the railway should be liable for the damage it causes'.[18]

3. *Some Non-Arithmetical Models of Coase's Theorem*

Lastly, before proceeding, we briefly review some previous attempts to

[15] Coase (n 1) 30. As an aside, Pigou was an English economist who had written an influential treatise on welfare economics. See AC Pigou, *The Economics of Welfare* (4th edn, Macmillan 1932).

[16] It is worth noting that Coase refers to 'Pigou's example' not for its own sake, but rather to refute Pigou's approach to economics. In this paper, however, we will not enter into this fray, ie the details of Pigou's approach. For a summary of Pigou's approach, and a critique of Coase's critique of Pigou, see Herbert Hovenkamp, 'The Coase Theorem and Arthur Cecil Pigou' (2009) 51 Arizona LR 633. See also Calabresi (n 6) 738.

[17] Or, in Coase's own words, the problem is a 'reciprocal' one. Coase (n 1) 2.

[18] ibid 33-34.

model Coase's Theorem to set the stage for our models of the Coase Theorem in s 4 below. In summary, although some scholars have tried to formally model the Coase Theorem or test it experimentally, we explain why these previous approaches are deficient.

The literature on the Coase Theorem is vast[19]; in addition, this literature is highly polarized: for every paper in defense of the Coase Theorem, it seems, there is a paper critical of Coase's Theorem. But within this contentious Coasian corpus, formal or analytical models of the Coase Theorem are few and far between. Instead, most analyses, explanations, and extensions of the Coase Theorem (including both defenses and criticisms of Coase's Theorem) are expressed either in arithmetical terms or simply in words.[20]

One notable and early exception, however, is Posner, who presents a graphical analysis of the problem of railway sparks.[21] Since Posner's model is analytical, like the models we present in this paper, it is more general than most statements of the Coase Theorem, which rely on artificial parameter values or fanciful arithmetical tables. The problem with Posner's model, though, is that it is not really Coasian in spirit because his model assumes that only one of the parties is able to avoid the externality in his model. In summary, Posner models the problem of railway sparks in

[19] See, for example, Steven G Medema, 'The Coase Theorem' in Cary L Cooper and Chris Argyris (eds), *The Encyclopedia of Managerial Economics* (Basil Blackwell 1996). In addition, Professor Coase's social cost paper is (still) the most-cited law review article of all time. See Fred R Shapiro and Michelle Pearse, 'The Most Cited Law Review Articles of All Time' (2012) 110 Michigan LR 1483, Table I, 1489 and 1504.

[20] Most of the academic literature in this field restates the Coase Theorem arithmetically with arbitrary or make-believe values. For a small sample this literature, see Varouj Aivazian and Jeffrey L. Callen, 'The Coase Theorem and the Empty Core,' (1981) 24 JLE 175, 176-179; Kenneth R Vogel, 'The Coase Theorem and California Animal Trespass Law' (1987) 16 JLS 149, 159; Robert Cooter, 'The Cost of Coase' (1982) 11 JLS 1, 2-4; A Mitchell Polinsky, *An Introduction to Law and Economics* (2nd edn, Little Brown 1989) 11-14; Stewart J Schwab, 'Coase, Rents, and Opportunity Costs' (1991) 38 Wayne LR 55, 73-74; Stephen G Medema, 'Legal Fiction: The Place of the Coase Theorem in Law and Economics' (1999) 15 Economics & Philosophy 209, 214-215. Likewise, for a small sample of the literature in which the Coase Theorem is expressed exclusively in words, see, for example, Richard A Epstein, 'A Theory of Strict Liability' (1973) 2 JLS 151; George J Stigler, 'Two Notes on the Coase Theorem' (1989) 99 Yale LJ 631; Michael R Butler and Robert F Garnett, 'Teaching the Coase Theorem: Are We Getting It Right?' (2003) 31 Atlantic Economic J 133, 133-135; FE Guerra-Pujol and Orlando I Martinez-Garcia, 'Clones and the Coase Theorem' (2011) 2 JL Social Deviance 43, 65-81.

[21] Richard A Posner, *Economic Analysis of Law* (7th edn, Aspen 2007) 52-54. As an aside, Posner's elegant model first appears in print in the third edition of his textbook. See Posner (n 3) 44-46.

which sparks emitted by railroad locomotives cause fires that destroy crops, since the crops of some farmers are planted next to the railroad tracks, that is, within close range of the flying, fire-causing sparks. Posner states that 'changing the number of trains is assumed to be the only way of changing the amount of crop damage'.[22] But Posner's assumption misses the whole point of Coase's analysis, the idea that harms are 'reciprocal': a harm is the product of a joint interaction, such as the railroad company's decision to run a given number of trains per day and the farmer's decision not to plant fire-resistant crops.[23]

Aside from Posner, a few other scholars have also presented non-arithmetical models of the Coase Theorem. Among the most promising such models, we would point out the formal models of Lee and Sabourian,[24] Anderlini and Felli,[25] Acemoglu,[26] and Hurwicz.[27] Leonid Hurwicz, for example, presents an elegant formal of the Coase Theorem, but his model, however, is of limited scope and usefulness, since it assumes zero transaction costs, and as Coase himself has noted, most Coasian interactions (or 'Coasian games') will most often occur under conditions of high transaction costs.[28]

Some scholars have focused on the problem of transaction costs and have tried to formally model the process of Coasian bargaining. For instance, Lee and Sabourian model Coasian interactions as a dynamic bargaining game.[29] In summary, Lee and Sabourian demonstrate that such interactions produce a large number of equilibria and conclude that the Coase Theorem is valid if and only if there are no transaction costs. Of course, in real-world interactions, strategic considerations may often

[22] Posner, *Economic Analysis* (n 21) 53.
[23] That is why our models of the Coase Theorem (see s 4 below) assume, unlike Posner's model, that either party (not just the railroad company, for example) can take steps to avoid or reduce the risk of harm.
[24] Jihong Lee and Hamid Sabourian, 'The Coase Theorem, Complexity, and Transaction Costs' (2007) 135 J Economic Theory 214.
[25] Luca Anderlini and Leonardo Felli, 'Transaction Costs and the Robustness of the Coase Theorem' (2006) 116 Economic J 223.
[26] Daron Acemoglu, 'Why Not a Political Coase Theorem?' (2003) 31 J Comparative Economics 620.
[27] Leonid Hurwicz, 'What Is the Coase Theorem?' (1995) 7 Japan World Economy 49.
[28] See, for example, Ronald H Coase, 'The Coase Theorem and the Empty Core: A Comment' (1981) 24 JLE 183, 187.
[29] Lee and Sabourian (n 24). Their dynamic bargaining model of the Coase Theorem is based on the work of Rubinstein. See Ariel Rubinstein, 'Perfect Equilibrium in a Bargaining Model' (1982) 50 Econometrica 97. See also Lutz-Alexander Busch and Quan Wen, 'Perfect Equilibria in a Negotiation Model' (1995) 63 Econometrica 545.

obstruct Coasian bargains, even when transaction costs are low, especially in situations of bilateral monopoly.[30] For their part, Anderlini and Felli model 'Coasian negotiations' as a two-stage bargaining game with ex ante negotiation costs and show that such ex ante costs may produce a hold-up problem, thus preventing the parties to the negotiations from reaching an efficient Coasian bargain.[31] Their model, however, is also of limited usefulness, since in many real-world interactions, Coasian bargains are often made even under the conditions of their model.

Next, we wish to say a few words regarding Acemoglu's fascinating bargaining model in his 2003 paper, 'Why not a political Coase Theorem?'[32] Although some scholars have attempted to extend the domain of the Coase Theorem to certain specified areas of politics,[33] Acemoglu presents a generalized model of Coasian interactions between rulers and citizens. In essence, Acemoglu presents a model of political bargaining and shows that the applicability of the Coase Theorem to politics is limited because of commitment problems inherent to the political process.[34] But to the extent such commitment problems can be solved, the conclusions of the Coase Theorem would apply, even to the domain of politics.

In any case, it is worth noting that these various bargaining models of Coasian interactions are not really models of the Coase Theorem per se, for these approaches model the decision whether to negotiate and whether to make a Coasian bargain; that is, they model ex post behavior after the externality has occurred. Coase, in contrast, was not concerned with ex post bargaining per se; he was concerned with the ex ante problem of harmful effects, that is, with avoiding or reducing externalities ex ante, either through legal rules or through Coasian bargaining. That is why our models of the Coase Theorem (see s 4 below) are *ex ante* models, not *ex post* models. In other words, we model the decision whether to produce the externality in the first place.

Other scholars, in contrast, have taken an experimental or behavioral approach to the Coase Theorem.[35] That is, instead of attempting to model

---

[30] See Robert Cooter, 'The Cost of Coase' (1982) 11 JLS 1.
[31] Anderlini and Felli (n 25).
[32] Acemoglu (n 26).
[33] See, for example, J Gregory Sidak, 'The Inverse Coase Theorem and Declarations of War' (1991) 41 Duke LJ 325.
[34] For an overview of the commitment problem, see chapter 2 of Thomas C Schelling, *The Strategy of Conflict* (rev edn, Harvard University Press 1980).
[35] For a small sample of this experimental literature, see Daniel Kahneman, Jack L Knetsch, and Richard H Thaler, 'Experimental Tests of the Endowment Effect and the Coase Theorem' (1990) 98 JPE 1325; Glenn W Harrison and Michael McKee,

the Coase Theorem formally, these researchers have tried to test the Coase Theorem experimentally. In summary, these experimental studies purport to test whether Coasian bargains will occur under artificial bargaining conditions with low transaction costs. The problems with the design and implementation of these experimental studies, however, are legion. Among other things, the main problems or design defects with these experimental tests of the Coase Theorem are that the objects subject to bargaining are low-value items, their prices are not set by markets but rather by the authorities conducting the experiments, and the human subjects participating in these experiments are not drawn from a random sample of the population.

Therefore, in place of artificial experimental studies, or complex *ex post* bargaining models, or verbal restatements of the Coase Theorem, or arithmetical analysis with arbitrary values, or instead of simply assuming that the Coase Theorem is true (as the late George Stigler would do[36]), in the remainder of this paper we present a simple analytic and game-theoretic treatment of Coasian games and the Coase Theorem.

IV. COASIAN GAMES

In this paper, a 'Coasian game' refers to any interactive, strategic, or game-theoretic model in which the interests of the players are conflicting due to the presence of negative externalities or harmful effects, such as stray cattle, airplane noise, or railway sparks. First, we present a simple two-player Coasian game in s 4.1 of the paper. Next, we present an even more generalized population model of Coasian interactions in s 4.2. Lastly, we return to Coase's simple model of farmer-rancher interactions and present an alternative farmer-rancher game in s 4.3 below.

1. *A Two-Player Coasian Game with Probabilistic Payoffs*

Our two-player Coasian game consists of a simultaneous-move game in which the players, whom we designate abstractly as Player A and Player B, share a simple strategy set: *cooperate* or *defect*. Our model is based on the following intuition: in the real world, when a person or a firm is engaged in a socially-useful activity, such as cattle ranching, his activity may produce a probabilistic risk of harm. For example, cattle may trespass on a neighboring farm and damage the farmer's crops, unless such crops are resistant to cattle, or a railroad locomotive may emit sparks and produce a

'Experimental Evaluation of the Coase Theorem' (1985) 28 JLE 653; Elizabeth Hoffman and Matthew L Spitzer, 'The Coase Theorem: Some Experimental Tests' (1982) 25 JLE 73.

[36] George J Stigler, 'Two Notes on the Coase Theorem' (1989) 99 Yale LJ 631.

fire, unless landowners next to the railroad tracks avoid storing inflammable substances, such as hay, too close to the railroad tracks. Each player in our model must thus decide whether to *cooperate* by paying a cost to avoid or reduce the risk of a harm, such as damaged crops or to *defect* by not paying any harm-avoiding costs.

Before proceeding, we wish to make an important observation about our model: *both* players—not just the player who is 'causing' the harm in the traditional sense—are able to cooperate by taking steps to avoid harming the other player. For example, consider again Coase's problem of cattle trespass. The rancher can avoid harming the farmer by fencing-in his cattle, but the farmer himself can avoid this harm by fencing-out the cattle or by growing cattle-resistant crops. Likewise, with respect to the problem of railway sparks, the owner of the railroad company may reduce the risk of fires by reducing the speed of the locomotives or by installing costly spark-arresters, but at the same time, landowners can also reduce the risk of fire by not storing any inflammable substances next to the railroad tracks. The larger point is that (i) all these risk-reducing or harm-avoiding measures are costly cooperative measures and (ii) both players (not just the harm-producing player) must decide whether to cooperate or defect. If a player decides to cooperate, that means he is willing to pay a cost to avoid harming the other player; if, however, a player decides to defect, that means he is not willing to pay such a cost and is, in effect, creating a risk that the other player will be harmed.

Now, returning to our Coasian game, recall that both players in our model have to decide whether to cooperate (invest in a harm-avoidance measure to reduce the risk of an externality) or defect (make no such investment in risk reduction). Given this simple strategy set, and given that there are only two players, there are four possible scenarios or Coasian interactions in this Coasian game:

Scenario #1	both players cooperate: a 'cooperation-cooperation' interaction
Scenario #2	player A defects, while player B cooperates: a 'defection-cooperation' interaction
Scenario #3	player A cooperates, but player B defects: a 'cooperation-defection' interaction
Scenario #4	both players defect: a 'defection-defection' interaction

Since this is a game-theoretic model, the payoffs depend on the strategies simultaneously chosen by the players at the beginning of our Coasian game, and the payoffs associated with each possible interaction of the

game may be expressed in 'normal form' as follows:[37]

		Player B *cooperate*	Player B *defect*
Player A *cooperate*		$b - c_1 - pc_2$	$b - c_1 - pc_2$
Player A *defect*		$b - pc_2$	$(1 - p)(-c_2)$

Figure 1

Normal-form payoff table.

where c_1 is the cost of avoiding a given harm (ie the cost of investing in a safety device, such as a spark arrester, or the cost of reducing one's activity level, such as running fewer trains); where c_2 is the cost of the harm if such harm occurs (ie crop damage caused by fires); where p is the probability of such harm occurring; and where b is the benefit of avoiding the harm.[38]

Now that we have defined strategy set of the players (*cooperate* or *defect*) and assigned payoffs, we shall explain the assumptions in our model and explain the logic of each possible interaction of our Coasian game as follows:

First, consider scenario #1: mutual cooperation. For illustration, assume player A is a landowner whose land adjoins a railroad line and player B is a railroad whose locomotives produce sparks. If both players cooperate by

[37] For simplicity, the payoffs expressed in this table are player A's payoffs (i.e the row player's) because player B's payoffs are the same as player's A payoffs when both players cooperate or when both players defect and are the exact opposite of player's A payoffs when the players play different strategies.

[38] Also, notice that the payoffs of this Coasian game – that is, the benefits and costs corresponding with each strategy – are expressed in abstract terms, rather than in arithmetical terms, in order to illustrate the underlying logic and structure of seemingly unrelated problems, such as the problem of cattle trespass, railway sparks, and other harmful effects. In addition, another advantage of expressing these values as abstract values is flexibility and generality; that is, our abstract model permits us to derive results for any actual value that these parameters might take.

investing in safety or reducing their activity levels,[39] then each player's payoff for mutual cooperation is equal to $b - c_1 - pc_2$, where c_1 is the cost of avoiding the harm, and pc_2 the cost of the harm (if it occurs) discounted by the probability of such harm occurring, and b the benefit of avoiding a given harm.[40] Moreover, notice that one of the terms, pc_2, is probabilistic in nature. The probabilistic nature of this cost distinguishes our model from many other game-theoretic models in law and economics in which costs (and payoffs) are usually fixed. In our model, by contrast, the payoffs are probabilistic because investment in a given harm-avoidance measure (eg spark arresters, fences, etc) merely reduces the probability that a harm will occur (eg damaged crops or the payment of money damages) but such investment does not eliminate this risk altogether.

Next, consider scenario #2, a mixed (defection-cooperation) interaction. If player A defects and player B cooperates,[41] then player A's 'temptation payoff' is $b - pc_2$, while player B's 'sucker's payoff' is $b - c_1 - pc_2$. The logic of these payoffs is as follows: player B receives a 'temptation' payoff $b - pc_2$ because he gets the benefit of player A's costly investment in harm-avoidance without having to pay this cost himself, but player A's payoff is $b - c_1 - pc_2$ because he ends up paying the cost of avoiding the harm. (Again, notice that the last term, pc_2, of both players' payoffs is probabilistic for the same reasons stated in the paragraph above.[42]) Now, in contrast to the scenario above, consider the converse situation (scenario #3). That is, if player A cooperates instead of defecting, and player B defects instead of cooperating, then the payoffs of the players are reversed: player A now receives the payoff $b - c_1 - pc_2$, while player B receives the temptation payoff $b - pc_2$ because in this case it is player B who avoids having to pay c_1, the cost of avoiding the harm.

Before proceeding, the reader may ask: if only one player is willing to invest in a costly harm-avoidance measure (as in scenarios #2 and #3 above), why does the term p, the probability of avoiding the harm, remain

[39] For example, player A, the landowner, cooperates by planting fewer crops next to the railroad line, while player B, the railroad, cooperates by installing costly spark arresters on its locomotives or by operating fewer locomotives.

[40] For player A, the landowner, b might consist of the value of reducing the risk of harm to his crops. For player B, the railroad, b might be value of avoiding the risk of a lawsuit from the landowner.

[41] For example, player A, the landowner, might decide to defect by planting inflammable crops next to the railroad line (thus increasing the risk of harm to his crops from railway sparks). In contrast, player B, the railroad, might nevertheless decide to cooperate by installing spark arresters to reduce the risk of fires.

[42] That is, player B's costly investment in a given harm-avoidance measure merely reduces the risk that such harm will occur, it does not eliminate this risk.

the same as when both players invest in harm-avoidance measures separately (as in scenario #1)? That is, why is the probability of avoiding a given harm constant? In our model, we assume for the sake of simplicity that when one player invests in safety, any additional investment in safety by the other player does not further reduce the probability of harm.[43] That is, we assume that when both players invest in safety, their joint investment is redundant.

Lastly, consider scenario #4: mutual defection. What are the payoffs when both players defect, that is, when neither player A, the landowner, nor player B, the railroad, invests in any harm-avoidance measure or reduces their activity levels? In our model, both players forego the benefit b and avoid paying the harm-avoidance cost c_1 and instead receive a mutual defection payoff consisting of $(1-p)(-c_2)$.[44] In essence, the players 'take their chances' when neither invests in safety or reduces their activity levels. The logic of this mutual defection payoff is as follows: when neither player is willing to invest in a costly harm-avoidance measure, then this set of choices creates a probabilistic risk that a harm will occur, and moreover, we further assume for simplicity that this probabilistic risk is equal to $1-p$. That is, we assume that if p is the probability of harm when at least one of the players pays a cost to avoid that harm, then the probability of harm must be $1-p$ when no one invests in safety.[45]

Given this payoff structure, and given our simplifying assumptions, which of these four Coasian scenarios is most likely to occur? Put another way, what is the optimal strategy or best response from the point of view of each Coasian player? Is there a stable Coasian equilibrium?

---

[43] In reality, such additional investment in harm-avoidance may reduce the risk of harm by some linear or marginally-declining amount, but we make the assumption of redundancy to keep our Coasian model as simple as possible.

[44] Notice that the mutual defection payoff is a function of c_2, not c_1. As one anonymous referee of this paper noted, making the mutual defection payoff a function of c_1 is problematic (and artificial) because c_1 refers to the cost of prevention, not the cost of the harm.

[45] This assumption, however, is open to debate. For instance, as one anonymous referee of this paper noted: if the probability of crop damage is only 0.1 when one of the parties builds a fence, then the probability of crop damage without a fence is not necessarily 0.9. (It could very well be higher or lower than 0.9 depending on the specifics of the situation.) Nevertheless, we make this simplifying assumption for ease of exposition and convenience, since our general assumption is that investment in safety tends to reduce the probability of harm, while the lack of such investment tend to increase this probability. Also, notice that if the magnitude of the harm to be avoided were less than the cost of avoiding it, then it would not make sense to invest in the harm-avoidance measure in the first place.

For his part, Coase famously asserts in his social cost paper that the players will negotiate and strike a Coasian bargain to solve the reciprocal harm/harm-avoidance problem, but only when transaction costs are zero.[46] This is the core of the Coase Theorem. But what happens when transaction costs are high, or when strategic behavior prevents the formation of Coasian bargains even when transaction costs are low?

If we take another glance at the game tree or at the payoff table of our Coasian game, the equilibrium path is not obvious. Since the payoffs are probabilistic and are expressed in variables, it is difficult to tell whether there are any dominant or dominated strategies or what the best responses of the players are. As a result, we will re-introduce the concept of probability, as well as the related idea of an 'expected payoff', in order to solve this game and find the existence of any possible equilibria.

Consider player A first.[47] Player A's expected payoff from playing a given strategy (*cooperate* or *defect*) depends on the probability P that player B might also play the same strategy as well as the probability $1–P$ that player B might choose a different strategy.[48]

Recall that player A has two choices in his strategy set. If player A cooperates, he will obtain the payoff $b - c_1 - pc_2$ with probability P (ie the probability that player B also cooperates), and he will also obtain the same payoff, $b - c_1 - pc_2$, with probability $1–P$ (ie the probability that player B defects). Player A's expected payoff of cooperating, which can be written as $E(C)$, is expressed formally as follows:

$$E(C) = (b - c_1 - pc_2)(P) + (b - c_1 - pc_2)(1–P)$$
$$E(C) = Pb - Pc_1 - Ppc_2 + b - c_1 - pc_2 - Pb + Pc_1 + Ppc_2$$
$$E(C) = b - c_1 - pc_2$$
(1.1)

In other words, when player A cooperates by investing in safety or reducing his activity level, his payoff is constant regardless of what player B does. By contrast, player A's expected payoff from defecting does depend on what player A does. In summary, player A receives the payoff $b - pc_2$

[46] That is, when 'the pricing system works smoothly'. Coase (n 1) 5.

[47] In fact, the analysis in the remainder of this section applies equally to both players since, for as we stated earlier in n 37, the payoffs in our simple model are symmetrical.

[48] For reference, notice that this type of probability (ie the probability P of the other player's strategy selection) is written as a capital letter to distinguish it from the earlier type of probability, that is, the probability p that a harm will occur if one or both of the players invests in safety.

when player B cooperates and the payoff $(1-p)(-c_2)$ when Player B defects. Since player B will cooperate with probability P and defect with probability $1-P$, we can express player A's expected defection payoff $E(D)$ as follows:

$$E(D) = (b - pc_2)(P) + [(1-p)(-c_2)](1-P)$$
$$E(D) = Pb - Ppc_2 + (pc_2 - c_2)(1-P)$$
$$E(D) = Pb - Ppc_2 + pc_2 - c_2 - Ppc_2 + Pc_2$$
$$E(D) = Pb - c_2(2Pp + P + p - 1)$$
(1.2)

What if we assume that $pc_2 = 0$ for simplicity; that is, what if we assume that the probability of harm is low, close to zero, when at least one of the players invests in safety or reduces his activity level. Under this assumption, player A's expected cooperation payoff is

$$E(C) = b - c_1$$
(1.1a)

and player A's defection payoff $E(D)$ becomes:

$$E(D) = Pb - c_2 + Pc_2$$
$$E(D) = P(b + c_2) - c_2$$
(1.2a)

Notice that the size of player A's defection payoff (equation 1.2a) is a function mostly of P, the probability the player A will cooperate. By contrast, player A's cooperation payoff is a function only of the terms b and c_1. In other words, player A's best response depends mostly on what strategy player B chooses. If player B decides to cooperate, ie $P = 1$, then $E(D)$ will be greater than $E(C)$ because player A's expected payoff for defecting will be b, while his expected payoff for cooperating will remain $b - c_1$.

Now, assume that player B decides to defect, ie $P = 0$. In this case, $E(C)$ will be greater than $E(D)$ because player A's expected payoff for defecting will be $-c_2$, while his expected payoff for cooperating will remain $b - c_1$. Assuming that $b > c_1$,[49] Player A should cooperate when B defects, and conversely, player A should defect when B cooperates.

[49] This is a reasonable assumption, since otherwise, it would not make sense to invest in safety when the cost of such investment is greater than the benefit to be received from such investment.

In other words, our model shows that player A's decision to cooperate or defect is not so much a function of legal rules or transaction costs but of player B's choice, which in turn is a function of player A's decision.[50] Is there any way around this circular result?

One possible solution is to deny the Coase Theorem: in the absence of an equilibrium solution to our Coasian game, the choices of both players might then be a function of the legal rules, contra the invariance thesis of the Coase Theorem. On this view, legal rules are a device for coordinating the choices of the players, specifically, a device for getting at least one of the parties to invest in safety or reduce his activity level. This analysis also confirms a central axiom of law and economics, that the applicable legal rule should impose liability on the party with the lowest cost of avoiding the harm.[51] Thus, under the assumptions of our model, we would expect the rule of legal liability to matter, even under conditions of low transaction costs.

2. *An n-Player Coasian Game with Probabilistic Payoffs*

Next, we present a multi-player evolutionary model of our Coasian game. In summary, our *n*-player evolutionary game works as follows:

(a) There is a large and well-mixed population of players.

(b) This population contains two types of players, cooperators and defectors.

(c) At the start of each round of play, two players are selected at random from the population and then, during each round of play, these two players play a Coasian micro-game.

(d) After each round of play, the player with the highest payoff in the micro-interaction not only survives but also produces a descendant-clone who asexually inherits the victor's player's type (ie if the victor was a cooperator, then his descendant is a cooperator).

(e) The player with the lowest payoff, in contrast, is eliminated from the population.

[50] Also, notice that this analysis is independent of the level of transaction costs.
[51] For the classic 'cheaper cost avoider' theory of tort liability, see Guido Calabresi, *The Costs of Accidents: A Legal and Economic Analysis* (Yale University Press 1970).

(f) Lastly, if the interaction ends in a draw or tie (ie cooperator-cooperator or defection-defection interactions), both contestants survive but neither produces a descendant.

The purpose of this game is to determine which strategy will spread through our population of Coasian players. Will cooperators outperform defectors, or will defectors displace cooperators, or will the population consist of a stable mix of cooperators and defectors? To answer these questions, we proceed in several stages.

First, we restate the expected payoffs corresponding to each possible Coasian micro-game. In summary, there are four possible micro-interactions in the *n*-player evolutionary game, as in the traditional two-player model: (1) mutual cooperation, or C|C for short; (2) cooperation-defection, or C|D; (3) defection-cooperation, or D|C; and (4) mutual defection, or D|D. Since the structure of the payoffs in the *n*-player game are the same as in the two-player game, the payoffs corresponding to each Coasian micro-game are as follows:

E(C|C) = [the payoff to a cooperator given that he interacts with another cooperator] = $b - c_1 - pc_2$
E(C|D) = [the payoff to a cooperator given that he interacts with a defector] = $b - c_1 - pc_2$
E(D|C) = [the payoff to a defector given that he interacts with a cooperator] = $b - pc_2$
E(D|D) = [the payoff to a defector given that he interacts with another defector] = $(1 - p)(-c_2)$

In summary, on the far left- and far-right hand sides of the table above, we have expressed the payoffs corresponding to each Coasian micro-interaction in mathematical form, while in the middle section, separated by brackets [...], we have 'translated' the mathematical notation into plain English for the non-mathematical reader.

Moreover, since this is a population model, the success of a given strategy is said to be 'frequency dependent' because the success or survival rate of a strategy depends not only on the frequency of the other strategy but also on that strategy's own frequency.[52] Since success or 'fitness' (rate of survival) is frequency dependent, we proceed to use the methods of evolutionary game theory to determine whether a strategy is an

[52] For an illustration of frequency dependency, see Richard McElreath and Robert Boyd, *Mathematical Models of Social Evolution* (University of Chicago Press 2007) 38.

'evolutionarily stable strategy' or ESS and to find the long-run evolutionary equilibrium of the population—that is, the frequency of cooperators and defectors over many generations.[53] Specifically, we wish to answer the following key questions: (i) is cooperation an evolutionarily stable strategy or ESS? In other words, are cooperators able to resist invasion by defectors? (ii) Likewise, is defection an ESS? That is, are defectors able to resist invasion by cooperators? (iii) Or, do Coasian interactions produce an evolutionarily stable mix of cooperators and defectors?

Let P be the frequency of cooperators in the population, and thus let $1 - P$ the frequency of defectors in the population. First, consider a population in which the frequency of cooperators is very high $(P \approx 1)$. With this population structure, cooperators rarely interact with defectors because the frequency of defectors is low $(1 - P \approx 0)$, and thus the average fitness of a cooperator, written as $W(C)$, is determined by his interactions with other cooperators in the population as follows:[54]

$$W(C) = w' + 1[E(C|C)] + (1 - 1)[E(C|D)]$$
$$W(C) = w' + E(C|C) + 0$$
$$W(C) = w' + b - c_1 - pc_2$$

At this point, consider the appearance of a rare defector mutant in this population of cooperators. Will this defector be able to spread across the population, gradually displacing the cooperators, or will the cooperators be able to resist invasion by the defectors? To answer this question, we must determine the average fitness of the rare defectors among the population of cooperators, and then compare the average fitness of such defectors with the average fitness of cooperators. Since defectors are rare $(1 - P \approx 0)$, the chance one defector will meet another defector is likewise small. As a result, the average fitness of a defector, written as $W(D)$, is determined by his interactions with cooperators as follows:

$$W(D) = w' + 1[E(D|C)] + (1 - 1)[E(D|D)]$$
$$W(D) = w' + E(D|C) + 0$$
$$W(D) = w' + b - pc_2$$

Thus, when we compare the average fitness levels of the majority

[53] For an overview of the ESS concept, see generally John Maynard Smith, *Evolutionary Game Theory* (Cambridge University Press 1982). See also George R Price and John Maynard Smith, 'The Logic of Animal Conflict' (1973) 246 Nature 15.

[54] Before proceeding, note that the parameter w' in our equations refers to the 'baseline fitness' or baseline survival rate of all the individuals in the population—that is, the probability of survival from generation to generation—and thus reflects the strength of selection on a given population. See McElreath and Boyd (n 51) 40-41.

cooperators and the rare defectors, we see that defectors have a higher average fitness than cooperators. Stated formally, we see that $W(D) > W(C)$ because $b - pc_2 > b - c_1 - pc_2$.[55] This means that defectors will outperform cooperators and thus spread across and invade the population of cooperators.

But now this state of affairs raises a new question: can a population of defectors resist invasion by cooperators? Consider a population in which the frequency of defectors is high $(1 - P \approx 1)$. With this population structure, defectors interact with other defectors most of the time, so the average fitness of a defector, W(D), is determined by his interactions with other defectors as follows:

$$W(D) = w' + 1[E(D|D)] + (1 - 1)[E(D|C)]$$
$$W(D) = w' + E(D|D) + 0$$
$$W(D) = w' + (1 - p)(-c_2)$$
$$W(D) = w' + pc_2 - c_2$$
$$W(D) = w' + c_2(p - 1)$$

Next, consider the appearance of a rare cooperator mutant in this Hobbesian population of defectors. Will the rare cooperators be able to invade the population and displace the defectors, or will the defectors be able to resist invasion by the cooperators? To answer this question, we must compare the average fitness level of the rare cooperators with that of the majority defectors. Since cooperators are rare $(P \approx 0)$, the average fitness of a cooperator is thus determined by his interactions with defectors as follows:

$$W(C) = w' + 1[E(C|D)] + (1 - 1)[E(C|C)]$$
$$W(C) = w' + E(C|D) + 0$$
$$W(C) = w' + b - c_1 - pc_2$$

Now, when we compare the average fitness levels of the majority defectors and rare cooperators, we see that the rare cooperators have a higher average fitness than the majority defectors do. This result also raises an intriguing question: will the population of cooperators and defectors continue to cycle depending on which group is in the majority, or is there an evolutionarily stable mix of cooperators and defectors?[56]

In any case, how does this result relate to Coase's Theorem? In summary,

---

[55] Notice that the baseline fitness terms, w', cancel out.

[56] One could easily find for this equilibrium mix of defectors and cooperators by setting $W(C)$ equal to $W(D)$, substituting p' for p, and solving for p'.

our result shows another dimension of the Coase Theorem. Recall that Coase himself was concerned with negative externalities, or 'the problem of harmful effects'.[57] Most of the literature on the Coase Theorem focuses on transaction costs, legal rights, bargaining, the endowment effect, and willingness to pay, and thus most commentators tend to focus exclusively on law, behavioral economics, or on economics proper: the benefits and costs of various negative externalities, such as the harmful effects produced by cattle ranching, crop farming, railroads, and so forth. In brief, the Coase Theorem asks two basic questions: what is the harm, and who should pay the cost to avoid this harm? Thus, under traditional economic or Coasian analysis, once the harm has been identified, the main questions are always economic in nature: 'who pays whom?'

Our analysis, in contrast, raises a different set of questions. Instead of 'who pays whom?', our analysis asks: which harm-avoidance measure more effectively reduces the probability or risk of harm? Unlike traditional economic or Coasian analysis, our analysis shows that what really matters is not the (social or private) benefits generated by a conflicting activities and not the (social or private) costs imposed by such activities, but rather what really matters is the effectiveness of the harm-avoidance measures that are available to the parties to address a given harm, and this insight is captured by the probabilistic payoffs, namely, the parameter p, in our models of Coasian games.

This insight is not necessarily inconsistent with main results of the Coase Theorem: the invariance thesis and the efficiency thesis. For example, the efficiency criterion is consistent with the proposition that courts and legislatures should impose legal liability on the party that can most effectively reduce the probability of a given harm, but notice that our emphasis is not on the cost of avoiding a given harm but rather on the probability of avoiding such harm. In many cases, cost and probability will be close proxies for each other, but in other cases, these issues may diverge: the 'cheaper cost avoider' may not necessarily be able to reduce the probability of a given harm more effectively than another party might. In other words, a different party might able to reduce the risk of such harm more effectively (although at greater cost) than the designated cheaper cost avoider. This possibility opens up a new avenue of research, a new door for the Coase Theorem to open.

[57] Coase (n 1) 1.

3. *An Alternative Coasian Game with High Transaction Costs, Fixed Legal Rules, and Deterministic Payoffs*

Lastly, we model a farmer-rancher game with high transaction costs (ie no Coasian bargaining among the players) and with fixed legal liability rules but with deterministic (non-probabilistic) payoffs. For this revised Coasian game, we now add the following assumptions:

> (a) The population is large, well-mixed, and composed of two ideal types: farmers *F* and ranchers *R*.
>
> (b) Individuals from this large, well-mixed population are selected at random and interact in pairs.
>
> (c) Individuals are not permitted to make side deals or Coasian bargains with each other (that is, we assume high transaction costs).
>
> (d) In the absence of a legal rule, no fence gets built.
>
> (e) When a legal rule is enacted (either fence-in or fence-out), there is full compliance with the rule; that is, if the rule is fence-in, all cattle ranchers comply with the rule and fence-in their lands, and by the same token, if the rule is fence-out, all farmers comply with the rule and fence-out their lands.
>
> (f) The cost of fencing is constant and the payoffs to farming and ranching are equal, or stated formally, $bR=bF$.

As before, we recognize that these simplifying assumptions are not necessarily consistent with real-world conditions. For example, in a real-world situation, the cost of fencing will vary depending on the size of one's land, and the revenues generated by farming and ranching will likewise vary depending on a wide variety of factors. Nevertheless, we make these artificial and unrealistic assumptions to simplify our mathematical analysis and test the main insight generated by the Coase Theorem: the conjecture that the rules of the game will have no effect on the allocation of resources when transaction costs are high.

Now, consider a large, well-mixed population consisting of farmers and ranchers. Ranchers receive a fixed payoff of bR, while farmers receive a fixed payoff of $bR - dp$, where d is the cost of the damages or harm to crops caused by stray cattle, p is the probability that this harm will occur (in the absence of a fence), and $dp > 0$. For now, assume there is no fencing rule or convention in place and that $bR=bF$.

Given this set of assumptions, we see that ranching is an evolutionarily stable strategy or ESS since the ranching payoff exceeds the farming payoff, since by definition $bR > bF - dp$. As a result, the population dynamic will be pro-rancher: when ranchers are common, farmers will not be able to invade a population of ranchers, and when farmers are common, ranchers will be able to invade the population and displace the farmers, and so either way, ranchers will always dominate the population in the absence of any fencing rule or convention.

But now consider what effect a fencing rule would have on our model. There are two possible rules: fence-in and fence-out. Assume both fencing rules are equally effective in solving the problem of stray cattle, so the main effect of either rule is simply to rearrange the payoff structure of farmer-rancher interactions, since fences are costly to build and maintain. Specifically, under a fence-in regime, the payoff to a rancher is $V(R)=bR - c$, where c is the cost of fencing-in the rancher's land, and likewise, the payoff to a farmer is $V(F)=bF - (1 - p)d$, where this last term is the probability that the farmer's crops are damaged even with a fence in place. To keep this model as simple as possible, we will assume that the fence-in rule neutralizes the problem of stray cattle, that is, we assume that $(1 - p)d = 0$. To recap, then, when ranchers are required to fence in their cattle, an individual rancher's payoff is reduced by the cost of fencing-in his land, while farmers receive a fixed payoff bF since the fence-in rule neutralizes the problem of stray cattle, ie $(1 - p)d = 0$.

Given a fence-in rule, we now see that farming will be an ESS because the farming payoff exceeds the ranching payoff, or $bF > bR - c$. Thus, under a fence-in regime, the proportion of farmers in the population will increase in frequency over time. This means the following population dynamic will occur: when farmers are common, ranchers will not be able to invade a population of farmers, but when ranchers are common, farmers will always invade the population and displace the ranchers.

Next consider, what happens when the applicable rule is fence-out, instead of fence-in. Under a fence-out regime, the payoff to a rancher is $V(R)=bR$, while the payoff to a farmer is $V(F)=bF - c$, since now it is the farmer who must pay the fencing costs.[58] In summary, given a fence-out rule, ranching will be an ESS because the ranching payoff exceeds the farming payoff, that is, $bR > bF - c$. As in the case with no legal rule, the population

[58] Again, for simplicity, we assume that $bR = bF$, ie, there is no reason to prefer farming over ranching or vice-versa, and we assume that the fence-out rule solves the stray cattle problem, so we ignore d.

dynamic will be pro-rancher: when ranchers are common, farmers will not be able to invade a population of ranchers, and when farmers are common, ranchers will always be able to invade the population and displace the farmers.

In summary, the lesson of this Coasian game is clear: the dynamic of the population over time is a function of the rules. That is, when transaction costs are high, or when Coasian bargaining is not possible, the payoffs of the players, and thus the outcome of the game, is dependent on the legal rule. This result thus confirms one of the conclusions or predictions of the Coase Theorem: when transaction costs are high, the choice of legal rule will determine the allocation of resources.

V. Conclusion

In closing, we concede that the Coasian games presented in this paper abstract from reality. Specifically, our models of the Coase Theorem are much more abstract and idealized than actual or real-life farmer-rancher interactions in many respects: the population of farmers and ranchers in our models are well-mixed and large, their corresponding strategies are simple and stylized, and the payoffs to each strategy are kept constant. In addition, we have omitted stochastic effects such as noise or errors from our model. Instead, we have decided to trade off realism for tractability. That is, we have intentionally designed our model of farmer-rancher interactions to be as simple as possible to illustrate the logic of the Coase Theorem.

We now wish to close this paper by looking towards the future and sketching some other possible Coasian games. Specifically, we briefly consider some variations to our model of the Coase Theorem and identify some new questions for future research:

Question #1 What happens when $bR \neq bF$?

One direction for future work is to relax the assumption of equal payoffs, such as making the payoffs vary inversely with the choice of legal rule. For example, with a pro-farmer fence-in rule, a rancher might respond by investing less in ranching (eg by decreasing his herd from n steer to $n - 1$ steer), while farmers might respond by investing more in farming (by planting more crops), and this change in investment levels will, in turn, affect the expected payoffs corresponding with each activity.

Question #2 What happens when the choice of legal rule is endogenous to the model?

That is, what happens when the players must not only decide how much to invest in farming or ranching but must also decide how much to invest in rent-seeking activities, such as lobbying or litigation, in order to obtain a favorable legal rule. Now, the payoffs of the players will be a function of their farming or ranching activities; their payoffs will also be a function of their lobbying or litigation activities as well, and since activities like lobbying and litigation tend to increase the probability of a favorable ruling, such a possibility also lends itself to a probabilistic analysis.

Question #3 What happens if we assume a different population structure?

That is, instead of assuming a large and well-mixed population, as we have done in this paper, what if we were to model the population structure graphically? For example, imagine a large number of evenly-sized towns distributed over a large square grid. Each town contains n number of plots of land with n number of farmers and ranchers, and in addition, each town must decide whether to adopt a pro-farmer rule (fence-in) or a pro-rancher rule (fence-out), with the choice of legal rule depending on which group is a majority in each town. Given this graphical configuration of the problem, we would then find what mix of pro-farmer and pro-rancher rules will result over the long run. That is, instead of modelling a population of farmers and ranchers, we would model a population of legal rules, with feedback effects between the population of legal rules and the population of farmers and ranchers in each town, since the choice of legal rule depends on the population dynamic in each town, and since the population dynamic in turn, depends on the choice of legal rule. Such an approach to the Coase Theorem, one with feedback loops between the legal rules and the economic activities of the actors, seems to be an especially promising area for future Coasian analyses.