



A quantitative quest for philosophical fairness in EU's competition procedure

Haukur Logi Karlsson

Thesis submitted for assessment with a view to obtaining
the degree of Doctor of Laws of the European University Institute

Florence, 15 September 2017

European University Institute
Department of Law

A quantitative quest for philosophical fairness
in EU's competition procedure

Haukur Logi Karlsson

Thesis submitted for assessment with a view to obtaining
the degree of Doctor of Laws of the European University Institute

Examining Board

Professor Giorgio Monti, EUI (supervisor)

Professor Dennis Patterson, EUI

Professor Davíð Þór Björgvinsson, University of Iceland

Professor Ioannis Lianos, UCL

© Haukur Logi Karlsson, 2017

No part of this thesis may be copied, reproduced or transmitted without prior
permission of the author

**Researcher declaration to accompany the submission of written work
Department of Law – LL.M. and Ph.D. Programmes**

I Haukur Logi Karlsson certify that I am the author of the work "A quantitative quest for philosophical fairness in EU's competition procedure" I have presented for examination for the Ph.D. at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the Code of Ethics in Academic Research issued by the European University Institute (IUE 332/2/10 (CA 297)).

The copyright of this work rests with its author. Quotation from this thesis is permitted, provided that full acknowledgement is made. This work may not be reproduced without my prior written consent. This authorisation does not, to the best of my knowledge, infringe the rights of any third party.

I declare that this work consists of 114.256 words.

The work was not recommended for language correction and thus appears here in its original form.

The work was funded through a project grant from the Icelandic Research Fund, grant no. 141274-053.

Reykjavík, 9 May 2017



Haukur Logi Karlsson

Abstract

The question of procedural fairness in EU's competition procedure has been discussed widely in the academic literature based on the traditional positivistic legal method; so far without a success in producing a consensus on where the practical limitations of the concept of procedural fairness ought to lie. This thesis sets out to approach the problem more fundamentally by propping beyond the concept of procedural fairness in the legal positivistic sense, and venture into the territory of moral and political philosophy for establishing a practical understanding of the more general concept of fairness in human relations. Once the concept of fairness has been properly revealed in practical terms, the thesis attempts to quantitatively translate this concept of fairness into the laws to facilitate the composition of a fair legal rule. To achieve this, a novel methodological model is constructed based on microeconomic tools. This model, the model of fair rules, is then used to assess two dilemmas of procedural fairness in the context of EU's competition procedure that have been solved by the CJEU based on the traditional juridical method. The results of the assessment suggest that methodological improvements can be made in the design of competition procedures with regards to facilitating procedural fairness. Such improvements would also have implications for the legal interpretive methodologies used by the EU courts.

Table of contents

Preface.....	i
01 The problem with fairness in EU's competition procedure.....	1
1. The elements	4
2. The problem.....	11
3. The academic debate and the Court's solution	13
4. The question and a path to an answer	23
02 The dark matter of law.....	27
1. Modern legal positivism and its greatest critic	28
2. The problem with the separation thesis	31
3. Beyond the rule of recognition	35
4. The social contract.....	41
4.1. Origins and evolution of the social contract.....	41
4.2. Social contracts and the theory of games.....	50
4.3. The ultimate priori of the social contract	55
4.4. Institutionalising the social contract	59
5. Law's dark matter	60
03 Tools for translating optimality into laws	65
1. Traditional tools for translating optimality	65
1.1. Optimising towards economic efficiency.....	66
1.2. Optimising towards moral appropriateness	70
1.3. The benchmark for possible methodological improvements	76
2. Alternative tools for translating optimality	77
2.1. The basics of decision theory	80
2.2. Risk, uncertainty, and the decision matrix.....	85
2.3. Expected utility theory	88
2.3.1. Von Neumann and Morgenstern	88
2.3.2. Savage and Harsanyi	90
2.4. Prospect theory	93

2.4.1.	Empirical invalidity of the rationality assumption	94
2.4.2.	The alternative model of prospect theory.....	95
2.4.3.	The location of the status quo.....	99
3.	Improving the traditional by using elements of the alternative	101
04	The model of fair rules	103
1.	The social contract and the role of the legislator	103
2.	Legislative actions as decision problems.....	108
3.	The elements of the model of fair rules.....	109
3.1.	The plan of the legislator	109
3.2.	Stakeholders and the reduction of moral and efficiency claims	111
3.3.	Interpersonal comparison of stakeholder claims through a preference index ..	113
3.4.	Actual stakeholder preferences registered into payoff matrices.....	119
3.5.	Identifying the fair and efficient legislative action	120
4.	The utility of a model for identifying fair rules	124
05	The essence of a procedure	127
1.	The procedure as an instrument of law	128
1.1.	Civil procedures as an instrument.....	130
1.2.	Criminal procedures as an instrument.....	136
1.3.	Administrative procedures as an instrument	139
1.4.	The instrumental essence of a law enforcement procedure.....	143
2.	The normativity of a procedure	147
2.1.	The deontological view	148
2.2.	The consequentialist view	153
2.3.	The normative essence of a procedure	159
3.	A procedure is	161
06	The norm and the process of EU competition law.....	169
1.	Recap of the historical development.....	170
2.	The process of competition	182
3.	The current consensus on the objective of EU's competition policy	186

07 Stakeholders, decision points, and EU’s competition proceedings	193
1. The stakeholders in a competition proceeding	194
2. The stakes in a competition proceeding	197
3. The procedural provisions in EU competition law	200
3.1. Decision points in EU’s competition law procedure	201
3.2. Instrumental and normative function	207
3.2.1. The procedural regulations	208
3.2.2. The procedural practices of DG Competition	212
3.2.3. The procedure before the Court of Justice of the EU	215
4. The Accuracy of EU’s competition procedure.....	216
08 Applied fairness in EU’s competition procedure	221
1. The case law on Article 47 of the Charter of Fundamental Rights	222
1.1. Review of the cases	222
1.2. The main fairness issues of the case law	232
2. Solving fairness issues by applying the model of fair rules	237
2.1. Case Study I: The issue of fairness in KME-Chalkor	240
2.2. Case Study II: The issue of fairness in Groupe Gascogne	245
2.3. Applying the model of fair rules to KME-Chalkor and Groupe Gascogne	250
2.3.1. Step 1: Identifying the legislative plan	251
2.3.2. Step 2: Preference function of the stakeholders	253
2.3.3. Step 3: Preference index for antitrust procedure	255
2.3.4. Step 4: Preference matrixes for the key stakeholders.....	258
2.3.5. Step 5: Finding the fair antitrust procedure	261
3. Assessing the application of the model of fair rules.....	265
09 Conclusions	267
1. Thesis summary	267
2. Concluding reflections.....	272
Bibliography	277

Preface

The topic of this research and the approach I have taken to address it is a cumulation of more than a decade of working and thinking about the problems involved. During my master studies at the University of Reykjavík I did a course in the autumn of 2006 on the European Convention of Human Rights under professor Oddný Mjöll Arnardóttir. Disconcertingly, at times I felt I could not understand how the Strasbourg judges reached some of their landmark decisions, in contravention to what I had believed until then to be the proper way of approaching the juridical method. This eventually pushed me towards reading more legal philosophy. Resultantly I started to understand the forces at work in elite constitutional decision making, especially the political context that often shapes the outcome of high stake adjudicative processes.

In February 2009, during my LLM studies at Stockholm University, I was introduced to the procedural fairness problems that would become the KME-Chalkor litigation during a lecture given by Nils Wahl, then judge at the Court of First Instance. I was so intrigued by how he presented the problems that I sent him an email the week after and asked him to be my LLM thesis supervisor. He promptly replied, and of course respectfully declined the request, citing workload. Later I was lucky enough to have Ulf Bernitz supervising my thesis, which was very inspiring; being guided by his great academic abilities and kind personality. After my LLM studies I went to Brussels and worked for few years in the EFTA apparatus responsible for the enforcement of the competition provisions of the EEA agreement. Influenced by the KME-Chalkor litigation, the lawyers of the Norwegian company Posten Norge around that time tested the same arguments in the context of the EFTA pillar of the EEA agreement. I only did a very minor research task on that case on behalf of the EFTA Surveillance Authority, but due to its stature within the organisation at the time, I could not escape being involved in the excitement about it and its fate before the EFTA Court.

When drafting my PhD research proposal in late 2010, I decided to focus on the issues that were causing excitement in Brussels at the time regarding the institutional design of the competition law procedure and its compatibility with procedural fairness standards in the wake of entering into force of the Lisbon Treaty. Initially, my idea was inspired by arguments laid out

in a paper by Slater, Thomas, and Waelbroeck (2008), and by opposing arguments made in a paper by Wils (2010). In February 2011, while I was still formulating my proposal, AG Sharpston delivered her opinion in the KME case and after reading her arguments I got convinced that there was job to be done in the field. The gap I identified in the practice and the literature was related to the revelation I had during my ECHR course years earlier; I felt that the very concept of procedural fairness as used by Sharpston in her opinion and as used in the leading academic literature on the KME-Chalkor litigation, was somehow counter intuitive. The aim of the research thus became to explore the philosophical foundations of the concept of procedural fairness and see if philosophy could solve, what the juridical method had in my mind failed to do in a satisfactory manner.

On the back of this proposal I was in May 2011 accepted into EUI's prestigious PhD program after the hardest interview I ever had, conducted by my eventual supervisor professor Giorgio Monti and by professor Dennis Patterson, who I would later learn to appreciate for his direct and methodological way of thinking about law and legal argumentation. Due to romantic reasons involving my now wife Áslaug, I extended the Brussels stay by a year and was granted with a permission to delay by one year the initiation of the PhD program, thus I finally moved to Firenze in September 2012. By that time, the courts in Luxembourg and Strasbourg had taken a stance on the procedural fairness issue in competition proceedings and the main contributions to the literature were out. The stage was thus set for me to see if I could develop my research proposal into a PhD with a modest contribution to the field.

The time at the EUI and in Firenze was nothing short of extraordinary. The Institute a relaxed tranquil working place, but still intensely intellectually stimulating. I will always remember my daily bike rides down the Badia hill in the soft twilight through the olive tree groves and the historic urban streets, with the silhouette of the iconic Duomo dominating the skyline. A perfect reward for a productive day of writing and researching. The relaxed pace of the Florentine way of life also suited me well, and the cheerful attitude of the Florentine people makes me smile when I think about them.

I have been lucky to be guided through the PhD process by my supervisor Giorgio Monti, who has with his unassuming personality and brilliant mind been a perfect role model for the aspiring academic. He has given me many good comments on numerous drafts of the thesis and has taken a light touch approach to the supervision, which has suited my style of working well. My external co-supervisor, Oddný Mjöll Arnardóttir, has also been tremendously helpful during this process. I initially asked her to join the project due to the good intellectual impression she gave me when I was her student at the University of Reykjavík. Since then she has grown into a national leader in legal academia and one of very few Icelandic legal scholars with an international reputation. In addition to commenting on my work she has facilitated crucial logistical assistance by securing with me a generous 3-year project grant from the Icelandic Research Fund (grant no. 141274-053) which made the research possible from a financial point of view. She also helped me integrate at the University of Iceland during the last part of the PhD process and has given me many good advices about the academic life in general.

The research also benefited from the generosity of judge Páll Hreinsson, who invited me to stay in the Icelandic cabinet at the EFTA Court in Luxembourg in early 2016. This stay was very helpful towards getting a practical perspective on the research. Coincidentally, I attended a reception with AG Sharpston at the EFTA court, but of course I was too shy to approach her and tell her how my PhD thesis was inspired by her KME opinion. I also owe gratitude to the law department of the University of Iceland. During the last year of the research the department provided me with an office space in Lögberg and the academic and the administrative staff showed me great collegiality during the last stretch of the work.

My fellow EUI researchers and friends Alastair Maciver, Juha Tuovinen, and Zane Rasnaca also contributed to my work by providing social and moral support through all the coffees and lunches at the Badia, and the occasional nights out at the infamous Bar Fiasco, or the slightly better respected Brewdog.

The true hero of my PhD saga is my beloved wife Áslaug. Not only did she put up with my academic endeavour by following me to Italy; she also accepted to marry me in the Palazzo

Vecchio (2014) and had our two children, Aría (2013) and Kári (2015), while stoically waiting for me to get this over with.

I would like to dedicate this work to my dear friend Kolbrún Ólafsdóttir (1971-2009) who left us too soon while pursuing her dream of studying at the LSE. In some of our numerous intercontinental phone calls, during my Stockholm years (2007-2009) and her Beijing and London years, I first discussed the prospect of doing a PhD one day. Well, here it is.

Reykjavík, 9 May 2017

The problem with fairness in EU's competition procedure

Legal practitioners often overlook the importance of the logical link between the laws and society's reigning political philosophy. When entangled in the process of solving legal problems as private practitioners, or as official decision makers of public institutions, the laws may seem a closed system that abides exclusively to its own internal logic, unaffected by external social and political realities. Nevertheless, the laws undeniably tend to mirror the human societies they serve. Novel technologies prompt legislative changes, major societal events are dealt with by amending the laws, and gradual cultural progressions only become permanently cemented through alterations in society's legal fabric. Laws do however not only mirror what society is; the laws also reflect what society wants to be. The laws can thus depart from the actualities of reality both due to a failure to adapt to what is, and due to societies failure to adapt to what ought to be, according to the desired reality prescribed by the laws.

Human society is by its very nature subject to constant change, fuelled by the fluidity of social interactions and the randomness of their consequences. The black letter laws are comparatively static; reminiscent of a snapshot at a given point in time that mirrors what society once was, or a reflection of what society once wanted to be. The tension between the dynamic nature of society and the static character of the laws creates multiple dilemmas and paradoxes for the practitioners of law. This tension can be ignored by considering the domain of laws as being a closed system of logical arguments that is self-sufficient in providing solutions to any question of law. By ignoring the temporal tension, the laws will however fail to reach their full potential in their mission to guide and reflect the society that they are meant to serve. Acknowledgment of the temporal tension reveals weaknesses in the traditional juridical method and suggest an opportunity for methodological improvements.

Among the most difficult problems of legal adjudication is to reconcile the two-pronged purpose of the laws of providing a sturdy guidance of the optimal behaviour, and at the same time to adjust flexibly to the contemporary mood. A lax attitude towards the former risks complacency with socially suboptimal behaviour, while rigidity towards the later risks imposition of behavioural standards no longer deemed socially desirable. This problem can

take many forms in legal practice and has been written about extensively in legal theory, often implicitly in the context of the doctrines of legal positivism (i.e. the law is a closed system), legal naturalism (i.e. the law is a part of a larger system of ethics), and legal realism (i.e. the law is what it is).

The problem of fairness in the laws is a variant of the temporal problem. Fairness can both be a substantive standard of optimality that society seeks to achieve through its laws, and a standard whose substantive meaning is sensitive towards the current social temperament. It should be safe to presume that the laws ought to be fair. Determining objectively whether a specific act of law, or a specific legal action constitutes fairness, poses the hard question.

The subject of this research is to explore the general problem of fairness in the laws, from the special point of view of European Union (EU) competition law procedure. The approach to the question of fairness in the laws is initially posed as a philosophical question, but will gradually gravitate towards a practical application in the context of EU competition procedure by using insights from the theory of economics.

The special case of EU competition law procedure provides an excellent case study of how the problem of fairness in the laws has been approached in practice by practitioners of EU law, by the EU institutions, and by the EU law focused academic literature. A recent discussion about fairness in the context of EU competition procedure centred on the *KME-Chalkor* litigation,¹ which raised doubts about the compatibility of EU's institutional approach to competition law enforcement with human rights obligations to ensure procedural fairness. The academic community identified a potential problem prior to the litigation, scrutinised it closely while it was ongoing, and has reflected upon it afterwards. Taken together this has become a sizable literature.² The views in the literature mirror the opposing claims of the parties to the case

¹ See three judgements delivered on 8 December 2011: Case C-272/09 P *KME Germany and Others v Commission* [2011] ECLI:EU:C:2011:810; C-386/10 P *Chalkor v Commission* [2011] ECLI:EU:C:2011:815; and C-389/10 P *KME Germany and Others v Commission* [2011] ECLI:EU:C:2011:816.

² See some of the main contributions; Denis Waelbroeck and Denis Fosselard, 'Should the decision-making power in EC antitrust procedures be left to an independent judge? – The impact of the European Convention of Human Rights on EC antitrust procedures' (1994) 14 *Yearbook of European Law* 111; Frank Montage, 'The case for a radical reform of the infringement procedure under regulation 17' (1996) 17 *European Competition Law Review* 428; Koen Lenaerts and Jan Vanhamme, 'Procedural rights of private parties in the community administrative

before the Court of Justice of the European Union (CJEU; the Court). Parallel to the *KME-Chalkor* litigation, similar issues were tried before the other EEA court on the Kirchberg plateau in Luxemburg and before the European Court of Human Rights (the ECHR) in Strasburg. The EFTA Court's judgement in the *Posten Norge*³ was delivered five months after the CJEU's judgment, but the ECHR's judgment in *Menarini*⁴ was delivered two months prior to the CJEU's judgement in *KME-Chalkor*.

process (1997) 34 Common Market Law Review 531; Wouter P. J. Wils, 'The combination of the Investigative and prosecutorial function and the adjudicative function in EC antitrust enforcement: a legal and economic analysis' (2004) 27 World Competition 201; Ólafur Jóhannes Einarsson, 'EC competition law and the right to a fair trial' (2006) 25 Yearbook of European Law 555; Eric Barbier de La Serre, 'Procedural justice in the European Community case-law concerning the rights of the defence: essentialist and instrumental trends (2006) 12 European Public Law 225; Donald Slater, Sébastien Thomas and Denis Waelbroeck, 'Competition law proceedings before the European Commission and the right to a fair trial: no need for reform?' (2008) The Global Competition Law Centre Working Paper Series - Working Paper No. 04/08. Also published as; Donald Slater, Sébastien Thomas and Denis Waelbroeck, 'Competition law proceedings before the European Commission and the right to a fair trial: no need for reform?' (2009) 5 European Competition Journal 97; Ian S. Forrester, 'Due process in EC competition cases: a distinguished institution with flawed procedures' (2009) 34 European Law Review 817; Fernando Castillo de la Torre, 'Evidence, proof and judicial review in cartel cases' (2009) 32 World Competition 505; Wouter P. J. Wils, 'The increased level of EU antitrust fines, judicial review and the ECHR' (2010) 33 World Competition 5; Jaime Flattery, 'Balancing efficiency and justice in EU competition law: elements of procedural fairness and their impact on the right to a fair hearing' (2010) 7 The Competition Law Review 53; Wouter P. J. Wils, 'EU anti-trust enforcement powers and procedural rights and guarantees: the interplay between EU law, national law, the Charter of Fundamental Rights of the EU and the European Convention on Human Rights' (2011) 34 World Competition 189; Marc Jaeger, 'The standard of review in competition cases involving complex economic assessments: towards the marginalisation of the marginal review?' (2011) 2 Journal of European Competition law & Practice 295; Peter Oliver, "'Diagnostics" – a judgment applying the Convention of Human Rights to the field of competition' (2012) 3 Journal of European Competition Law & Practice 163; Anne-Lise Sibony, 'Annotation of the judgment of the Court in Case C-272/09 P KME Germany and others v. Commission' (2012) 49 Common Market Law Review 1977; Marco Bronckers and Anne Vallery, 'Fair and effective competition policy in the EU: which role for authorities and which role for the courts after Menarini?' (2012) 8 European Competition Journal 283; Renato Nazzini, 'Administrative enforcement, judicial review and fundamental rights in EU competition law: a comparative contextual-functionalist perspective' (2012) 49 Common Market Law Review 971; Heike Schweitzer, 'Judicial review in EU competition law' in Damien Geradin and Ioannis Lianos (eds), *Handbook on European competition law* (Edward Elgar Publishing 2013) 491; Ingrid Vandenborre and Thorsten Goetz, 'EU competition law procedures' (2012) 3 Journal of European Competition Law & Practice 578; Nils F.W. Hauger and Christoph Palzer, 'Investigator, prosecutor, judge ... and now plaintiff? The Leviathanian role of the European Commission in the light of fundamental rights' (2013) 36 World Competition 565; Eric Barbier de la Serre, 'Standard of review in competition law cases: Posten Norge and beyond' in Carl Baudenbacher, Philipp Speitler and Bryndís Pálmarsdóttir (eds), *The EEA and the EFTA Court: decentred integration: to mark the 20th anniversary of the EFTA Court* (Hart Publishing 2014); and Wouter P. J. Wils, 'The compatibility with fundamental rights of the EU antitrust enforcement system in which the European Commission acts both as investigator and as a first-instance decision maker' (2014) 37 World Competition 5.

³ See judgement of the EFTA Court of 18 April 2012 in Case E-15/10 *Posten Norge v EFTA Surveillance Authority* [2012] EFTA Ct Rep 246.

⁴ See judgement of the ECHR of 27 September 2011 in Case No 43509/08 *A. Menarini Diagnostics S.r.l. v. Italy* [2011] ECLI:CE:ECHR:2011:0927JUD004350908.

The context provided by the Court's decisional practice on the fairness of EU's competition procedure and the accompanying body of academic literature can be posed as a foreground to the more fundamental philosophical questions of fairness in the laws, and more specifically the question of fairness of procedural law, which lurk in the background.

As a starting point, let us quickly sketch up the main elements of the foreground, before engaging with the more fundamental philosophical fairness problem in the background, which forms the primary object of this research.

1. The elements

Leading up to the *KME-Chalkor* litigation, parallel developments had occurred in two unrelated fields of EU law that, when taken together, gave rise to doubts about the compatibility of EU's competition procedure with obligations to ensure procedural fairness. More specifically, these doubts concerned the procedure for imposing fines in cartel cases and the way in which the Court exercised its power to review such decisions.

The former development occurred following the modernisation of the competition enforcement regime with the entering into force of Regulation 1/2003⁵ and is best explained using data published by the Directorate General (DG) Competition:

⁵ Council Regulation (EC) 1/2003 on the implementation of the rules on competition laid down in Articles 81 and 82 of the Treaty [2002] OJ L001.

Cartel cases where fines were imposed by the Commission 1990 – 2016

(amounts adjusted for Court judgments)⁶

<u>Period</u>	<u>Fine amounts</u>	<u>Number of cases</u>	<u>Average fine per case</u>
1990 - 1994	344.282.550 €	10	34.428.255 €
1995 - 1999	270.963.500 €	10	27.096.350 €
2000 - 2004	3.157.348.710 €	30	105.244.957 €
2005 – 2009	7.920.497.227 €	33	240.015.068 €
2010 – 2014	7.608.375.579 €	30	253.612.519 €
(2015 – 2016)	4.091.507.000 €	11	371.955.182 €
Total	23.392.974.566 €	124	188.653.021 €

In the 1990s, the average cartel fine approximately amounted to € 30 million. By the early-mid 2000s, the fines started to increase substantially, reaching an average of € 285 million in the period 2010-2016.⁷ This increase in the level of fines coincided with the modernisation program of the late 1990s and early 2000s, and was supported by a seemingly conscious use by DG Competition of reproachful rhetoric to increase the social stigma of engaging in a conduct prohibited by EU’s competition provisions.⁸ An example of this can be seen in a speech given by Mario Monti, then Commissioner for Competition, in Stockholm in September 2000:⁹

⁶ Source: DG Comp’s website, updated on 12 December 2016 <<http://ec.europa.eu/competition/cartels-statistics/statistics.pdf>> accessed 30 December 2016. This table merges tables 1.4 and 1.10 in DG Comp’s document and adds an average column. The amounts concern fines imposed for infringements against Article 101 TFEU and its predecessors, i.e. ex Article 81 EC and ex Article 85 EC. In this table, a cartel case concerns a single proceeding against various undertakings concerned, and may involve more than one infringement. Only those cartel cases where a fine was imposed were considered for the purpose of this table.

⁷ Note though that the bare numbers say nothing about the seriousness of the underlying breach or the size of the firms involved, which will influence the size of the fine and thus influence how comparable each instance is with another. Commission officials have nonetheless conceded that there has been a noticeable increase in the size of cartel fines in recent decades. See Wouter P. J. Wils, ‘The compatibility with fundamental rights of the EU antitrust enforcement system in which the European Commission acts both as investigator and as a first-instance decision maker’ (2014) 37 *World Competition* 5, 8; Wouter P. J. Wils, ‘The increased level of EU antitrust fines, judicial review and the ECHR’ (2010) 33 *World Competition* 5, 10-12. See also Fernando Castillo de la Torre, ‘Evidence, proof and judicial review in cartel cases’ (2009) 32 *World Competition* 505, 506. Wils however maintains that the increase is irrelevant for the purposes of assessing compliance with Article 6(1) ECHR, since the Engel criteria assesses the maximum permissible fine, the level of which has remained at 10% of annual corporate turnover since the implementation of Regulation 17 in the 1960s.

⁸ For examples of this rhetoric see quotes listed in; Donald Slater, Sébastien Thomas and Denis Waelbroeck, ‘Competition law proceedings before the European Commission and the right to a fair trial: no need for reform?’ (2008) *The Global Competition Law Centre Working Paper Series - Working Paper No. 04/08*, 14-15

⁹ See Mario Monti, ‘Fighting Cartels Why and How? Why should we be concerned with cartels and collusive behaviour?’ (2000) 3rd Nordic Competition Policy Conference, Stockholm <http://europa.eu/rapid/press-release_SPEECH-00-295_en.htm> accessed 9 September 2016.

'Cartels are cancers on the open market economy, which forms the very basis of our Community. By destroying competition they cause serious harm to our economies and consumers. (...) In the words of Adam Smith there is a "tendency for competitors to conspire". This tendency is of course driven by the increased profits that follow from colluding rather than competing. We can only reverse this tendency through tough enforcement that creates effective deterrence. The risk of being uncovered and punished must be higher than the probability of earning extra profits from successful collusion.'

Describing the conduct of breaching Article 101 TFEU as the infliction of a cancer on the economy, and accusing the entities engaged in such activity of conspiracy against the general public, was probably not what the original signatories of the EEC Treaty of Rome envisioned, when they delegated the supposedly minor administrative issue of enforcing the competition provisions of the Treaty to the Commission in the early 1960s,¹⁰ but is consistent with a current trend in competition law enforcement in many other jurisdictions.¹¹

The later development that raised worries about EU's competition procedure's compliance with fairness standards concerned the constitutional project of the EU. In line with ordoliberal ideas popular at the time, the original EEC Treaty of Rome was viewed by some as a sort of an economic constitution, referring to the imperativeness of the free movement principles branded as the four freedoms and the ancillary provisions on competition. Later, the Court started gradually to recognise traditional rights principles as being part of the constitutional framework of the EU system of laws, despite their absence from the Treaty of Rome. This was in part due to a necessity following the establishment of the supremacy doctrine in *Costa v*

¹⁰ See further discussion on the intentions of the Treaty signatories with regards to the competition provisions in; David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 *Harvard International Law Journal* 97, 103-07; and more generally on the history of EU's competition law by the same author in; David Gerber, *Law and Competition in the Twentieth-Century Europe: Protecting Prometheus* (Oxford University Press 2001).

¹¹ On the existence of a worldwide trend of criminalising cartel offences see Peter Whelan, 'Cartel criminalization and the challenge of "moral wrongfulness"' (2013) 33 *Oxford Journal of Legal Studies* 535, 536.

Enel,¹² which created the potential for EU laws to override constitutional rights in the Member States; including traditional rights provisions of the moral kind.

In *Internationale Handelsgesellschaft* of 1972, the Court claimed that ‘*respect for fundamental rights forms an integral part of the general principles of law protected by the Court of Justice. The protection of such rights, whilst inspired by the constitutional traditions common to the Member States, must be ensured within the framework of the structure and objectives of the Community.*’¹³ From then on, it has been assumed that traditional moral rights form part of the constitutional framework of the EU, and since the judgment in *Rutili*¹⁴ in 1975, the rights under the European Convention on Human Rights (ECHR) have been explicitly referred to in the Court’s case law.¹⁵

Although implicitly recognised as being part of the primary law of the EU from the 1970s through the Court’s case law, an effort was not made to codify an explicit list of rights for the EU until in 1999. Then the European Council decided to commission the drafting of a Charter of Fundamental Rights (the Charter) to a senior body that adopted the name ‘the European Convention’. The Commission, the European Parliament, and the Council of Ministers proclaimed the draft as the Charter of Fundamental Rights in the Nice summit of 2000, but decided that ‘*the question of the Charter’s force [should] be considered later.*’¹⁶

An updated version of the Charter was supposed to become part of the European Constitution of 2004,¹⁷ which failed in the ratification process. A reference was also made to an updated version of the Charter in the Lisbon Treaty of 2007, with the intention of giving it a comparable hierarchal legal status as the founding Treaties. When the Lisbon Treaty entered into force on 1

¹² See Case 6/64 *Costa v ENEL* [1964] ECLI:EU:C:1964:66.

¹³ See Case 11/70 *Internationale Handelsgesellschaft mbH v Einfuhr- und Vorratsstelle für Getreide und Futtermittel* [1972] ECLI:EU:C:1970:114.

¹⁴ See Case 36/75 *Roland Rutili v Ministre de l’intérieur* [1975] ECLI:EU:C:1975:137. The ECHR is explicitly mentioned in paragraph 32 of the judgment.

¹⁵ See further on the early development of fundamental rights protection within the EU in; Ólafur Jóhannes Einarsson, ‘EC competition law and the right to a fair trial’ (2006) 25 *Yearbook of European Law* 555, 556-59.

¹⁶ See European Council – Nice 7-10 December 2000: Conclusions of the presidency <http://www.europarl.europa.eu/summits/nice1_en.htm> accessed 12 September 2016.

¹⁷ See Treaty establishing a Constitution for Europe [2004] OJ C310/1.

December 2009, the Charter of Fundamental Rights¹⁸ thus acquired a Treaty status within the EU system of laws, which consequently for the first time made rights of the moral kind an explicit part of the codified constitutional framework of the European Union.

For the purposes of the *KME-Chalkor* litigation, this gradual constitutional development with regards to moral rights gave rise to an argument that the institutional arrangement of competition enforcement, which was instituted by the Treaty of Rome in 1957 and by Regulation 17¹⁹ in 1962 and thus predated this development, was no longer compatible with the recognition of the right to a fair procedure enshrined in the newly codified Article 47 of the Charter of Fundamental Rights.

Independently, these two separated but parallel developments, of criminalisation of competition law breaches and the constitutionalising of moral rights, could each warrant a reconsideration of whether the institutional architecture of competition law enforcement was compatible with the current norms of procedural fairness. When combined, these distinct developments formed a powerful argument that required close attention by the stakeholders of the competition law enforcement regime. Arguably, the threshold of rights protection had risen over time: first following the implicit recognition of rights in the EU system of law; and later through an explicit codification. At the same time, the protective interests had also increased: on one hand as the result of increased social stigma against competition law breaches; and on the other hand, as the result of increased economic consequences for those caught committing such breaches. If we think about the standard of criminality and the standard of a moral entitlement to a fair process as two separate constants against which the factual context of a case is assessed, the argument of the *KME-Chalkor* litigation was that the substance of the rights constant had changed, and that the factual circumstances with regards to the standard of criminality had changed.

The narrative about the gradual criminalisation of competition law breaches fits uncomfortably with the traditional categorisation of the adjudicative processes for dealing with potential

¹⁸ See Charter of Fundamental Rights of the European Union [2010] OJ C 83/391.

¹⁹ See Council Regulation No 17 (EEC): First Regulation Implementing Articles 85 and 86 of the Treaty [1962] OJ No. 013.

breaches against public law, which normally are categorised as either criminal or administrative. Usually matters of lesser importance are dealt with based on the less expensive and the more efficient mechanism of the administrative procedure. Matters of greater importance for the state or the individuals involved, are hence dealt with through the more expensive and cumbersome criminal procedure. A narrative that suggest a gradual change within a system of binary categorisation needs to explain precisely when and how an object ceases to be one thing and becomes the other thing, i.e. in this case criminal or administrative type of a proceeding.

In the abstract, the punishment for a breach against legally protected public interests is usually rationalised through one of two modes of arguing: either through a deontological argument, in which a criminal procedure would be warranted if the moral stigma of being found guilty of a breach is great (i.e. retributive rationale); or through a consequentialist argument, in which a criminal procedure would be warranted if the consequences at stake would be great for any of the stakeholders (i.e. deterrence rationale).²⁰

The level of the seriousness of the punishment for being found in breach thus usually correlates with the level of the stigma, or with the level of the consequences of the breach. Importantly, this typical correlation is not a necessary relation. If a punishment is designed based on deterrence theory, it is possible that the most efficient level of deterrence is achieved through a low-level punishment, irrespective of the importance of the interests at stake for the relevant stakeholders.²¹ Leaving such extraordinary circumstance aside, it should in most cases be safe to assume that breaches against public interests that are subject to severe punishments should ideally be dealt with through a criminal procedure, rather than an administrative procedure.

²⁰ See generally on the theory of punishment in criminal law in; H.L.A Hart, 'The Presidential Address: Prolegomenon to the Principles of Punishment' (1959) 60 Proceedings of the Aristotelian Society 1.

²¹ Usually a severe punishment would warrant a criminal procedure due to the potential consequences for the accused, but it is also possible that a morally reprehensible crime does not need a hefty punishment for achieving deterrence effect, in which case the argument for a criminal procedure would need to rest on a deontological rationale, rather than on a consequentialist rationale. Further on this point see Peter Whelan, 'Cartel criminalization and the challenge of "moral wrongfulness"' (2013) 33 Oxford Journal of Legal Studies 535, 540-43.

The two main elements underpinning the *KME-Chalkor* cartel litigation, which have been explained above, suggest that a cartel is a breach against legally protected public interests. Breach against these public interests is both considered morally reprehensible (e.g. DG Comp's rhetoric) and is subject to a sizable pecuniary penalty. Irrespective of whether the enforcement rationale is based on the theory of retribution or deterrence, it seems that in the abstract these kinds of breaches should fall within the criminal sphere of public law enforcement, rather than the administrative sphere. The exception noted above with regards to a deterrence rationale of a punishment seems inapplicable; the fine is high rather than low, which in any case will create the potential for grave consequences that require a cautious approach to adjudication.

The legal landscape in *KME-Chalkor* was however complicated by the fact that the competition procedure at stake had initially been designed in the 1960s as an administrative procedure. Even after the Commission had sharpened its rhetoric against competition law breaches and started to multiply the fines in the early-mid 2000s, the legislator resisted when regulation 17 was modernised by explicitly maintaining the non-criminal definition of the procedure for imposing fines in competition cases. This can be seen in Article 23(5) in the successor Regulation 1/2003,²² which restates the repelled Article 15(4) of Regulation 17/62: '*Decisions taken pursuant to paragraphs 1 and 2 shall not be of a criminal law nature.*'

Things were further complicated by developments in the interpretation of the European Court of Human Rights (the ECoHR) of how a distinction should be made between a criminal procedure and an administrative procedure. For the purposes of the ECHR the distinction mattered; Article 6 only set the standard of procedural fairness for criminal and civil procedures, not for administrative procedures. This had gradually lead the ECoHR to expand the concept of a criminal procedure to expand the applicability of the procedural guarantees dictated by the ECHR. This policy had however backfired when citizens started to demand their day in court to argue over parking fines and other minor issues based on the ECHR. The ECoHR attempted to strike a balance in this sense with the *Jussila* judgement,²³ where it abandoned the binary approach to the categorisation of criminal and administrative procedures in favour

²² See Council Regulation (EC) No 1/2003 on the implementation of the rules on competition laid down in Articles 81 and 82 of the Treaty [2003] OJ L 1/1.

²³ See Judgment in Case no. 73053/01 *Jussila v Finland* [2006] ECLI:CE:ECHR:2006:1123JUD007305301.

of a gradual approach in which procedures could be somewhat criminal in nature, but not so that they necessitated the full set of procedural guarantees provided by Article 6 of the ECHR:

‘There are clearly “criminal charges” of differing weight. What is more, the autonomous interpretation adopted by the Convention institutions of the notion of a “criminal charge” by applying the Engel criteria have underpinned a gradual broadening of the criminal head to cases not strictly belonging to the traditional categories of the criminal law, (...). Tax surcharges differ from the hard core of criminal law; consequently, the criminal-head guarantees will not necessarily apply with their full stringency’²⁴

The third element in the problem in the *KME-Chalkor* litigation was thus that the actual procedure for imposing fines in competition cases had been designed to comply with the more relaxed procedural fairness standard of the administrative procedure, and the legislator had restated his intention to refrain from making the standard of fairness more rigid in line with what applies to criminal procedures explicitly in the modernised procedural regulation from 2003. Additionally, it was not entirely clear how the ECoHR would categorise EU’s enforcement procedure in the field of competition law, and thus whether, or to what extent the procedural guarantees stipulated in Article 6(1) of the ECHR and the accompanying case law would apply.

2. The problem

In the *KME-Chalkor* litigation, the Court was confronted with a problem consisting of elements that pointed towards contradictory solutions. On one hand, competition law breaches carried social stigma and were punished by hefty fines, which indicated a need for a criminal procedure to ensure procedural fairness for the parties involved. On the other hand, the actual existing procedure had initially been designed as an administrative procedure and the legislator had recently restated his intention to keep it within the administrative sphere for the purposes of procedural fairness guarantees. The case law of the ECoHR did at that point not provide a definite answer about how the dilemma about the applicable standard of procedural fairness should be resolved.

²⁴ See Case no. 73053/01 *Jussila v Finland* [2006] ECLI:CE:ECHR:2006:1123JUD007305301, para 43.

Referring to the temporal problem described earlier, we can see that there had been changes in how the Commission wanted competition law breaches to be perceived by the public over the decades since the enactment of Regulation 17. The gradual recognition, and later the codification of fundamental rights into the EU legal system, can be viewed as a response to changes in the public perception and expectations about moral rights in the laws, and thus how rigidly the standard of procedural fairness should be interpreted in the context of competition law enforcement.

The social stigma, the hefty fines, and the gradual rise and explicit recognition of the right to a fair procedure, did by the time of the *KME-Chalkor* litigation suggest that breaches to the competition law regime should be perceived as of a criminal nature and should accordingly be dealt with based on the criminal procedural modality. At the same time, however, the written law of the procedural regulation explicitly ordered that the criminal procedural modality should not be used for imposing fines for breaches against the competition provisions of the TFEU, suggesting that such breaches were not serious enough to warrant the procedural guarantees reserved for those accused of the most serious breaches against public order.

The actual problem, which the Court was confronted with in the *KME-Chalkor* cases, was of course much more specific than is indicated by the description of the broad elements above. The three elements formed a legal-factual texture, which dictated the range of possibilities to answer the more specific challenge made against the Court's doctrine regarding the standard of review of the Commission's competition decisions. The specific issue of the *KME-Chalkor* cases, of concern for the purposes of this research, was the question whether the CJEU's doctrine of providing the Commission with deference to assess complex economic facts in competition cases, and thus limit its judicial review of the Commission's decisions, was compatible with the right to a fair trial enshrined in Article 47 of the Charter of Fundamental Rights. These more specific issues will be analysed later; for the purposes of this introductory chapter it suffices to summarise the problem and the Court's solution, and to review the academic literature on the cases.²⁵

²⁵ The *KME-Chalkor* cases are analysed in more details in chapter 8.

3. The academic debate and the Court's solution

Doubts about the compatibility of the EU's competition law regime with principles of procedural fairness can be traced back to the 1970s and the 1980s,²⁶ but for the purposes of the *KME-Chalkor* litigation the first important contribution to the academic literature was Waelbroeck's and Fosselard's article of 1994 in the *Yearbook of European Law*.²⁷ The arguments raised in this article set the tone for the ensuing debate about procedural fairness in EU's competition law regime, and its compatibility with the procedural safeguards articulated in Article 6 of the ECHR, and later Article 47 of the Charter of Fundamental Rights. Importantly, the argument was raised in the article about the incompatibility of the CJEU's doctrine of granting the Commission with deference in assessing complex technical and economic facts, with the requirement of the ECHR that a tribunal with full jurisdiction should review criminal charges.²⁸

Another important contribution from the initial phase of the debate, was an article by Montage, where he articulated three problems from the point of view of the accused with the enforcement regime under Regulation 17 and suggested that the best remedy would be that the Court of First Instance (now the General Court) should take the initial prohibition decision in competition cases. Accordingly, he argued that DG Comp's role in competition proceedings should end with the issuance of a statement of objections (SO).²⁹

In an article from 1997 in the *Common Market Law Review*, a judge and a legal secretary at the Court of First Instance jointly responded to the charge that the alleged deference doctrine was incompatible with the requirements of Article 6 of the ECHR. They conceded that the Court had repeatedly stated in its judgments that there was a certain discretion enjoyed by the Commission, but they argued that the Court did nonetheless exercise a full jurisdiction within

²⁶ See further in Ian S. Forrester, 'Due process in EC competition cases: a distinguished institution with flawed procedures' (2009) 34 *European Law Review* 817, 819-20.

²⁷ See Denis Waelbroeck and Denis Fosselard, 'Should the decision-making power in EC antitrust procedures be left to an independent judge? – The impact of the European Convention of Human Rights on EC antitrust procedures' (1994) 14 *Yearbook of European Law* 111.

²⁸ See Denis Waelbroeck and Denis Fosselard, 'Should the decision-making power in EC antitrust procedures be left to an independent judge? – The impact of the European Convention of Human Rights on EC antitrust procedures' (1994) 14 *Yearbook of European Law* 111, 125-33.

²⁹ See Frank Montage, 'The case for a radical reform of the infringement procedure under regulation 17' (1996) 17 *European Competition Law Review* 428.

the meaning of the ECHR on the account of what the Court effectively did in such cases. The system could thus be viewed as compatible with the ECHR due to a paradox in how the Court expresses himself and in how it acts.³⁰ In an article from 2006, Einarsson reaches a similar conclusion based on a detailed review of the ECoHR's, the General Court's, and the CJEU's case law. Einarsson argues that if the EU courts did in the past not review sufficiently the Commission's administrative process in competition cases, the current case law (at the time, i.e. in 2006) indicated that the level of scrutiny was in line with the requirements of Article 6 of the ECHR as interpreted by the ECoHR.³¹

By the time Einarsson's article was published in 2006, the academic debate had more or less settled on the view that: (i) the Commission's competition law enforcement procedure was criminal in nature within the autonomous meaning of Article 6 ECHR; (ii) the EU courts would need to exercise full judicial review jurisdiction over the result of the administrative procedure before the Commission in order to comply with the requirements of the ECHR as interpreted in the case law of the ECoHR; and (iii) the actual review conducted by the EU courts was no longer subject to the deference doctrine originating in *Consten and Grundig*³² of 1966.

A succession of events in late 2006 and in 2007 gave rise to a reconsideration of whether the consensus still held. On 23 November 2006, the Grand Chamber of the ECoHR handed down the judgment in *Jussila v Finland*, where it reopened the question of which procedural guarantees should be applicable to procedures that did not belong to the traditional categories of criminal law, but which did fall under the criminal heading of Article 6 ECHR based on the *Engel criteria*. The judgment did not concern a competition procedure, but the logic of the finding indicated that it could be applicable to EU's competition procedure as well. On 17 September 2007, the Court of First Instance handed down the *Microsoft* judgment. Interestingly, for the purposes of applicability with Article 6 ECHR, the deference doctrine for

³⁰ See Koen Lenaerts and Jan Vanhamme, 'Procedural rights of private parties in the community administrative process (1997) 34 Common Market Law Review 531, 560-62. The co-author of the article, Judge Lenaerts, is from 2015 the President of the Court of Justice of the EU.

³¹ See Ólafur Jóhannes Einarsson, 'EC competition law and the right to a fair trial' (2006) 25 Yearbook of European Law 555, 612-14.

³² See Joined Cases 56 and 58/64 *Consten and Grundig v Commission* [1966] ECLI:EU:C:1966:41, p 347.

complex economic and technical matters was restated subject to certain conditions.³³ Finally, on 13 December 2007 the Member States signed the Lisbon Treaty and thus signalling the eventual 1 December 2009 entering into force of Article 47 of the Charter of Fundamental Rights, which codified procedural rights equivalent to those of Article 6 ECHR into the constitutional structure of the EU.

In a much-cited article from 2008, Slater, Thomas, and Waelbroeck,³⁴ restated many of the arguments initially raised by Waelbroeck and Fosselard in 1994. They argued that the seriousness of competition law breaches warranted an initial decision taken by an independent tribunal within the meaning of Article 6 ECHR, and that in any case the judicial review conducted by the Court of First Instance was not a full jurisdiction review within the meaning of the ECHR. Forrester, Microsoft's counsel in the *Microsoft case* and now a judge at the General Court, also criticised the Commission's competition procedure in a 2009 article in the *European Law Review*, arguing that the power over the result of a public prosecution should not be in the hands of the politically appointed Commissioners of the Commission due to a risk of political bias.³⁵

The Commission's officials were quick to respond to these renewed speculations and doubts about the fairness of the competition procedure. In 2009, Torre wrote an article³⁶ in defence of the system in place, and in 2010 Wils did the same.³⁷ The main argument of Wils was made with reference to the *Jussila* judgment of the ECoHR: the competition procedure should not be considered to be of a hard-core criminal nature, and thus the procedural guarantees of

³³ See Case T-201/04 *Microsoft v Commission* [2007] ECLI:EU:T:2007:289, paras 87-89. The actual formulation is somewhat strange; the deference doctrine is articulated in paras 87 and 88, but then reduced and conditioned in para 89 with reference to para 39 of Case C-12/03 P *Commission v Tetra Laval* [2005] ECLI:EU:C:2005:87.

³⁴ Donald Slater, Sébastien Thomas and Denis Waelbroeck, 'Competition law proceedings before the European Commission and the right to a fair trial: no need for reform?' (2008) The Global Competition Law Centre Working Paper Series - Working Paper No. 04/08 <<https://www.coleurope.eu/content/gclc/documents/GCLC%20WP%-2004-08.pdf>> accessed 3 January 2017. Also published as; Donald Slater, Sébastien Thomas and Denis Waelbroeck, 'Competition law proceedings before the European Commission and the right to a fair trial: no need for reform?' (2009) 5 *European Competition Journal* 97.

³⁵ See Ian S. Forrester, 'Due process in EC competition cases: a distinguished institution with flawed procedures' [2009] 34 *European Law Review* 817, 831-33. Forrester was also Chalkor's counsel in the *KME-Chalkor* litigation.

³⁶ Fernando Castillo de la Torre, 'Evidence, proof and judicial review in cartel cases' (2009) 32 *World Competition* 505.

³⁷ Wouter P. J. Wils, 'The increased level of EU antitrust fines, judicial review and the ECHR' (2010) 33 *World Competition* 5.

Article 6(1) ECHR should not apply with their full stringency. This, Wils argued, permits the first decision in competition cases to be reached through an administrative procedure. Addressing the issue of deference in light of the *Microsoft case*, Wils resorted to the same argument as Lenaerts and Vanhamme mustered in their *Common Market Law Review* article of 1997; the CFI in *Microsoft*, first grants deference to the Commission in paragraph 87 and 88, then takes most of it back in paragraph 89, and finally in practice executes a full judicial review leaving no deference to the Commission.³⁸ Jaeger, the President of the General Court, argued along similar lines in his 2011 article in the *Journal of European Competition Law & Practice*.³⁹

The entering into force of the Lisbon Treaty by late 2009 had the practical implication that it became easier to plead the points raised in the academic debate through a direct litigation before the EU courts with reference to Article 47 of the Charter of Fundamental Rights, which supposedly provided at least equivalent procedural guarantees as where available under Article 6 of the ECHR. Already by 2010 the *KME-Chalkor* appeal litigation⁴⁰ was underway with the promise of settling the longstanding disagreement about the compatibility of the infringement procedure that resulted in a fine in competition cases, with the procedural fairness standards resulting from Article 6 ECHR and the accompanying case law of the ECoHR, and now also the corresponding standard in Article 47 of the Charter of Fundamental Rights.

Shortly before the CJEU could reach its decision in the *KME-Chalkor* cases, the ECoHR had the opportunity to clarify its position on the applicability of Article 6(1) of the ECHR with regards to competition law proceedings in the judgment in the *Menarini* case which was delivered on 27 September 2011.⁴¹ In *Menarini* the ECoHR found the Italian competition law procedure, which like the EU's procedure was an administrative procedure at the first instance, to be criminal in nature according to the *Engel criteria*. The ECoHR nonetheless found the procedure compatible with Article 6(1) of the ECHR, since Italian courts had full jurisdiction to review the

³⁸ Wouter P. J. Wils, 'The increased level of EU antitrust fines, judicial review and the ECHR' (2010) 33 *World Competition* 5, 18 and 26-28.

³⁹ Marc Jaeger, 'The standard of review in competition cases involving complex economic assessments: towards the marginalisation of the marginal review?' (2011) 2 *Journal of European Competition Law & Practice* 295, 313.

⁴⁰ The General Court handed down its judgment in the cartel cases against Chalkor and KME (Cases T-21/05 and T-25/05) on 19 May 2010, and the appeals were filed on 30 July and 3 August 2010.

⁴¹ See Case No 43509/08 A. *Menarini Diagnostics S.r.l. v. Italy* [2011] ECLI:CE:ECHR:2011:0927JUD004350908.

administrative decision.⁴² In *Menarini* the ECoHR thus found competition proceedings to be criminal in nature, but that they should, even so, not be subject to the strict interpretation of the procedural conditions of Article 6(1) ECHR. A competition decision imposing a penalty of a criminal nature could thus be taken through an administrative procedure, if an independent and impartial court or a tribunal had full jurisdiction to review the finding. This meant that the *Jussila* distinction between different degrees of criminality was extended to competition proceedings, in the sense that they did not belong to the traditional hard-core criminal types of proceedings, but were somewhat criminal nevertheless.

The *KME-Chalkor* litigation ended with the CJEU's judgments on 8 December 2011. The Court avoided the question of criminal or non-criminal nature of the proceedings, which had been one of the main points of the preceding academic debate and one of the central points discussed during the court proceedings. The Court also chose to approach the problem exclusively from the point of view of Article 47 of the Charter; it did not mention Article 6(1) of the ECHR.⁴³ The Court concluded that the General Court had full jurisdiction to review the Commission's fining decisions in cartel cases both in terms of facts and law, and thus the procedure was compatible with the standard of procedural fairness articulated in Article 47 the Charter of Fundamental Rights. To be able to do so, the Court quashed the deference doctrine that had been reinstated by the *Microsoft* judgment, and argued along the lines of Leanarts and Vanhamme in their 1997 article in the *Common Market Law Review* and Wils in his more recent article of 2010 in *World Competition*. The Court argued that irrespective of the deference doctrine the General Court says it is using; in practice, it is conducting a full judicial review.⁴⁴

⁴² See Case No 43509/08 A. *Menarini Diagnostics S.r.l. v. Italy* [2011] ECLI:CE:ECHR:2011:0927JUD004350908, paras 59 and 63-64. See also a short case note on *Menarini* in; Peter Oliver, "'Diagnostics' – a judgment applying the Convention of Human Rights to the field of competition" (2012) 3 *Journal of European Competition Law & Practice* 163.

⁴³ Sibony points out that the Court could avoid answering the question about the criminal nature of the proceeding by only referring to Article 47 of the Charter, since the classification only has relevance in terms of Article 6 of the ECHR; Article 6 ECHR only applies to civil and criminal proceedings, while Article 47 of the Charter is not conditioned on any specific type of a proceeding. See Anne-Lise Sibony, 'Annotation of the judgment of the Court in Case C-272/09 P *KME Germany and others v. Commission*' (2012) 49 *Common Market Law Review* 1977, 1989-95.

⁴⁴ The *KME-Chalkor* cases are reviewed in detail in Chapter 8.

On April 18 2012, few months after the result in the *KME-Chalkor* litigation, the EFTA Court delivered a judgment in *Posten Norge*, which dealt with similar topics as had been the central issues in *Menarini* and *KME-Chalkor*.⁴⁵ In its judgment, the EFTA Court approached the problem from the point of view of Article 6 ECHR, and was quick to conclude with reference to *Menarini* and AG Sharpston's opinion in *KME*⁴⁶ that the said procedure before the EFTA Surveillance Authority was criminal in terms of Article 6 ECHR. The EFTA Court however noted, referring to *Jussila*, that the extent of the procedural guarantees provided by Article 6 ECHR 'must be determined with regard to the weight of the criminal charge at issue'.⁴⁷ The procedural guarantees were thus not to be considered as absolute, but rather as a balancing exercise in which the nature of the crime at stake was a large factor. On an inference of this rationale, the Court concluded that the competition procedure at stake could be compatible with Article 6(1) ECHR, even if the initial criminal penalty decision was taken at an administrative stage, if the administrative decision could later be appealed to an impartial and an independent tribunal with full jurisdiction.⁴⁸

The EFTA Court also discussed the deference doctrine with regards to complex economic and technical facts. The EFTA Court argued that its review of economic and technical facts was limited to a legality review, but that this limitation did not restrict the EFTA Court in conducting a full jurisdiction review in terms of Article 6 ECHR. Interestingly the EFTA Court explained that the limitation inherent in a legality review bared the EFTA Court from substituting its own assessment for ESA's assessment because it considered its own assessment of the facts to be more correct; the EFTA Court could only annul ESA's assessment if the inference drawn by ESA could not be substantiated by the relevant facts and was thus in breach of the principle of legality.⁴⁹

⁴⁵ See Case E-15/10 *Posten Norge v EFTA Surveillance Authority* [2012] EFTA Ct Rep 246.

⁴⁶ AG Sharpston was more explicit in her opinion than the CJEU in concluding that EU's competition procedure should be considered criminal in terms of Article 6 ECHR. See Case C-272/09 P *KME Germany and Others v Commission* [2011] ECLI:EU:C:2011:63, opinion of AG Sharpston.

⁴⁷ See Case E-15/10 *Posten Norge v EFTA Surveillance Authority* [2012] EFTA Ct Rep 246, para 87-88.

⁴⁸ See Case E-15/10 *Posten Norge v EFTA Surveillance Authority* [2012] EFTA Ct Rep 246, para 91.

⁴⁹ See Case E-15/10 *Posten Norge v EFTA Surveillance Authority* [2012] EFTA Ct Rep 246, paras 98-101.

While the judgments in the *KME-Chalkor* were a bit cryptic in terms of providing answers to the main points of argument in the foregoing academic debate about the procedural fairness of EU's competition procedure, the *Menarini* judgment and the *Posten Norge* judgment brought more clarity. The two later judgments are explicit in that the *Jussila* doctrine of differing degrees of criminality applies to the field of competition law procedure and that consequently the procedural guarantees of Article 6(1) ECHR do not apply with their full stringency in that field of law due to the lacking degree of criminality. The *KME-Chalkor* judgements follow the same rationale, but do so implicitly and without reference to Article 6(1) ECHR and the *Jussila* doctrine, and in fact without much reference to any proper rationalisation.⁵⁰

Several authors have taken stock of the legal situation after this trio of judgments.⁵¹ Sibony notes the absence of a discussion about the nature of competition proceedings in terms of procedural typology and the choice of the Court not to mention the Article 6 of the ECHR in the *KME-Chalkor* judgments. She suggests several potential explanations that concern inter- and intra-institutional politics. Ultimately, she considers the result a disappointment; the guardian of the Treaty did '*not expose a theory of guardianship*'.⁵² Barbier de la Serre considers the EFTA Court in *Posten Norge* going beyond what the CJEU did in *KME-Chalkor* in terms of

⁵⁰ Sibony notes this as a missed opportunity for the CJEU. See Anne-Lise Sibony, 'Annotation of the judgment of the Court in Case C-272/09 P *KME Germany and others v. Commission*' (2012) 49 *Common Market Law Review* 1977, 2002.

⁵¹ See for example; Peter Oliver, "'Diagnostics" – a judgment applying the Convention of Human Rights to the field of competition' (2012) 3 *Journal of European Competition Law & Practice* 163; Anne-Lise Sibony, 'Annotation of the judgment of the Court in Case C-272/09 P *KME Germany and others v. Commission*' (2012) 49 *Common Market Law Review* 1977; Marco Bronckers and Anne Vallery, 'Fair and effective competition policy in the EU: which role for authorities and which role for the courts after *Menarini*?' (2012) 8 *European Competition Journal* 283; Renato Nazzini, 'Administrative enforcement, judicial review and fundamental rights in EU competition law: a comparative contextual-functionalist perspective' (2012) 49 *Common Market Law Review* 971; Heike Schweitzer, 'Judicial review in EU competition law' in Damien Geradin and Ioannis Lianos (eds), *Handbook on European competition law* (Edward Elgar Publishing, 2013) 491; Ingrid Vandenborre and Thorsten Goetz, 'EU competition law procedures' (2012) 3 *Journal of European Competition Law & Practice* 578; Nils F.W. Hauger and Christoph Palzer, 'Investigator, prosecutor, judge ... and now plaintiff? The Leviathanian role of the European Commission in the light of fundamental rights' (2013) 36 *World Competition* 565; Eric Barbier de la Serre, 'Standard of review in competition law cases: *Posten Norge* and beyond' in Carl Baudenbacher, Philipp Speitler and Bryndís Pálmarsdóttir (eds), *The EEA and the EFTA Court: decentred integration: to mark the 20th anniversary of the EFTA Court* (Hart Publishing 2014); and Wouter P. J. Wils, 'The compatibility with fundamental rights of the EU antitrust enforcement system in which the European Commission acts both as investigator and as a first-instance decision maker' (2014) 37 *World Competition* 5.

⁵² Anne-Lise Sibony, 'Annotation of the judgment of the Court in Case C-272/09 P *KME Germany and others v. Commission*' (2012) 49 *Common Market Law Review* 1977, 2002.

claiming jurisdiction to control the assessment of facts. He also notes a problem associated with a mismatch between what the CJEU says it does, and what it effectively does in terms of judicial review; if the Court's actions are inconsistent with its rhetoric, there is no way for an outsider to determine when the judicial review is in line with the required standards of procedural fairness.⁵³

Bronckers and Vallery point out a paradox in the CJEU's argument in *KME*; the General Court had in its judgment under appeal repeatedly referred to a discretion it granted the Commission with, but the CJEU puzzlingly stated that this was no hindrance for the purposes of full judicial review; '*[i]n our view, it is unfortunate that the appearance of a less than full review by the General Court was not criticised by the Court of Justice. Justice must not only be done, but also be seen to be done.*'⁵⁴ Bronckers and Vallery note that the EFTA Court with its *Posten Norge* judgment explicitly embraced the lessons of *Jussila* and *Menarini*, but that the CJEU still has to confront the concept of judicial review more directly. They suggest that '*perhaps the proper solution is to distinguish Commission decisions imposing fines, which fall within the criminal sphere protected by Article 6 ECHR, from other decisions.*' By this, the Commission's discretion could be eliminated for a particular type of decisions, without having to eliminate it across the board; '*[f]ining decisions must be deemed criminal cases from an ECHR perspective; other competition law decisions more resemble classic administrative law.*'⁵⁵ Bronckers and Vallery warn that pushing the fairness claim of moving the initial decision power to a court could imply inefficiencies in some Member States; '*[i]n other words, for a transitional period an approximation of fairness may have to be accepted, with improvements being made to the administrative process, in order not to jeopardise the effectiveness of competition law.*'⁵⁶

⁵³ Eric Barbier de la Serre, 'Standard of review in competition law cases: *Posten Norge* and beyond' in Carl Baudenbacher, Philipp Speitler and Bryndís Pálmarsdóttir (eds), *The EEA and the EFTA Court: decentred integration: to mark the 20th anniversary of the EFTA Court* (Hart Publishing 2014) 418 and 427-28.

⁵⁴ Marco Bronckers and Anne Vallery, 'Fair and effective competition policy in the EU: which role for authorities and which role for the courts after *Menarini*?' (2012) 8 *European Competition Journal* 283, 292.

⁵⁵ Marco Bronckers and Anne Vallery, 'Fair and effective competition policy in the EU: which role for authorities and which role for the courts after *Menarini*?' (2012) 8 *European Competition Journal* 283, 294-96.

⁵⁶ Marco Bronckers and Anne Vallery, 'Fair and effective competition policy in the EU: which role for authorities and which role for the courts after *Menarini*?' (2012) 8 *European Competition Journal* 283, 297.

Wils considers the debate about the compatibility of the EU competition enforcement system as such, in which the Commission both investigates and decides, with Article 6 of the ECHR, to be over following *Jussila and Menarini*; *[t]he only question which is still open to debate is whether the General Court, when reviewing European Commission decisions, exercises sufficient jurisdiction to meet the ‘full jurisdiction’ standard laid down in the case law of the European Court of Human Rights. Either the General Court does exercise ‘full jurisdiction’, and there is no problem, or it does not, in which case the solution is for the court to modify its practice.*⁵⁷ Addressing the issue of the mismatch between the rhetoric on deference and the actual actions of the General Court, Wils notes that *‘[w]hat is decisive is whether the General Court in fact exercises full jurisdiction, not any general statements which the Courts may make as to its powers. It is nevertheless also important that the General Court is seen to exercise full jurisdiction. For this reason, potentially misleading general statements should be avoided.’*⁵⁸

Nazzini considers the CJEU falling short of adopting a full correctness standard of judicial review in *KME-Chalkor*, implying that some elements of deference remain. He also considers the argument for saving the current system advocated in the literature by the Commission’s officials, such as Wils, which apparently was used by the CJEU in *KME-Chalkor*, to be *‘far from satisfactory.’*⁵⁹ Nazzini suggests that *KME-Chalkor* signals the demise of the deference standard of review in competition cases in favour of the correctness standard. He does however offer an alternative to retain some of the elements of the deference doctrine; *‘The alternative is to preserve, when appropriate, a deferential standard of review but with the necessary procedural counterbalance of functional separation between the prosecutor and the decision-maker within the Commission.’*⁶⁰ The minimal fix, suggested by Nazzini, would require the establishment of a separate adjudicative unit outside of DG Competition, which would report directly to the

⁵⁷ Wouter P. J. Wils, ‘The compatibility with fundamental rights of the EU antitrust enforcement system in which the European Commission acts both as investigator and as a first-instance decision maker’ (2014) 37 World Competition 5, 11.

⁵⁸ Wouter P. J. Wils, ‘The compatibility with fundamental rights of the EU antitrust enforcement system in which the European Commission acts both as investigator and as a first-instance decision maker’ (2014) 37 World Competition 5, 24.

⁵⁹ Renato Nazzini, ‘Administrative enforcement, judicial review and fundamental rights in EU competition law: a comparative contextual-functionalist perspective’ (2012) 49 Common Market Law Review 971, 995-97. The argument he is referring to concerns the contention that the General Court supposedly conducts a full review irrespective of what he explicitly states in his judgments with regards to deference granted to the Commission.

⁶⁰ Renato Nazzini, ‘Administrative enforcement, judicial review and fundamental rights in EU competition law: a comparative contextual-functionalist perspective’ (2012) 49 Common Market Law Review 971, 999.

Commissioner responsible for the competition portfolio. This would thus be a functional separation, rather than a structural separation of prosecutorial and adjudicative powers.⁶¹

To summarise the current state of play in the *KME-Chalkor* saga, it is now generally accepted that institutionally it is permissible in terms of Article 6 ECHR and 47 of the Charter to organise competition law enforcement based on a two-tier adjudicative system, in which the former is an administrative level. This is however only permissible if the adjudicative function at the later tier is conducted by an impartial and an independent tribunal with a full jurisdiction to review the results of the administrative tier. The court practice of the EFTA Court, the CJEU, and the ECoHR is towards reducing or eliminating any deference granted to the administrative authorities, but the academic debate has not yet reached a consensus on how to optimise procedural fairness in competition proceedings. Some want to retain some elements of the deference doctrine, while others want to sever the adjudicative function from the administrative function through changes in the institutional structure. The consensus about the current legislative landscape thus only extends to *what currently is*; a consensus about what ideally *ought to be* in terms of procedural fairness in competition proceedings is still nowhere in sight.

Notably, the current consensus of the courts on the permissibility of deciding a competition procedure at the first instance through an administrative procedure, rests on the assumption that the criminal nature of the acts that are the objects of such procedures, exist in different degrees of severity and that the associated procedural guarantees granted to defendants should take note of the nature of the act on a scale of criminal gravity. The implication of this apparent consensus of the three European courts is that the binary system of categorising procedures within the sphere of public law as either administrative or criminal is effectively obsolete for accurate descriptive purposes. The revelation that procedures can be somewhat criminal in nature, but not entirely, and that procedural design should somehow reflect these varying degrees of criminality, signals a rupture from a tradition of categorical absolutes in procedural design and a succession of a procedural architectural regime based on a notion of

⁶¹ Renato Nazzini, 'Administrative enforcement, judicial review and fundamental rights in EU competition law: a comparative contextual-functionalist perspective' (2012) 49 *Common Market Law Review* 971, 1002-04.

balancing. A methodological approach to this balancing is still underdeveloped in the academic literature.

The absence of categorical procedural imperatives for determining the object of procedural fairness, and the lack of a comprehensive balancing methodology for achieving procedural fairness through procedural design, ambiguates the notion of procedural fairness. The legal concept of procedural fairness is thus in need of a sound philosophical rationale and a comprehensive methodology of execution to escape the perils of ambiguity.

4. The question and a path to an answer

The foreground to the main topic of this research, that I have now outlined based on the *KME-Chalkor* litigation, shows how a specific problem with procedural fairness in EU competition procedure has been dealt with in the literature and in the decisional practice of the CJEU and its sister courts in Luxemburg and Strasburg. Over time, a consensus has emerged that the concept of procedural fairness, as articulated in the European Convention of Human Rights and the Charter of Fundamental Rights, is not an absolute term but a term subject to some type of balancing. What remains underdeveloped, both in the academic literature, and in the decisional practice of the courts, is to explain further the underlying normative elements and the methodology of this balancing.

The main research question of this research thus focuses on the larger philosophical and methodological issues in the background of the *KME-Chalkor* litigation; *how do we know when a procedure is fair?* The context of EU's competition law procedure and the concept of procedural fairness as dictated by Article 47 of the Charter is, due to the decisional practice and the literature summarised above, a good starting point and a good point of reference in developing the concept of procedural fairness philosophically and in developing a methodology for testing the level of procedural fairness in legislation.

Thinking along the lines of the traditional juridical method, the methodology for answering the research question would be similar to the one used by the courts and the academic literature summarised above; locate and define the problem in relation to the relevant sources of law,

and then use these sources to extract the law's stance, which dictates the solution to the problem at hand. As we have seen, the answer to the question of procedural fairness in the context of EU competition law according to this methodology depends on how the CJEU interprets Article 47 of the Charter.

As discussed above, this solution to the problem of procedural fairness is not entirely satisfactory due to the temporal problem. How do we know whether the CJEU's interpretation of the concept of procedural fairness in relation to Article 47 of the Charter is synchronised with what preferably ought to be considered fair in the present now? Especially considering that the concept of procedural fairness enshrined in the Charter is supposed to be analogous to the concept of the ECHR which was designed almost 70 years ago; before the Hart-Dworkin debate, before law and economics, and before John Rawls theory of justice. One answer would submit that Article 47 of the Charter and the Court's interpretations dictate the actual substance of procedural fairness. Another answer would suggest that Article 47 of the Charter and the Courts interpretations are an expression of a more fundamental concept of procedural fairness that exists independently and beyond the laws and thus derives its substantive content from other sources than the laws, as understood in the traditional sense.

The path to establishing the basic elements of procedural fairness in the context of EU competition law procedure thus starts at a familiar junction; do the laws dictate what is fair, or does fairness find expression through the laws? If the former is true, we need not look any further than at the constitutional texts and the Court's interpretation of them to establish what constitutes procedural fairness in each context. If, however, the latter were true, an autonomous conception of procedural fairness would be needed to determine if a given procedural rule is compatible with the concept of procedural fairness.

The initial claim of this thesis, on which the subsequent successive research problems are based, is articulated in chapter two. There I take the view that fairness is an autonomous concept that conditions the content of laws in a democratic society. In arriving to that conclusion, I explain how the mainstream theories of law (i.e. legal positivism and legal naturalism) fail to explain the moral implication of the political choice of a democratic form of

government. Positivism inherently assumes the irrelevance of morality, while naturalism inherently assumes the imperativeness of morality, both irrespective of the chosen form of government. Under my approach, morality's link to the laws is contingent upon the *a priori* choice of the form of government. A democratic choice necessarily implies laws obedience to a moral principle of egalitarian fairness.

Having made the initial claim, I turn next to finding a methodology for implementing the claimed concept of fairness. Chapter three reviews the leading methodologies of balancing in legislative design and introduces concepts from decision theory, a branch of micro economics theory that deals specifically with optimising decision making. Chapter four introduces a novel methodology for implementing the concept of fairness into legislative decision making; aimed at improving shortcomings of the orthodox methods, based on fundamental concepts from decision theory, and focused on the economics based concept of expected utility.

Having claimed a concept of fairness and designed a methodology for implementing it, chapter five starts preparing practical application by analysing the general normative and functional properties of law enforcement procedures, both of which are essential for extending the claimed concept of fairness to a concept of procedural fairness. This analysis reveals law procedure's normative object of optimisation and the functional process through which a procedure seeks optimisation. The extraction of these properties makes the domain of law enforcement procedures compatible for an analysis based on the previously established concept of fairness and for practical implementation based on the optimisation methodology described in chapter four.

Chapters six and seven extend the analysis of a law procedure in the abstract, to the specific domain of EU's competition procedure by identifying: firstly, the normative objective of EU's competition law; and secondly, the functional design of EU's competition procedure. The analysis confirms in the actual procedural design of EU's current competition procedure, the basic structure of procedural fairness identified in the abstract for procedures in general. This allows for the application of the initial fairness concept, on legislative design dilemmas within the specific domain of EU's competition procedure. Chapter eight, finally describes two recent

procedural design dilemmas in EU's competition procedure and shows how the dilemmas could be resolved based on the claimed concept of fairness and through the described optimisation methodology.

The thesis thus aims at: (i) claiming a philosophical conception of fairness in the laws, (ii) building a methodology for implementing that concept in legislative design, (iii) and showing how both the concept and the methodology can be applied to procedural design dilemmas in EU's competition procedure. If successful, the concept and the methodology should provide a philosophical foundation for answering *why a law procedure is fair*, and a methodological platform for establishing *when a law procedure is fair*.

The dark matter of law

In the 1980s, several independent studies in particle physics suggested the existence of an unknown matter that decisively affected how the observable universe functioned. From observed disturbances in the gravitational fields of large galactic structures, it seemed as if something was out there in the vast emptiness of the dark space.⁶² This unknown phenomenon became known as dark matter and is, in the standard model of particle physics (*the Lambda-CDM model*), assumed to account for 84.5% of the total mass of all matter in the universe. This phenomenon has not yet been seen or empirically detected, but it is presumed to exist due to gravitational effects affecting the visible matter of the universe.⁶³

A mainstream view of legal philosophy holds that various types of codes of conduct, that collectively can be referred to as human made positive law, can determine human behaviour. On this view, it is up to humans to decide how they behave, and for our purposes, to decide the standard or the object of fairness in the laws. A competing view holds that although the laws are human inventions or artefacts, their substance is subject to restraints deriving from a phenomenon that is integrally part of the human condition, but at the same time beyond the reach of positive human intervention. We can call this phenomenon the dark matter of law; i.e. a legal substance that is not part of the codified positive law, but that yet intangibly surrounds and influences any provision of the tangible law, similarly to the force of the dark matter on the visible matter of the galaxies.

Viewed in broader philosophical terms, this essentially is a debate about the extent of philosophical determinism and to what degree humans have free will. In legal philosophy, variants of this debate have been discussed under various headings for many centuries in the western civilisation. In earlier times, natural law and divine law were considered as constants that restrained the extent to which humans ought to exercise their free will. Accordingly,

⁶² See for example; George R. Blumenthal, S.M. Faber, Joel R. Primack, and Martin J. Rees, 'Formation of galaxies and large-scale structure with cold dark matter' (1984) *Nature* 517.

⁶³ For a short explanation of the concept of dark matter and the Standard Model in laymen terms see short notes on CERN's website; <<http://home.cern/about/physics/dark-matter>> <<http://home.cern/about/physics/standard-model>> accessed 19 October 2016.

human made laws were to be anchored in the laws of the nature, or in the divine will. The modern version of the debate stretches back to the birth of modern legal positivism in the monographs of Hans Kelsen from 1934 and H.L.A Hart from 1961, where they argue for a theory of positive law that is unbound by external ethical constraints.⁶⁴

To this date, the mid-20th century legal positivism of Kelsen and Hart, and their followers, must be considered an orthodoxy of legal philosophy and the benchmark against which alternative theories of law can assume to be weighed against. If we are thus to advance a theory of law that assumes the inclusion of a matter that exists beyond the positive law, i.e. the dark matter of law, the starting point needs to be the orthodoxy which the heterodoxy seeks to improve or replace.

1. Modern legal positivism and its greatest critic

Hart's legal positivism made an important contribution to the debate about supposed gaps in the law. Before Hart, the debate was polarised between the natural law contention that laws were a complete system with no gaps,⁶⁵ and the legal realist approach that claimed that rules only exist when they are applied, at which point their practical meaning was largely influenced by extra-legal factors.⁶⁶ Between these approaches, Hart argued for the *open texture of law*. His theory capitalised on the insight of the realists, that law cannot be predetermined for every possible future instance, but argued that laws do nonetheless cover a definite core of instances. Hart considered that at the margins of the core there are sometimes legally undetermined cases that judges have a discretion to decide. On Hart's view, they do so by making a new rule due to the absence of an applicable rule.⁶⁷

⁶⁴ See Hans Kelsen, *Pure theory of Law* (first published 1934, University of California Press 1967); and H.L.A. Hart, *The concept of law* (first published 1961, 2nd edn, Clarendon Press 1994).

⁶⁵ This approach is by Hart and some commentators labelled 'formalism' (see for example; Scott J. Shapiro, *Legality* (Belknap Press, 2011) 259-61) but I will refer to it as natural law approach since that is essentially what the approach is.

⁶⁶ Hart referred to legal realism as '*rule-scepticism*' which is usefully descriptive. The term legal realism is however the normal terminology. Brian Leiter describes what he calls '*the core claim of realism*' in the following way: '*[I]n deciding cases, judges respond primarily to the stimulus of the facts of the case, rather than to legal rules and reasons.*' See Brian Leiter, 'American Legal Realism' (2002) The University of Texas School of Law Public Law and Legal Theory Research Paper No. 042, 6-7 <<https://ssrn.com/abstract=339562>> accessed 3 January 2017.

⁶⁷ See H.L.A. Hart, *The concept of law* (first published 1961, 2nd edn, Clarendon Press 1994) 124-36.

Against this aspect of Hart's legal theory, Ronald Dworkin launched a powerful attack.⁶⁸ The basis for his objection was the intuition, and arguably empirical fact, that judges never engage legal reasoning as if they are making a new rule. Their findings are always formulated on the assumption they are applying existing legal norms. He further made a distinction between rules and principles, which Hart had neglected.⁶⁹ On Dworkin's account, rules are binary concepts that either apply or not; for example, either you break the speed limit or you do not. Principles however, are dynamic concepts that can conflict with each other and must thus be weighted in terms of importance by judges in hard cases. When such conflict occurs, Dworkin argued, judges do not make a new rule; they resort to moral reasoning to discover which principle should prevail. On Dworkin's account, the legal system is complete. There are always legal principles to discover that are morally appropriate. The single right answer to a question of law is thus the solution with the highest aggregated moral appropriateness.

Dworkin's claim was that if positivism only considers social facts that frequently run out as sources of law, then positivism cannot be true because in the reality, judges do not run out of legal sources. This claim has two premises that can be attacked. First, that social facts frequently run out, and secondly that judges never run out of legal sources. This simple claim was energetic enough to divide legal positivism into two factions, based on which of Dworkin's premise they considered false.⁷⁰

An objection to the former premise builds on a move to include moral arguments as the sources of laws and thus reduce the instances when the law run out greatly. This is made possible by accepting moral considerations as part of the laws on the condition that the *rule of recognition* validates them. Since the *rule of recognition* must be a social rule, the law is

⁶⁸ The original objection was published as an essay; Ronald Dworkin, 'The Model of Rules' (1967) 35 The University of Chicago Law Review 14. See also reprinted as; Ronald Dworkin, 'The Model of Rules I' in *Taking Rights Seriously* (Harvard University Press, 1977) 14.

⁶⁹ Hart later acknowledged that he overlooked this distinction, but he nonetheless considered that his theory could accommodate this distinction. See H.L.A. Hart, 'The Postscript' in *The concept of law* (first published 1961, 2nd edn, Clarendon Press 1994) 259-63.

⁷⁰ See further Shapiro's analysis in; Scott J. Shapiro, *Legality* (Belknap Press 2011) 267-73 regarding Dworkin's central claim in *The Model of Rules* and on the different responses to the claim.

ultimately decided by social facts alone. This approach is sometimes labelled as *inclusive positivism* because of the inclusion of moral considerations.⁷¹

An objection to the latter of Dworkin's premise relies on a different move. This objection concedes that judges run frequently out of legal sources. However, when that occurs, judges make new rules based on extra-legal norms that ultimately decide the case at hand. These norms could be moral norms, but since the judge is no longer engaged in application of legal sources, the separation of laws and morals holds. This approach is usually labelled as *exclusive positivism*, signalling the exclusion of moral considerations from the domain of laws.⁷²

Dworkin was not convinced by these responses.⁷³ The inclusive approach, he claimed, was a '*version of legal positivism*' that '*is best described as anti-positivism*'.⁷⁴ By this, he meant that these kinds of claims effectively conceded that morality is part of the laws but somehow still claimed that very broadly defined social conventions could be seen as social facts on which the laws were exclusively dependent. By doing this, Dworkin believed, positivism was saved by abandoning everything that is positivistic about it. Dworkin did not hold the exclusive approach in high esteem either. He considered Raz's account to be reliant on counter intuitive premises about the authority of laws. To explain this problem Dworkin branded Raz's approach as '*Ptolemaic dogma*', referring to the obsolete geocentric theories of the Greek-Egyptian astrologist Claudius Ptolemy (≈ 100 CE - 170 CE).⁷⁵

⁷¹ This approach is often associated with Jules Coleman and his followers. Hart endorsed this approach in *The Postscript*. For general discussion see Kenneth Einar Himma, 'Inclusive Legal Positivism' in Jules L. Coleman and Scott Shapiro (eds), *The Oxford Handbook of Jurisprudence and Philosophy of Law* (Oxford University Press 2004) 125.

⁷² Joseph Raz is the key theorist associated with this approach. For general discussion see Andrei Marmor, 'Exclusive Legal Positivism' in Jules L. Coleman and Scott Shapiro (eds), *The Oxford Handbook of Jurisprudence and Philosophy of Law* (Oxford University Press 2004) 104.

⁷³ Dworkin responded comprehensively to the attempts by positivist theorist to save legal positivism from his critique of Hart in a book review of Jules Coleman's book. He aimed mainly at the specific approaches of Raz and Coleman, but the arguments can be seen as having general application against these two forms of legal positivism. See Ronald Dworkin, 'Thirty years on: Book Review of The Practice of Principle: In Defence of a Pragmatist Approach to Legal Theory by Jules Coleman' (2002) 115 *Harvard Law Review* 1655.

⁷⁴ See Ronald Dworkin, 'Thirty years on: Book Review of The Practice of Principle: In Defence of a Pragmatist Approach to Legal Theory by Jules Coleman' (2002) 115 *Harvard Law Review* 1655, 1665.

⁷⁵ See Ronald Dworkin, 'Thirty years on: Book Review of The Practice of Principle: In Defence of a Pragmatist Approach to Legal Theory by Jules Coleman' (2002) 115 *Harvard Law Review* 1655, 1655.

Dworkin's simple objection split modern legal positivism into two incompatible positions.⁷⁶ This was however not a knockout blow, although it raised serious doubts about the plausibility of the positivist project. We must thus take a closer look at legal positivism's defining thesis about the reasonableness of separating laws and morals.

2. The problem with the separation thesis

Hart's account of legal positivism was initially an attempt to reform legal positivism as described by Jeremy Bentham, and in particular by John Austin in *The Province of Jurisprudence Determined* published in 1832. This type of legal positivism was grounded on an assumption that the laws were essentially the commands of the sovereign that could be implemented involuntarily by force. The command and the threat of force was thus an integral part of the concept of law. Hart argued that the *command theory* could not explain why the orders of the armed bandit should not be considered instances of legally valid rules.⁷⁷

In response to the insufficiency of the *command theory* of Bentham and Austin, Hart suggested *the rule of recognition*, which became the essential innovative feature of his version of legal positivism. The rule is laid out in the following way:

*[W]hat is crucial [about the Rule of Recognition] is the acknowledgement of reference to the writing or inscription as authoritative, i.e. as the proper way of disposing of doubts as to the existence of the rule. Where there is such an acknowledgement there is a very simple form of secondary rule: a rule for conclusive identification of the primary rules of obligation.*⁷⁸

⁷⁶ This split still remains, as is apparent by a comment recently made by Brian Leiter, a self-proclaimed exclusivist, about a fellow positivist Scott Shapiro: 'Shapiro self-identifies as a positivist, of course, and in his earlier work [...] offered a new argument for "hard" or "exclusive" legal positivism [...]. But his more recent work makes so many confused criticisms of Hart's positivism and concedes so much to the anti-positivist views as to make it unclear whether the resulting theory really honors the Sources and Conventionality Theses.' See Brian Leiter, 'Legal Positivism About the Artifact Law: A Retrospective Assessment' in L. Burazin, K. E. Himma, C. Roversi (eds) *Law as an Artifact* (Oxford University Press, forthcoming 2017) note no 9. <<https://ssrn.com/abstract=2870877>> accessed 16 January 2017.

⁷⁷ See chapters 2-4 in; H.L.A. Hart, *The concept of law* (first published 1961, 2nd edn, Clarendon Press 1994). See also H.L.A. Hart, 'Positivism and the Separation of Law and Morals' (1958) 71 *Harvard Law Review* 593, 603; 'Law surely is not the gunmen situation writ large, and legal order is surely not to be thus simply identified with compulsion.'

⁷⁸ See H.L.A. Hart, *The concept of law* (first published 1961, 2nd edn, Clarendon Press 1994) 95.

On this view, the laws create authoritative obligations. The laws do not simply state descriptively what *is*, they state normatively what *ought* to be done according to the law.⁷⁹ Importantly, the *conclusive* authority for the substance of a primary rule is a secondary rule that dictates which elements can form a basis for a judgement of legality under the primary rule. The secondary rule is thus the *rule of recognition*.

The principal purpose of the *rule of recognition* was to provide an authority for rules that made a distinction between the order of the gunman and the order of the legitimate official. This Hart managed to do without resorting to moral authority, by creating a master rule of other rules that legitimised orders of a predefined sort. If the order of the gunman were not sanctioned by the master rule as a source of laws, it could not become an instance of a valid law.

To emphasise the core element of legal positivism it is helpful to visualise the puzzle of legality as a question that needs to be answered based on certain facts.⁸⁰ The facts that are necessary for the answer can be called legal facts, which distinguish them from other facts that are not relevant. The task is then to determine the composition of the legal facts to enable extraction of the answer to the question of legality. In its basics, positivist theory irreducibly maintains that all legal facts are determined by social facts alone.⁸¹ What is meant by social facts remains flexible, but in a broad sense it refers to an entity that can be observed by the five human senses and subsequently described. A social fact thus observably *is* and refers to *a posteriori* knowledge.

Consequentially, the question of legality on the positivistic view must be answered with reference to a knowledge that is observably present, while in contrast, the naturalistic approach would additionally consider moral facts that exist irrespective of what can be observed in the present. Moral facts refer to *a priori* knowledge that is intrinsically true and

⁷⁹ Hart emphasises the normative aspect of social rules as compared with social habits. The element of *ought*, *should* and *must*, crucially distinguishes between optional habitual behaviour and socially required behaviour. See H.L.A. Hart, *The concept of law* (first published 1961, 2nd edn, Clarendon Press 1994) 9-10, 55-58.

⁸⁰ See Scott J. Shapiro, *Legality* (Belknap Press 2011) 25-27; and Mark Greenberg, 'How Facts Make Law' (2004) 10 *Legal Theory* 157.

⁸¹ See Scott J. Shapiro, *Legality* (Belknap Press 2011) 27.

thus independently authoritative. Moral facts command how things *ought to be*, but do not describe how things presently *are*. The answer to the legality question in the naturalistic view is thus derived from a combination of how things *are*, and how they morally *ought to be*.

The question of legality is a question of *a priori* knowledge, that is, what ought to validly direct an action. It is not a question of *a posteriori* knowledge. The point of reference is not what results of an action, but instead what causes an action. Hart's concept of law, by ascribing to normative concepts like authority and obligation, is in this sense a description of causal effects. This, according to Shapiro, potentially puts positivism at odds with Hume's fork.⁸²

Hume formulated a maxim that dictates that an *ought* cannot be derived from an *is*.⁸³ It means that for the output to be necessarily true, the input has also to *be a priori* true knowledge. If, however the input is dependent on what is observable, the truth of output is contingent on the limits of the observation. To put this in context; the *rule of recognition* is an observable social rule that produces rules about how things ought to be. The output is thus formulated as universally true; it is based on the *a priori* knowledge that the *rule of recognition* is authoritative and that there is an obligation to abide to it without any reservation. The problem is however that the *rule of recognition* itself is formulated as descriptive and observable and is thus contingent upon what can be observed and described. The question then becomes; is it possible to create something universally true, out of premises that are contingent upon what can be positively observed? Hume would deem that problematic. The output of contingent premises must be contingent as well.

⁸² See Scott J. Shapiro, *Legality* (Belknap Press 2011) 47.

⁸³ See David Hume, *Treatise of Human Nature* (first published 1738-40, David Fate Norton and Mary J. Norton eds, Oxford University Press 2007) Book III, Section I, last paragraph: 'I can't forbear adding an observation that may be found of some importance. In every system of morality I have met with I have noticed that the author proceeds for some time reasoning in the ordinary way to establish the existence of a God, or making points about human affairs, and then he suddenly surprises me by moving from propositions with the usual copula 'is' (or 'is not') to ones that are connected by 'ought' (or 'ought not'). This seems like a very small change [Hume writes 'This change is imperceptible', but he can't mean that literally], but it is highly important. For as this 'ought' (or 'ought not') expresses some new relation or affirmation, it needs to be pointed out and explained; and a reason should be given for how this new relation can be—inconceivably!—a deduction from others that are entirely different from it. Authors don't ordinarily take the trouble to do this, so I recommend it to you; and I'm convinced that paying attention to this one small matter will subvert all the vulgar systems of morality and let us see that the distinction between vice and virtue is not based merely on the relations of objects, and is not perceived by reason.' – This is taken from Jonathan Bennett's reformatted and annotated version for easier reading see: <<http://www.early-moderntexts.com/authors/hume>> accessed 1 November 2016.

Hart explained that his master rule, that validates other rules, exists based on a social practice that is concordantly followed by the officials of the system. Its existence is thus a matter of a social fact.⁸⁴ By this move, Hart attempted to defuse the concept of the *rule of recognition* of its normative charge as the ultimate source of legality. The absolute source of the rule that dictates legality is thus contingent on the social fact that the officials of the system follow it. Shapiro argues that this spin creates a new problem. If the *rule of recognition* is plainly an expression of a social practice, the question becomes about which social practices form the *rule of recognition* and which do not? It cannot be that everything that constitutes a social practice is an element of the law's master rule. Another *rule of recognition* would be needed to recognise the specific social practices that form the foundations of the first level master rule. It would be possible to repeat the regression infinitely.⁸⁵

It does thus not seem possible to give an entirely descriptive account of a normative claim without a reference to a normative authority in the way Hume suggested. One will always perceive through the senses more information than is relevant for describing a normative claim. To decide which information is relevant a further normative judgement is required. What ought to be taken as an instance or element of a given normative claim, out of the total sum of instances that are perceived? This cannot be answered without a reference to something that ultimately is the normative reason for the normative claim.⁸⁶

Hart's trick to replace the *command model* with the *rule of recognition* seemed ingenious for a while. Upon closer scrutiny, the move has been shown to be problematic. The *rule of recognition* temporarily created the illusion that a normative claim about the laws could be

⁸⁴ See H.L.A. Hart, *The concept of law* (first published 1961, 2nd edn, Clarendon Press 1994) 110: '[T]he rule of recognition exists only as a complex, but normally concordant, practice of the courts, officials, and private persons in identifying the law by reference to certain criteria. Its existence is a matter of fact.' See also pages 254-59 in the *Postscript* where Hart discusses Dworkin's criticism of the *practice theory*. On page 259 Hart describes the core claim in the following way: '[T]he law of a system is identified by criteria provided by a rule of recognition accepted in the practice of the courts (...)'.
⁸⁵ See Scott J. Shapiro, *Legality* (Belknap Press 2011) Chapter 4, especially 95-96.
⁸⁶ Burazin recently suggested that laws and legal systems could be perceived as institutional artefacts that had an author, with an intention, and which had been collectively accepted. The normativity of the legal system, understood in this way, would come through the second condition about the author's intention. See Luka Burazin, 'Can there be an artefact theory of law?' (2016) 29 *Ratio Juris* 385, 397-99.

based on a descriptive premise. A closer look at the premise however shows that to maintain the illusion, one would need to repeat the move indefinitely. Attempts of legal positivists, after Hart, bear the blunt mark of this hopeless regression. The arguments have become ever more abstract, narrower in scope of application, and increasingly detached from the practical reality of legal practitioners.⁸⁷

Later, Dworkin deemed these attempts to save the positivist project as uninteresting, because they did not have anything interesting to contribute outside the narrowly defined discipline of legal philosophy; not to more abstract disciplines such as political philosophy, or to less abstract disciplines such as the practice of lawyers and judges.⁸⁸ Perhaps that is the key to appreciating the merits of the legal positivist project; it may work within the narrow confines of the discipline of legal theory, but once the forces of the practical reality and the wider philosophical context start to pull, the project reaches its limits of utility.

3. Beyond the *rule of recognition*

To illustrate further the obstacle Hume's fork creates for legal positivism that is based on the legacy of Hart⁸⁹ it is helpful to conduct a simple thought experiment. Imagine a society that has laws and those laws derive their authority from the fact that they have been enacted in a way compatible with a master rule that declares how rules should be enacted. The people of this society know that they have obligation to follow the law because they recognise the authority of the master rule to clarify that issue. They thus have a valid *a priori* reason to follow the law. Few of the citizens might however start to wonder why the master rule can dictate which laws

⁸⁷ Brian Leiter summarises the status of the debate in a recent article and proclaims (once again) a positivist victory over the Dworkinians and the naturalists. If it is a positivist victory, it may however have been Pyrric; the claim left standing is reduced to a narrow statement that positivism is the best theory to explain what the ordinary man thinks about the laws: '*if we take seriously Hart's explicit theoretical aim of doing justice to what the ordinary man understands about the modern municipal legal system, then we have no better theory than positivism*'. See Brian Leiter, 'Legal Positivism About the Artifact Law: A Retrospective Assessment' in L. Burazin, K. E. Himma, C. Roversi (eds) *Law as an Artifact* (Oxford University Press, forthcoming 2017) at note no 39. < <https://ssrn.com/abstract=2870877> > accessed 16 January 2017.

⁸⁸ See Ronald Dworkin, 'Hart's Postscript and the Character of Political Philosophy' (2004) 24 *Oxford Journal of Legal Studies* 1, 36-37. Brian Leiter responded to this criticism on the legal positivist project in a critical article, see Brian Leiter, 'The end of Empire: Dworkin and the Jurisprudence in the 21st Century' (2005) 36 *Rutgers Law Journal* 165, 178.

⁸⁹ The objection raised against legal positivism in this section mainly focuses on the inclusive tradition which Hart sanctioned in *the Postscript* as the right interpretation of his project. The exclusive position advocated by Raz and few of his former students is caught by this objection but slightly differently and will only be dealt with briefly.

are valid. Who decided so in the first place and more importantly, on which grounds? Could it be that the master rule is based on faulty premises and thus everything that derives from it as well?

The sceptical citizen in the example needs *a priori* reason, a reason that exists *a priori* and independently of the master rule itself, for it to be cognitively plausible that a rule could possess the suggested authority.⁹⁰ This would of course not be a problem if the master rule possessed authority in the virtue of itself, instead of being a product of a human action. In that case, it could be claimed that since the rule does not rely on anything previous, its authority is inherent, which eliminates the need for *a priori* reason to explain its authority.

Raz's argument, about the autonomous authority of the laws, seeks to overcome this problem of normativity by using this strategy. This is however problematic in the positivist world where the laws are the expressions of manmade social facts. The laws can hardly assume autonomously *a priori* authoritative role, as if they had nothing to do with their authors and their intentions.⁹¹ This is however not the case on the Hartian version of legal positivism. The master rule is the product of a past cognitive human decision, or a series of decisions, which he labelled social practices and others have labelled social conventions.⁹² Deferral to social conventions or practices begs a question; based on which ultimate reasons were these past decisions taken?

⁹⁰ John Gardner and Timothy Macklem make an attempt to discount this objection by comparing it to the chicken and the egg logical problem. They are surely mistaken in that regard; it does not take an instance of law to make a human in the same way as it takes an egg to make a chicken. It does however take a human to make an instance of law (e.g. the *rule of recognition*), just as it takes a chicken to make an egg. See further John Gardner and Timothy Macklem, 'Scott J Shapiro: Legality' (2011) Notre Dame Philosophical Reviews <<http://ndpr.nd.edu/news/27609-legalty/>> accessed 19 January 2017.

⁹¹ This argument of the exclusivists is somehow deeply counterintuitive and I have not yet seen a plausible explanation of how the initial instance of a law is possible without an *a priori* reason of some kind. Leiter tries unconvincingly to provide explanations in a recent paper and then ultimately tries to impose the burden of proof on those that are not convinced by the autonomous authority of the *rule of recognition*. See Brian Leiter, 'Legal Positivism About the Artifact Law: A Retrospective Assessment' in L. Burazin, K. E. Himma, C. Roversi (eds) *Law as an Artifact* (Oxford University Press, forthcoming 2017) at notes no 15-39. <<https://ssrn.com/abstract=2870877>> accessed 16 January 2017.

⁹² See for example; Jules Coleman, *The Practice of Principle: In Defence of a Pragmatist Approach to Legal Theory* (Oxford University Press 2003).

If humans act rationally, a specific reason would normally be needed for an action to make sense. Acting mistakenly or not, the human mind consistently seeks to direct human action towards maximisation of something that he believes to be of a value. It could of course be that the human mind in fact acts randomly, although it would be hard to make sense of the human condition if that was the case. If the human mind acted randomly, as opposed to rationally, an authoritative facility like the master rule could be formed by a coincident without a deliberate *a priori* reason justifying it. Hart's thesis of the social practice could be read in that way; people start social practices without any *a priori* reason randomly. This reading of the source of the *rule of recognition* however runs into a problem when required to explain how an initial random act becomes a social practice without a deliberate decision to replicate the random act. Could something be labelled as a social practice that is both formed and followed randomly?

It seems as if something that is practiced socially, must at least be supported by a reason of no harm. If something happens initially through a random occurrence, an action would not be taken to replicate the occurrence unless it would prospectively be believed to be at least harmless compared with alternative courses of action. Most often, such deliberate actions would in fact only be taken based on something believed to be *a priori* beneficial. Thus, a social practice can only be sustained and formed by a deliberate action to act in a specific way, and a deliberate action must be based on *a priori* reason of some kind.

For illustrative purposes, we can apply this logic to two well know thought experiments: The question of *why we drive on this, or that, side of the road*, is sometimes asked in classrooms of legal theory. It could be deemed as a historical coincident, on which side of the road one ought to drive, and further explained as an authoritative rule that is not supported by any *a priori* reason (at least not a reason that is valid today); it simply sprung out of a social practice or a coordination convention. However, by thinking beyond the social practice, or the coordination convention, the proper answer to the question quickly emerges. Why do we drive on the side of the road instead of in the middle of the road? Most people have no trouble answering this question and most would give the same reason why it makes sense to drive on the side of the road. Left or right does not really matter. The key thing is to avoid colliding when traveling in

opposite directions and by that solve a coordination problem efficiently. The very first instance of someone driving on the right side of the road instead of on the left side might be a complete coincidence, but for others to make a social practice out of doing the same would require *a priori* reason such as; it makes sense that we all do the same to avoid collisions.

Another thought experiment that is often used as an example in legal theory classrooms is the social etiquette of taking hats off when entering a church. Initially, this seems an entirely random practice, which is not supported by any particular *a priori* reason. A brief contemplation however reveals incentives that might suffice as reasons. Obviously, it might infuriate fellow churchgoers if one were to breach the code of etiquette; they might feel offended and their feelings might be hurt in a place where they seek mental refuge and healing. The potential of being lynched by an angry mob provides an incentive to just take off our hats and thus avoid any confrontations. We thus take off our hats in churches to respect the feelings of others and that in combination with the potential of becoming subject to social sanctions becomes the *a priori* reason to comply with the rule of etiquette.⁹³

What initially seem random rules of etiquette or coordination conventions, are at a closer look usually founded on and maintained by normative *a priori* reasons. The former example above relies on the efficiency of the solution it provides to a coordination problem. The later example relies on the credibility of the threat of sanction that supports it. At first sight, these reasons seem unrelated, but at a closer look, it becomes apparent that they can both be translated into economic incentives that are sufficient to give a reason to act. Based on this rationale, we can predict that even the most obscure social practices are bound to be the result of deliberate decisions that are based on, or maintained, by incentives that reason with us cognitively.

⁹³ Leiter takes rules of etiquette as an example of rules that do not have any creator or any initial intention: *'Think of etiquette: even the particular norms of etiquette are rarely intentionally created, and the institution of regulating the boundaries of informal social interaction by, for example, norms of politeness, respect and courtesy—“etiquette”—is a kind of human social practice that was not created by anyone or any group for any particular purpose, [...]'*. See Brian Leiter, 'Legal Positivism About the Artifact Law: A Retrospective Assessment' in L. Burazin, K. E. Himma, C. Roversi (eds) *Law as an Artifact* (Oxford University Press, *forthcoming 2017*) at note no 21. <<https://ssrn.com/abstract=2870877>> accessed 16 January 2017. Arguably, rules of etiquette do have an intentional purpose, contrary to what Leiter maintains, and initially must necessarily have been created by someone. Often, rules of etiquette are meant to express certain feelings in human interaction. A random gesture can thus become a proxy for such an expression if the meaning of the proxy is understood as such by the relevant population. On the other hand, the gesture might seem odd to outsiders that do not know which expression the gesture is meant to convey.

We have seen that the plausibility of Hart's theory hinges on the emergence of social practices that do not have any purposeful *a priori* reason for existing. This premise is however defeated by Hume's fork, as is apparent by the difficulty of explaining how creating and maintaining a social practice is possible without an anchor in *a priori* reason for action.

It should be noted that the other large branch of legal theory, the natural law tradition, does not share the '*ought from an is*' problem of legal positivism. The natural law tradition claims that the ultimate authority that supports any legal rule is its moral appropriateness. Morality, on this view, is *a priori* to any man-made law and is as such independently authoritative; one ought to follow the law because it is the morally right thing to do. Morality, unlike the *rule of recognition*, is authoritative in the virtue of itself. A person does not need any further reasons than moral appropriateness to have a valid reason for any given action.

Despite this quality, the natural law tradition has its own set of controversies. A typical objection capitalises on the indeterminacy of the concept of morality; a judgment of moral appropriateness often lies in the subjective eye of the beholder and is thus hard to establish objectively. Natural law theory is also often challenged with the paradox of the laws of Nazi Germany, which supposedly were not entirely compatible with universal values of morality, but were nonetheless laws that formed a valid reason for action. This is a powerful objection; if moral appropriateness is a function of legality, how can then immoral laws exist?

The two main branches of legal theory rely, when pushed to their limits, on two incompatible assumptions. Natural law theory relies on the assumption of the moral *a priori*, and positivist legal theory relies on the assumption of a social practice that does not have a normative *a priori*. In their absolute forms, these assumptions cannot coexist. The debate about these positions has hit a dead-end within the confines of the discipline of legal theory. One can simply decide on which side of the fence to work by adopting whichever of the two incompatible assumptions. To move forward and seriously assess their merits we must slide for a moment into the domain of political philosophy. The discipline that deals with the origin and being of the political systems that ends up producing laws. To understand why and how some laws are

morally bound while others are not, we need to understand the organs that the laws serve and to which they owe their existence.

For the purposes of the main research question of this thesis, i.e. how do we know what is a fair procedural rule, the choice of a legal theoretical assumption is the key to an answer. Using the positivist assumption of the non-normative social practice, implies that fairness in the context of laws is whatever the legal system recognises as fair. Using the naturalist assumption about the *a priori* normative element of morality implies that fairness in the laws is a derivative from something that exists as a constant independently of the laws. The answer we are seeking for the purposes of this research is fundamentally normative, which means according to *Hume's fork*, that the answer is not to be found in the description of the positive law, but rather in the *a priori* element (i.e. in the dark matter) that lends the concept of fairness in the laws meaning independently of the positive laws.

Before we enter the domain of political philosophy, a categorical distinction needs to be made. From above we recall that one of the principle tasks of Hart's theory was to rescue positivist legal theory from reliance on Austin's *command model* that could not be distinguished for the gunman situation. He was thus trying to build a theory that did not rely on the threat of violence as the ultimate authority of laws. A similar development occurred much earlier within the context of political philosophy. The rise and rediscovery of the idea of democracy during the enlightenment period, created a demand for a political theory about the potential origins and being of political organisation that did not rely on the ultimate authority of the coercive monarch. Coercive authority in that sense is the antithesis of democracy. When we examine the foundations of a social practice or a convention we do so from the perspective of democracy, and that is surely the context within which positivist legal theory has been discussed since the publication of Hart's monograph. The task is thus to explain Hart's social practices without reliance on coercion as an ultimate reason. Social contract theory is the branch of political philosophy that deals with the origins and being of political organisations from this non-coercive point of view.

4. The social contract

The idea of the social contract is, in its basics, a simple metaphor. The metaphor of the contract tries to explain the being and function of political organisations, using insights from how we understand actual contracts and agreements. Over time understanding of contracts and bargaining has developed and with it the metaphor of the social contract. In the following passages, I will shortly review the historical development of the idea, before exploring in more depth a specific modern perspective that uses insights from game theory. For the purposes of explaining the origin and being of social practices and conventions on which positivist legal theory relies, and the element of morality on which natural law theory relies, the focus of discussion will be on the normative element of the social contract.

4.1. Origins and evolution of the social contract

The idea of human society as a form of an implicit agreement between free citizens precedes the modern era by millennia.⁹⁴ John Rawls's sophisticated and celebrated modern version draws on ideas developed during the enlightenment period. To understand fully his moves and motives it is thus helpful to analyse shortly the texts on which he relied, and more interestingly on the texts he did not cite.

Niccolò Machiavelli's *Il Principe* (e. *the Prince*) of 1513 was a revolutionary text in many ways. Although not the main topic of the monograph, Machiavelli discussed briefly the logic of power in a principality that was founded through the favour of the citizens. Machiavelli referred to it as the '*civil principality*', which by definition was not founded on, or held together through coercion or violence: '*A man who becomes ruler through popular favour [...] must keep the people well disposed towards him. This will be easy, since they want only no to be oppressed.*'⁹⁵ For many centuries, Machiavelli had a bad reputation due to his cynical outlook on the function of politics and power. By modern standards his vision is however still very current and builds on the presumption of a rational human mind responsive to self-preserving, and at times, cynical incentives. Machiavelli recognised that in the absence of coercion, the political

⁹⁴ See for example Plato's *Crito*.

⁹⁵ See Niccolò Machiavelli, *The Prince* (first published 1513, Quentin Skinner and Russel Price eds, Cambridge University Press 1988) 36.

establishment of the civil principality must govern through the consent of the citizens. This can be taken as an early description of a social contract.

Thomas Hobbes predicted grimly in the *Leviathan* of 1651,⁹⁶ that without an agreement between people, the inherent nature of humans would throw them into war of all against all. He observed that humans are rather homogeneous regarding physical and mental abilities, and that this should predispose them towards similar needs and desires, which ought to create high demand for specific materials. High demand in turn leads to low supply, and thus the war in the state of nature over the limited supplies becomes inevitable with its alarming consequences:

*'In such conditions there is no place for hard work, because there is no assurance that it will yield results; and consequently no cultivation of the earth, no navigation or use of materials that can be imported by sea, no construction of large buildings, no machines for moving things that require much force, no knowledge of the face of the earth, no account of time, no practical skills, no literature or scholarship, no society; and—worst of all—continual fear and danger of violent death, and the life of man solitary, poor, nasty, brutish, and short.'*⁹⁷

Hobbes believed that these dreadful prospects and the potential of achieving greater prosperity should make humans preconditioned towards reaching an agreement with their fellow beings; once reason has compelled men to seek peace, they *'should be willing (when others are too) to lay down [their] right to everything, and should be contented with as much liberty against other men as [they] would allow other men against [themselves].'*⁹⁸ The primary motive for entering voluntarily into an agreement, according to Hobbes, is rationality from an individual point of view. The individual recognises, Hobbes argues, the personal benefit of

⁹⁶ An online version can be found here: <<http://www.earlymoderntexts.com/authors/hobbes>> accessed 1 November 2016.

⁹⁷ See Thomas Hobbes, *Leviathan* (first published 1651, Richard Tuck ed, 2nd edn, Cambridge University Press 1996) Part I, Ch 13.

⁹⁸ See Thomas Hobbes, *Leviathan* (first published 1651, Richard Tuck ed, 2nd edn, Cambridge University Press 1996) Part I, Ch 14.

cooperation and he is prepared to make the effort of realising this potential on the condition that others do the same.

An essential element in Hobbes's elaboration of the social contract is the recognition of what would be called *cheap talk* in modern economics. Hobbes argues that words and agreements have no meaning in themselves. It is all about creating the right conditions, on the basis of which, the rational agent acts. In Hobbes mind, it was thus essential for the success of the social contract, that it could be enforced by '*some coercive power*'.⁹⁹ Although this conclusion in retrospect was perhaps unwarranted, its central insight about the role of incentives as the motor of the rational human being is still relevant. This sets him apart from many of his contemporaries that routinely used religiously based axioms in their theories and from later theorists in the deontological tradition.

John Locke in his *Second Treaties of Government* of 1689¹⁰⁰ contributed to social contract theory differently. Locke's starting point is the natural human condition. In that state, *the state of nature*, humans are free to act as they please so long as they abide to *the law of nature*.¹⁰¹ The law of nature burdens humans with the obligation to refrain from harming '*anyone else in his life, health, liberty, or possessions.*' Locke's foundation for this law has not aged particularly well in the age of secular science. In short, he argued, that since humans were all made by God and are thus all the servants of his deputy Jesus, they are by nature not allowed to harm themselves nor other equally ranked properties of God.¹⁰²

The leap from the *state of nature* to a human society must be based on consent, according to Locke, and this consent is triggered by the person's '*intention of better preserving himself, his liberty and property*', to better ensure the enforcement of the law of nature.¹⁰³ In Locke's view,

⁹⁹ See Thomas Hobbes, *Leviathan* (first published 1651, Richard Tuck ed, 2nd edn, Cambridge University Press 1996) Part I, Ch 14-15.

¹⁰⁰ An online version can be found here: <<http://www.earlymoderntexts.com/authors/locke>> accessed 1 November 2016.

¹⁰¹ See John Locke, *Two Treaties of Government* (first published 1689, Peter Laslett ed, Cambridge University Press 1988) Para 4.

¹⁰² See John Locke, *Two Treaties of Government* (first published 1689, Peter Laslett ed, Cambridge University Press 1988) Para 6.

¹⁰³ See John Locke, *Two Treaties of Government* (first published 1689, Peter Laslett ed, Cambridge University Press 1988) Para 95.

the power of society over men is limited by the scope of the powers they had themselves in the state of nature.¹⁰⁴ This forms the initial moral obligation of society against its citizens. However, contrary to Hobbes, this obligation is not based on the willingness of a rational person to concede rights, but instead on the natural law that cannot be conceded by anyone due to its divine origins.

Jean-Jacques Rousseau argued, in his influential *Du contrat social ou Principes du droit politique* (e. *the Social Contract*) of 1762,¹⁰⁵ that self-preservation is the primary natural motive of humans. This motive, in turn, controls human logic in the state of nature.¹⁰⁶ Hardship in the state of nature, argued Rousseau, may however compel humans to take refuge in the company of others in a civil society, which can only happen through an agreement.¹⁰⁷ In the state of nature, Rousseau explained, the individual had natural liberty to act on impulses and instincts, while in the civil state he has civil liberties that are limited by the moral obligation to conform to the general will.¹⁰⁸ The power of the sovereign, which personifies the general will, is in turn restricted by a special requirement:

*'[T]he social compact creates an equality among the citizens so that they all commit themselves to observe the same conditions and should all have the same rights. Thus, from the very nature of the compact, every act of sovereignty—i.e. every authentic act of the general will—obliges or favors all the citizens equally; so that the sovereign recognizes only the body of the nation and doesn't distinguish among the individuals of whom it is made up.'*¹⁰⁹

¹⁰⁴ See John Locke, *Two Treatises of Government* (first published 1689, Peter Laslett ed, Cambridge University Press 1988) Paras 131, 134-35.

¹⁰⁵ An online version in English can be found here: <<http://www.earlymoderntexts.com/authors/rousseau>> accessed 1 November 2016.

¹⁰⁶ See Jean-Jacques Rousseau, *The Social Contract* (first published 1762, Christopher Betts tr, Oxford University Press 1994) Book I, Ch 2.

¹⁰⁷ See Jean-Jacques Rousseau, *The Social Contract* (first published 1762, Christopher Betts tr, Oxford University Press 1994) Book I, Ch 4 and Ch 6.

¹⁰⁸ See Jean-Jacques Rousseau, *The Social Contract* (first published 1762, Christopher Betts tr, Oxford University Press 1994) Book I, Ch 8.

¹⁰⁹ See Jean-Jacques Rousseau, *The Social Contract* (first published 1762, Christopher Betts tr, Oxford University Press 1994) Book II, Ch 4.

The sovereign cannot treat one of its parts more preferably than the others. Equality thus becomes the initial moral obligation of the sovereign against the citizens within the confines of the social contract. In this, Rousseau realises that the act of cooperating on a mutual task, involves *a priori* commitment to behave in a specific way and that by doing so one can expect others to be required to do so as well. The intention of giving equality and the expectation of receiving equality, according to Rousseau, is a prerequisite for forming a social contract and its primary output once established.

Immanuel Kant criticised Hobbes's views on political theory¹¹⁰ in part II of his short essay; *In Theory and Practice* from 1793.¹¹¹ However, Kant's most important innovation with regards to later developments of social contract theory is not as such found in his description of the foundations of political authority; his theory of the *categorical imperative* proved far more important. The *categorical imperative* was first described in *Grundlegung zur Metaphysik der Sitten* (e. *Groundwork for the Metaphysics of Morals*) published in 1785.¹¹² His argument takes off as follows:

'Every thing in nature works according to laws. Only a rational being has the capacity to act according to the representation of laws, i.e. according to principles, or a will. Since reason is required for deriving actions from laws, the will is nothing other than practical reason. If reason determines the will without fail, then the actions of such being that are recognized as objectively necessary are also subjectively necessary; i.e. the will is a capacity to choose only that which reason, independently of inclination, recognizes as practically necessary, i.e. as good.

(...)

¹¹⁰ The critique was addressed against Hobbes's political theory as expressed specifically in *De Cive*; Thomas Hobbes, *On the Citizen* (first published 1642, Richard Tuck and Michael Silverthorne eds, Cambridge University Press 1998). The *Leviathan* is however Hobbes's most important work on political philosophy.

¹¹¹ The full name in the original German version is *'Über den Gemeinspruch: Das mag in der Theorie richtig sein, taugt aber nicht für die Praxis'* (e. On the common saying: That may be correct in theory, but it is of no use in practice). An online version in English is available here: <<http://www.sussex.ac.uk/Users/sefd0/tx/tp2.htm>> accessed 30 December 2016.

¹¹² An online version in English can be found here: <<http://www.earlymoderntexts.com/authors/kant>> accessed 1 November 2016.

The representation of an objective principle in so far as it is necessitating for a will is called a command (of reason), and the formula of the command is called imperative.

All imperatives are expressed by an ought, and by this indicate the relation of an objective law of reason to a will that according to its subjective constitution is not necessarily determined by it (a necessitation). They say that to do or to omit something would be good, but they say it to a will that does not always do something just because it is represented to it that it would be good to do it. Practically good, however, is what determines the will by means of representations of reason, hence not from subjective causes, but objectively, i.e. from grounds that are valid for every rational being, as such. It is distinguished from the agreeable, as that which influences the will only by means of sensation from merely subjective causes, which hold only for the senses of this or that one, and not as a principle of reason, which holds for everyone.¹¹³

Kant then continued to pronounce all imperatives as either categorical or hypothetical, meaning they either objectively or subjectively command the rational will of a person. Kant's central thesis was that moral duties are categorical imperatives. They imply commands that apply to the will of any rational person, irrespective of the ends this person may decide to pursue. Kant contrasts the categorical imperative with the hypothetical imperative. The hypothetical imperative implies a command that is conditioned upon a will towards a certain end. The hypothetical imperative thus commands that one must do something if a certain end is being pursued, while the categorical imperative does not rely on any specific end. It simply commands categorically in any case.¹¹⁴

Acting morally, according to Kant, is thus acting in a way that could be seen as universally rational. The rational will always commands a practically good action, and if the action can be considered universally good irrespective of specific ends, it is a morally good action. Or, as Kant

¹¹³ See Immanuel Kant, *Groundwork for the Metaphysics of Morals* (first published 1785, Mary Gregor and Jens Timmermann eds, 2nd edn, Cambridge University Press 2012) 26-27.

¹¹⁴ See Immanuel Kant, *Groundwork for the Metaphysics of Morals* (first published 1785, Mary Gregor and Jens Timmermann eds, 2nd edn, Cambridge University Press 2012) 28-34.

put it: *'act as if the maxim of your action were to become by your will a universal law of nature.'*¹¹⁵

After Kant and following the rise of economics and utilitarianism in the 18th and 19th centuries, the social contract idea as a basis for theorising about the organisation of society became a marginal topic of academic discourse. By the mid-20th century, the idea of the social contract however made a comeback. John Rawls has largely been credited for its revival through the success of his influential 1971 monograph, *A Theory of Justice*, where he launched a powerful attack against utilitarian theories of justice by building on the social contract tradition as described by Locke, Rousseau and Kant.¹¹⁶ Rawls intentionally left out Hobbes's *Leviathan* on his list of sources of inspiration, stating simply that *'[f]or all its greatness [it] leaves special problems.'*¹¹⁷ The key ideas in Rawls's social contract are the *original position* and the *veil of ignorance*.

Rawls explained that his concept of *'the original position of equality corresponds to the state of nature in the traditional theory of the social contract.'* By rephrasing the description of the natural state prior to the existence of the social contract, Rawls sought to emphasise the hypothetical nature of the contract. It was merely meant as a thought experiment to visualise what standard of justice persons would agree on, in a vacuum of equal bargaining power.¹¹⁸ To reach the level of equal bargaining power, Rawls realised that the parties to the agreement must be oblivious about their private position, when negotiating about the standard of justice. Thus, he invented the *veil of ignorance*, behind which the negotiation takes place. The core of Rawls's argument of what happens behind the veil can be found in the following passage:

'Since all are similarly situated and no one is able to design principles to favour his particular condition, the principles of justice are the result of a fair agreement or

¹¹⁵ See Immanuel Kant, *Groundwork for the Metaphysics of Morals* (first published 1785, Mary Gregor and Jens Timmermann eds, 2nd edn, Cambridge University Press 2012) 34.

¹¹⁶ Rawls mentions David Hume, Adam Smith, Jeremy Bentham and John Stuart Mill as the main proponents of the utilitarian ideas he seeks to oppose. See John Rawls, *A Theory of Justice* (1971, 2nd edn, Belknap Press 1999) xvii-xviii.

¹¹⁷ See John Rawls, *A Theory of Justice* (1971, 2nd edn, Belknap Press 1999) 10, footnote 4.

¹¹⁸ See John Rawls, *A Theory of Justice* (1971, 2nd edn, Belknap Press 1999) 11.

*bargain. For given the circumstances of the original position, the symmetry of everyone's relations to each other, this initial situation is fair between individuals as moral persons, that is, as rational beings with their own ends and capable, I shall assume, of a sense of justice. The original position is, one might say, the appropriate initial status quo, and thus the fundamental agreements reached in it are fair.'*¹¹⁹

Rawls social contract becomes fair if the negotiators are subject to three conditions: they are rational, morally bound, and in a symmetrical bargaining position. The moral requirement sets Rawls apart from Hobbes, who insisted on the cynical nature of humans. This difference could explain Rawls's unwillingness to cite Hobbes as an early advocate of the tradition he claimed to build on.

From these premises Rawls argued that utilitarian justice principles, which strive for maximising the total utility irrespective of its distribution, are not compatible with his version of the social contract. He claimed that no rational person seeking its own means, would agree to cooperate on the terms that only others should benefit from the cooperation. Rawls suspects that an agreement, between rational moral persons in the original position, would be reached on two principles of justice that he describes in the following way:¹²⁰

*'First: each person is to have an equal right to the most extensive scheme of equal basic liberties compatible with a similar scheme of liberties for others.[...] Second: social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone's advantage, and (b) attached to positions and offices open to all.'*¹²¹

The content of the hypothetical social contract would accordingly become these two principles.¹²² The main challenge however, is to explain why a rational deliberation would result in these two principles, rather than something entirely different. Rawls applied deductive reasoning and specific contingencies to support that point. He assumed that his

¹¹⁹ See John Rawls, *A Theory of Justice* (1971, 2nd edn, Belknap Press 1999) 11.

¹²⁰ See John Rawls, *A Theory of Justice* (1971, 2nd edn, Belknap Press 1999) 13.

¹²¹ See John Rawls, *A Theory of Justice* (1971, 2nd edn, Belknap Press 1999) 53.

¹²² See John Rawls, *A Theory of Justice* (1971, 2nd edn, Belknap Press 1999) 102.

rational persons would pursue their ends within the contingencies of some specific circumstances; at some point, a stable equilibrium is formed between the pursued ends and the specific circumstances. On Rawls's view, this does however not guarantee that the equilibrium is just. Therefore, his rational person must also be moral, i.e. with a sense of justice. This adds to the specific circumstances and thus alters the feasible equilibrium point, by leaving out unjust situations. His original position thus eliminates all specific circumstances except for the moral condition. By doing this, he argued that all agreements reached in the original position are fair.¹²³ His whole argument thus depended on the moral inclination of people.

The moral character of people in Rawls's model was expressed in two ways. Firstly, the rational behaviour of people was restricted by a condition of disinterest in the relative position of others. Rawls rational person thus only strives towards its own ends, but does not act on envious emotions by sabotaging the prospects of others. In this Rawls's theory differs from Hobbes's, which argued for the natural tendency of humans to compete in situations of scarcity. Secondly, the condition of a sense of justice compels men to honour agreements that are made: *'in reaching an agreement, then, they know that their undertaking is not in vain: their capacity for a sense of justice insures that the principles chosen will be respected.'*¹²⁴ This again puts Rawls at odds with Hobbes, which remarked; *'covenants without the sword are merely words, with no strength to secure a man at all.'*¹²⁵ In this sense, Rawls relied on a moral *a priori* in the same sense as Locke, but with a sophisticated Kantian spin. Rousseau and Hobbes however argue that cooperative behaviour is dependent on a certain moral attitude that is triggered by the right incentives.

The assumption Rawls made about the moral character of humans is problematic regarding our task of exploring the possibility of forming a social practice or a convention without reference to a normative *a priori*. The approach of Hobbes, in relying only on the rational choices of individuals, seems to strike a middle ground between relying on a vague moral assumption in the naturalist sense, and relying only on descriptive premises in the positivist

¹²³ See John Rawls, *A Theory of Justice* (1971, 2nd edn, Belknap Press 1999) 103-4.

¹²⁴ See John Rawls, *A Theory of Justice* (1971, 2nd edn, Belknap Press 1999) 125.

¹²⁵ See Thomas Hobbes, *Leviathan* (first published 1651, Richard Tuck ed, 2nd edn, Cambridge University Press 1996) Part II, Ch 17.

sense. A more recent literature on social contract theory helps explaining that despite dominance of the sinister nature of humans, in the way Hobbes suggested, there remains a possibility to form cooperation without relying on coercive authority or an unexplained moral value. This literature uses the language of game theory to demonstrate its point.

4.2. Social contracts and the theory of games

David Gauthier was among the first authors to model the social contract in terms of game theory.¹²⁶ He argued for a Hobbesian understanding of morality, where morality is the result of a rational choice. Thus, cooperation through a social contract did not require any *a priori* moral virtues. The object of moral appropriateness was simply the rational choice in the bargaining dilemma of cooperating with others. Gauthier modelled the bargaining problem of the social contract as a version of the prisoner's dilemma, in which individual players are pinned against each other by the temptation of gaining by acting selfishly. Brian Skyrms¹²⁷ and Kenneth Binmore¹²⁸ have argued for a different modelling. Binmore's version is sophisticated and deserves to be elaborated further for the purposes of advancing an understanding of the origin and being of social practices and conventions.

By portraying the problem of the social contract on the basis of various games in the language of game theory we can quickly establish that cooperation is often an efficient equilibrium solution. Usually however, there are also efficient equilibrium solutions that are not cooperative. This leaves us with the problem of why a cooperative equilibrium would be selected rather than an individualistic. This of course assumes that there is not any *a priori* moral reason in the deontological sense that resolves the issue; we are speaking of rational choice in the game theoretical sense.

¹²⁶ See David Gauthier, *Morals by Agreement* (Oxford University Press 1986).

¹²⁷ See Brian Skyrms, *Evolution of the Social Contract* (1996, 2nd edn, Cambridge University Press 2014); and Brian Skyrms, *The Stag Hunt and the Evolution of the Social Structure* (Cambridge University Press 2004).

¹²⁸ See Ken Binmore, *Game Theory and the Social Contract Vol. I: Playing Fair* (The MIT Press 1994); Ken Binmore, *Game Theory and the Social Contract Vol II: Just Playing* (The MIT Press 1998); and Ken Binmore, *Natural Justice* (Oxford University Press 2005).

Binmore builds an interesting theory based on the *stag hunt game*¹²⁹ that provides a solution to this issue. He claims that a '*social contract is the set of common understandings that allow the citizens of a society to coordinate their efforts*'. In his mind, positive law or moral sentiments do not keep societies together. The social contract is a rational solution to a coordination problem and '*does not need any glue*' to exist. Binmore argues that three conditions ranked in an order of priority must be fulfilled to achieve a successful social contract. The primary condition is stability. If not stable, a contract will not exist for long and is thus of negligible importance. To become stable, the common understanding and the coordinated behaviour must form a Nash equilibrium, in which each citizen's strategy is the best response to the strategies of other citizens. Efficiency is the secondary condition. The stable equilibrium must be efficient in comparison with competing social contracts to survive. Efficiency helps selecting out of the total available Nash equilibriums, a viable set of efficient equilibriums. Several equilibriums can however result in a comparable efficiency, which leads us to the third condition. A test of fairness weeds out of the efficient equilibriums a single right equilibrium upon which the social group should coordinate. The social contract must thus be: stable, efficient, and fair.¹³⁰

The question of why a cooperative equilibrium would be chosen, rather than an individualistic equilibrium, in a game is answered at the second tier of Binmore's formulation. Binmore explains that instead of resorting to normatively anchored concepts like trust, duty and authority, game theory can explain the phenomenon based on the *folk theorem*. The *folk theorem* predicts that in a repeated game with an infinite time frame, a cooperative efficient equilibrium can be maintained despite the potential of a one-off gain by cheating on the cooperation. If, however, the game has a definite end in sight, it will unravel in a race of

¹²⁹ The Stag Hunt game is about the dilemma of whether to cooperate with others on the hunt for achieving the great price of the stag, or whether to rely solely on yourself and hunt for the less valued hare. The catch is that if you choose to cooperate and the others won't you end up with nothing since cooperation is required to accomplish the great enterprise of hunting the stag. There are two Nash equilibriums in the Stag Hunt game. Either everyone hunts the stag or everyone hunts the hare. The Stag Hunt game is derived from Rousseau's story of the stag hunt: '*In this manner, men may have insensibly acquired some gross ideas of mutual undertakings, and of the advantages of fulfilling them: (...) If a deer was to be taken, every one saw that, in order to succeed, he must abide faithfully by his post: but if a hare happened to come within the reach of any one of them, it is not to be doubted that he pursued it without scruple, and, having seized his prey, cared very little, if by so doing he caused his companions to miss theirs.*' See Jean-Jacques Rousseau, *A Discourse on the Basis and Origin of inequality Among Men* (first published 1754, Bedford/St. Martin's 2010) Part II.

¹³⁰ See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 3-14.

opportunisms.¹³¹ This still only shows that it *can* be explained as rational to cooperate, despite an incentive not to cooperate in a repeated game, but this does not show that a cooperative equilibrium *will* be chosen.

The nudge that eventually pushes humans towards cooperative social contract equilibriums can be explained, according to Binmore, based on the theory of natural selection.¹³² Inspired by Dawkins's theory of the selfish gene,¹³³ Binmore explains that individuals within a group seek to maximise their personal fitness by acting in accordance with the prevailing efficient equilibrium within the group. This group equilibrium is not necessarily settled on the most efficient equilibrium possible, other potentially more efficient group equilibriums might exist. If this group would encounter another group that was settled on a more efficient equilibrium, it soon must either amend its ways to withstand competition, or face extinction from the game of life. This over time forces groups to settle on Pareto optimal equilibriums, where nothing goes to waste and nothing can be improved, without undermining the group's strategy against other competing groups. This is how the stag hunt will be chosen as a social contract, over the hare hunt in the *stag hunt game*. The group that settles on the less efficient hare hunt will not succeed in the game of life when competing with a group that hunts the stag. In principle, the same argument applies if the competing groups are playing the prisoners dilemma, or any other game that includes a Nash equilibrium that is not Pareto optimal.¹³⁴

This, however, is not enough to solve the equilibrium selection problem. Even if natural selection pushes towards Pareto optimality, many equilibrium solutions are still available. Pareto optimality can be achieved in various ways through different distribution of the payoffs from the efficiency gain. A further tool is thus needed to coordinate on a single efficient equilibrium to avoid destabilisation due to incoherency. Binmore argues that this selection tool

¹³¹ See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 10. It should be noted that the folk theorem is usually used to show that cooperation is possible even where there is a strong incentive to cheat like in the prisoner's dilemma game. It will thus also apply to show the possibility to cooperate when the incentive not to cooperate is lesser than in the prisoner's dilemma game, which is the case with the stag hunt game.

¹³² See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 7-14.

¹³³ Reference is made to the evolutionary theory argued for in; Richard Dawkins, *The Selfish Gene* (Oxford University Press 1976).

¹³⁴ Professor Simon Deakin raised an objection during our discussion about an earlier draft of this text, that perhaps the social contract was not a Stag hunt game. This is a valid objection, but does not in principle matter for the soundness of the argument. The point about the intergroup superiority of Pareto optimality stands.

has the same deep roots in the human condition as the efficiency selection tool. When faced with a simple problem of distribution of an efficiency gain, people somehow by instinct feel that a certain distribution is right, while another is wrong. We all know how to split a cake fairly as an example. This intuitive feeling is a sense of fairness, which Binmore believes, is the ultimate arbitrator of equilibrium selection.¹³⁵

According to Binmore, utilitarian and egalitarian theories of fairness are the main contenders for providing a tool to solve the selection problem.¹³⁶ Utilitarian solution is at a point where the weighted sum of payoffs is largest, while the egalitarian solution is '*the efficient outcome at which each player's weighted gain is equal.*'¹³⁷ To simplify we can assume that two utilitarians would agree on their highest combined payoff, while two egalitarians would agree on the equilibrium that leaves them both equally off, irrespective of the total combined payoff.

Binmore uses Rawls's theory of justice as fairness as a sophisticated example of egalitarianism and John Harsanyi's utilitarian theory of the social contract as a sophisticated example of utilitarianism. Binmore thinks Rawls is right, but for the wrong reasons. He argues that the thought experiment of the *original position* and the *veil of ignorance*, somehow intuitively make perfect sense, but that the reference to Kantian ethics fails to explain why the idea seems so right. Binmore's thesis is that Rawls's theory describes fairness norms that we use to solve countless everyday small-scale coordination problems. Problems that we solve so effortlessly, that we hardly realise we are solving problems. The fairness norms we use effortlessly in our daily life are deeply embedded into our existence. In fact, Binmore believes, they are written into our genes. He argues that evolutionary pressures could have created such a genetic prescription and influenced how this biological mechanism works together with our cultural heritage in choosing equilibriums in the game of life.¹³⁸ Binmore summarises his position in the following way:

¹³⁵ See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 14.

¹³⁶ See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 23.

¹³⁷ See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 29, 31.

¹³⁸ See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 15-17.

*'[M]y theory of fairness (...) is the claim that all fairness norms in actual use share the deep structure of Rawls' original position. This deep structure is biologically determined, and hence universal in the human species, but the standards of interpersonal comparison that the original position needs as inputs are culturally determined, and hence vary with time and place.'*¹³⁹

To explain why a utilitarian concept of fairness does not work, Binmore asks a rhetorical question: *'why should I maximize the sum of utilities rather than my own?'*¹⁴⁰ This highlights the contingency that utilitarian theory depends on. In a world of rational agents, an equilibrium based on utilitarian distribution of payoffs, cannot be formed, or maintained without external enforcement. Rational agents will always prefer to maximise their own utilities rather than the total utilities of the group, unless policed to do otherwise. Those who proportionally gain less from a utilitarian solution will have an incentive to team up with others in the same position and deviate from the social contract. In theory, the utilitarian solution is the best solution in terms of maximising utility, but in practice external enforcement may be impractical or impossible, which reduces its feasibility and congruently elevates the feasibility of alternative equilibrium solutions that are not contingent on external enforcement.¹⁴¹

Binmore builds a simple formula to explain how an egalitarian fairness solution is not subject to this dependency on external enforcement. He assumes that in a coordination problem, two players would agree on a Nash bargaining solution, which is the maximum product of their gains from the point of disagreement. He then adjusts the social indexes of both players to match either a utilitarian or an egalitarian fairness solution, both of which are fixed at the point of the Nash bargaining solution. If then the available equilibriums of efficiencies are suddenly expanded and thus an incentive created to bargain a new social contract, the two standards of fairness lead to different equilibriums. The utilitarian solution seeks to maximise the total utilities, which can only be achieved by external enforcement because some players will have incentive to oppose the new distribution of utilities, which is potentially asymmetric with their current social indexes. The egalitarian solution is however always symmetric, so that none of

¹³⁹ See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 18.

¹⁴⁰ See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 149.

¹⁴¹ See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 163.

the players has an incentive to oppose the new distribution. Everyone receives an increase in payoffs proportional to their current social indexes. The egalitarian solution is thus not dependent on external enforcement; no one has an incentive to deviate.¹⁴²

The consequence of this formulation is that a social contract cooperation, based on an egalitarian distribution of payoffs, can be formed, and sustained spontaneously without an external enforcement. A social contract cooperation based on utilitarian distribution however requires external enforcement to stay in place. This provides the last piece in the puzzle of how a single equilibrium, out of the many available Pareto optimal equilibriums, not only can, but also will be selected. Without external enforcement, a rational person would seek an equilibrium solution that achieved egalitarian distribution of payoffs. Other Pareto optimal equilibriums would need an external intervention to nudge in their favour, to override the spontaneous bargaining processes that by default seeks an egalitarian distribution. By this Binmore achieves to explain the being of the social contract based on human rationality in a Machiavellian manner as Hobbes and Rawls did before him, but escapes having to rely on Hobbes's external enforcement assumption and Rawls's Kantian inspired moral *a priori*.

4.3. The ultimate priori of the social contract

The social contract thought experiment, in its various forms, usually tries to explain the logic of the move from the pre-cooperative stage of humans to the cooperative stage. Hobbes saw cooperation as a rational efficiency enhancing steep, away from the gruesome prospects in the state of nature. Locke saw cooperation as a natural consequence of a morally bound human nature, using divinity as the ultimate moral authority. Rousseau recognised the principle of equality as inherited in cooperation, both as its prerequisite and its consequence, but egalitarian attitude was on Rousseau's view subject to a deliberate choice, and thus not a destined human behaviour. Kant explained morality as a categorical imperative, which secularised morality as an ultimate authority for human action. Rawls built on Locke's morally bound human nature, but with morality founded on Kant's categorical imperative instead of theology. He then explained how his morally bound person would form a social contract based on Rousseau's principle of egalitarian fairness.

¹⁴² See Ken Binmore, *Natural Justice* (Oxford University Press 2005) 158-159, 173-175.

From this, we can see two distinct ideas emerge about the origins of morality in cooperation. On one hand, we have human morality that is an *a priori* to all human action, which implies that any product of a human action is morally bound. This is the approach of Locke and Rawls. On the other hand, we have morality that is subject to the free will and created by deliberate reciprocal human interaction, which accordingly is not an *a priori* to all human action. This is the selfish reciprocity of Hobbes and the egalitarian nature of Rousseau's social cooperation. On this later view, morality only exists when humans restrain themselves and interact in a cooperative way; i.e. acting morally is the act of the rational mind, not a predefined part of the human condition.

These two approaches to morality in cooperation face different challenges. The former has a similar obstacle as the naturalist legal theorist; if humans are inherently moral, then how are immoral actions possible? The later faces a problem of stability; if humans are only subject to their free will, then how is moral cooperation possible without becoming to a quick end due to opportunistic behaviour?

Binmore gives us further clues about how to approach this problem. He argues, through game theory modelling, that it is rational to act morally. Additionally, he uses anthropological and biological theories to depict a plausible evolutionary explanation about how this could have occurred. By this he tries to prove that there are scientific explanations for the phenomenon that philosophers have through the centuries referred to as morality. Instead of refereeing to a Kantian moral *a priori*, he shows mathematically that egalitarian distribution of payoffs is a superior equilibrium solution in the game of life. The crucial point is however, that other equilibriums can be maintained, but only artificially through external enforcement.

Hobbes was on to something when he realised that humans need a plausible reason to behave themselves in cooperation with others. Rousseau realised that an egalitarian attitude was required within the confines of a workable social contract. Locke and Kant knew that a moral constant would enable successful cooperation. Rawls melted these ideas together into a sophisticated theory, but failed to realise the truth of Hobbes's cynicism.

What we can draw from this is that there certainly is a moral constant, which is imprinted into our rational minds. There is a single right way to solve coordination problems in human cooperation spontaneously, which involves egalitarian distribution of the gain from the efficient solution. Egalitarian fairness is thus the basis of morality and moral behaviour in the company of others. Egalitarian fairness can be reduced into a singular entitlement, and a corresponding singular duty; each has the moral entitlement to be treated with relative equality, and each has the duty to treat others with relative equality. Interaction based on these principles results in a stable Pareto efficient cooperative equilibrium, where no one has an incentive to withdraw his cooperativeness.

From the literature about the social contract, we can gather that the crucial point is to explain the motor of the moral attitude. What compels the free human mind to act morally? Hobbes's insight was that the motor was a matter of simple logic; either cooperate, or perish in the state of nature. He sensed that it was ultimately a choice whether to cooperate and thus whether to act morally, and that cooperative attitude depended on a proper incentive for the rational mind. Rousseau advanced Hobbes intuition by predicting that an egalitarian attitude was required to form and maintain a social contract. Rawls's reliance on Kantian ethics, in modelling his morally bound rational agent, only holds if free will is removed. Even though there is a morally right attitude, it does not necessarily follow that the agent will always act morally, unless he is devoid of free will. Assuming free will, Kant's agent intrinsically has a choice between acting on the hypothetical, or the categorical imperative.

The alternative to the moral attitude in interpersonal relations, which Binmore showed through the cooperative solution of the *stag hunt game* and through the *folk theorem*, can be explained using the example of another game. The *hawk-dove game* is about competition for scarce resources in which the worst outcome results in a mutually destructive conflict between two hawks. If, however one yields and acts as the dove, the hawk receives a larger share than the dove, but the payoffs are nonetheless Pareto optimal. If, however both yield the result is not Pareto optimal since higher total payoffs can be reached through another solution. In the *hawk-dove game*, there are three Nash equilibriums. Two pure strategies, where each acts the

opposite role to the other, and one mixed strategy where the players randomly choose which role to act. The random strategy only makes sense if the players cannot decide who acts which role, but this strategy will not lead to optimal efficiency because sometimes both will act as hawks and sometimes both will act as doves. The pure strategies on the other hand rely on the capabilities of either player to maintain a hawk position and convince the other of his destiny as a dove. The optimal equilibrium solution to this game thus requires an attitude that is not morally compatible; one player would have to subdue the other. We can easily imagine political organs that are formed on the blueprint of this game and its equilibrium solution. However, as the Pareto optimal solution suggests, such organs rely on coercion to remain stable.

The importance of the attitude with which interpersonal relations are approached, also becomes apparent if the situation is depicted as the *prisoner's dilemma game*. The Pareto optimal result of the *prisoner's dilemma* is achieved by a cooperative attitude on staying silent and thus get away with the crime. The *prisoner's dilemma* involves a strong individual incentive to confess and to frame the others; but the others recognise this incentive as well and thus have an incentive to do the same. The sensible strategy is thus for all actors to confess and blame the others, in order not to be the only non-confessing culprit. This forms the Nash equilibrium for the *prisoner's dilemma*, and it is not Pareto optimal. The problem for a group that settles on the confess-confess equilibrium, is not internal instability, but instead competitive disadvantage with external groups that have found a way to cooperate on the Pareto optimal solution of the game by approaching the task with an egalitarian cooperative attitude. The *folk theorem* shows this.

We can now see that morality enters the social contract through the attitude with which interpersonal interaction is approached. The display of different game scenarios with different behavioural incentives shows this. If the interaction is approached competitively, like in the *hawk-dove game*, optimal stability requires coercion, while the *stag hunt* and the *prisoner's dilemma* acquire stable optimality through cooperative egalitarian attitude. Morality in cooperation is thus simply a question of people's strategic attitude. If a group decides to cooperate without an external enforcement mechanism, it must do so based on an egalitarian principle of fairness. Otherwise, a stable Pareto optimal equilibrium cannot be maintained

spontaneously. The output of egalitarian cooperation will necessarily bear the mark of the foundational principle. Any attempts to distribute payoffs differently will either lead to destabilisation of the cooperation, or call for coercive enforcement to maintain stability. The social contract thus creates a moral commitment upon formation, which in essence is the requirement of a cooperative attitude, which implies a compliance with an egalitarian principle of distribution.

The view on the social contract and morality that has been articulated above shows how the idea of morality as an *a priori* concept can be harmonised with the concepts of *free will* and *rationality*. Cooperation approached based on an egalitarian fairness principle is *a priori* the morally optimal approach. This optimality is achieved through the long-term efficiency prospects in the game of life. The *rational* long-term strategy on an individual level, is thus to use this principle of *morality* for guidance when exercising *free will* (i.e. be cooperative and flourish). By exercising *free will* opportunistically, long-term efficiency is put at risk for a short-term gain. Such a behaviour is irrational in the game of life and thus immoral (i.e. be egoistic and perish). In Kantian terms, this is the difference between acting on the categorical imperative and the hypothetical imperative.

4.4. Institutionalising the social contract

Systems of public governance can be seen as devices to coordinate the selection of social contract equilibriums. The social group's balance of power is transcribed through the selection of governance structure and is thus decisive in determining the distribution of payoffs. In groups where the balance of power is decentralised, the ability to externally enforce equilibrium solutions is limited and thus they tend to organise on a solution that is fair in the egalitarian sense. In groups where the power is consolidated with elites, or a tyrant, the selection of an equilibrium can be sustained irrespective of fairness concerns, as long as it remains rational not to oppose those in power due to the risk of punishment.

Democracy, seen in this context, is a system of governance that seeks to decentralise power by giving every member of the group an equal worth when it comes to coordinating on social contract equilibriums.

The fundamental difference between democracy and other systems of governance rests in the ultimate normative reason for action within the system. Democracy must rely on an equitable system of payoff distribution because of its inherited decentralised power structure. If the system ignores the egalitarian *a priori*, it risks destabilisation or an authoritarian mutation. Authoritarian systems of government do not need such an *a priori*, because through consolidation of power and enforcement capabilities they are able to unilaterally dictate actions.

Democracy is thus the practical application of the abstract notion of the egalitarian social contract. The institution of democracy is in essence about avoiding consolidation of power. In that way, democracy is the operationalization and the deliberate maintenance of a spontaneous equilibrium selection device. It is an institution established to maintain and foster the moral egalitarian attitude in interpersonal interaction.

The idea of the social contract explains on how democracy is possible without ultimate reliance on coercive authority. The game theoretical approach of Binmore shows that the social contract is possible without reliance on Kantian ethics as its ultimate justification. In a world of rational beings, the roots of democracy lay in simple social and coordination practices that are formed and maintained based on the rational decision to act cooperatively. The decision to act cooperatively implies compliance with a specific method for distributing the potential payoffs that the cooperation ripens. The specific distribution method necessarily prescribes relatively equal shares to all equity holders in the relevant social or coordination practice. Complex social organs that are organised based on democracy are thus ultimately striving for equal distribution of the efficiency gain that the collective organ enables. Translated into the language of moral philosophy, this is the ultimate moral obligation of the organ and its members.

5. Law's dark matter

Social contract theory has shown us that simple social and coordination practices can be created based on different equilibrium solutions and that the specific solution that is practiced

depends on the attitude with which the participants approach the interaction. The problem for legal positivism built on Hart's premises, is that social practices are necessarily contingent facts. The contingency on which they rely is the attitude with which the practice was created and is maintained. A *rule of recognition* that is based on contingent social facts, such as social practices that are the result of a certain type of social interaction, are dependent on the same contingency as its foundational premises.

Hart's social practices are simply coordination equilibriums aimed at efficiency enhancements. Mysterious social practices that appear randomly or simply exist for no *a priori* reason do not exist in a world that assumes the existence of *free will*. Social practices and conventions are the output of the collision of rational human beings, which are playing the game of life. As we have seen, the function of a social practice depends on the attitude with which it is maintained. Social practices can be maintained based on the ultimate authority of the one with the greatest capacity to force obedience through the threat of violence. Social practises can also be maintained through cooperative attitude based on an egalitarian moral *a priori*. Seen in this way, the *rule of recognition* is simply a stable equilibrium solution to a certain coordination problem within a certain system of governance. The system of governance dictates how equilibriums within that system are selected. Authoritarian systems have a greater choice over the selection, due to their ability and willingness to apply force to maintain stability. Democratic systems have decentralised power structure and must thus select equilibriums that can be maintained with minimal enforcement effort.

As we have seen, the only efficient equilibriums that can be maintained without enforcement are those who respect the *a priori* of egalitarian fairness. We can thus see that if the *rule of recognition* is the product of a social practice within a democratic system of governance, the moral *a priori* of egalitarian fairness ultimately binds it. If, however the *rule of recognition* is the product of a system of governance that does not rely on the egalitarian *a priori* as the ultimate source of authority, the necessary embeddedness of this particular *a priori* in the governance system is no longer required.

Legal positivism is defined by the refusal to recognise a necessary connection between law and moral norms. This thesis holds in marginal cases like that of the laws of Nazi Germany, where the organ that produced the laws was not morally conditioned. However, when we talk about laws in the context of organs that are organised on the principle of democracy, the separation thesis of legal positivism does not survive. Laws that are created by an organ that maintains stability due to a specific moral attitude of its members cannot create an output that contradicts the principal reason for the organs stable being. A *rule of recognition* and derivative laws that contradict the principle of egalitarian fairness cannot survive in a democratic system of governance, because they do not form a stable equilibrium solution to the coordination problem they are meant to solve. Anomalies can appear, but they will always be gradually corrected or ironed out to comply with the cooperative attitude. There is thus a necessary connection between laws and the moral principle of egalitarian fairness in every political organ that builds on the principle of democracy.

The task of this chapter was to explore whether fairness in the laws was subject to a deliberate decision, or whether fairness in the laws was determined by an *a priori* that independently controlled whether the laws are fair or not. The traditional approaches to legal philosophy, under the headings of legal positivism and natural law theory, cannot answer this question because both theories assume the answer in their foundational premises. Legal positivism assumes that fairness in the laws is decided by the laws, while natural law theory assumes morality as an *a priori* which decides what is legally fair.

By exploring the concept of the social contract, in the context of political philosophy, it is possible to get beyond the primary assumptions made in the mainstream theories of legal philosophy. Social contract theory shows that social interaction, in a world that assumes free will, is sensitive to the attitude with which it is approached. Egoistic attitude dissolves spontaneous cooperation, while cooperative attitude maintains it. Social contract theory also shows that egoistic cooperation can be maintained through coercion.

Social contract theory shows that the nature of the concept of law is sensitive towards the foundational attitude with which the political organisation that the laws serve was founded

upon. If the laws serve a political organ that is founded upon a cooperative social contract (i.e. democracy), the ultimate *a priori* of the laws is the egalitarian fairness principle on which the political organ was founded. If the laws serve a political organ that permits egoistic social cooperation, the ultimate *a priori* is a threat of violence, which accordingly provides the laws with its ultimate authority.

As we saw with Hart's failed attempt to escape the *command model*, positivist legal theory must choose an *a priori* if it does not want to commit the fallacy of inferring *an ought* from *an is*. In doing so, it has a choice of sliding into the egoistic or the cooperative social contract. The former choice implies coercive form of government, while the later implies democratic form of government. We can also see that natural law theory assumes democratic form of government. The insight from social contract theory also explains why the laws of Nazi Germany were possible, and how that concession fails to defeat natural law theory; natural law theory holds if the political organ is founded on a cooperative premise, the laws of Nazi Germany were however the product of a coercive political organ.

The task of this chapter was to challenge legal positivism's central thesis about the severability of legality and morality, and to establish a normative *a priori* on which a thesis of procedural fairness should be built. I have now shown that in the case of democratic system of governance the positivist thesis does not hold, and I have described a normative requirement, which the choice of democracy lays on every action within political organs that ascribes to it. This normative requirement, building on Binmore's work on the social contract, is a Pareto efficiency condition and a condition of egalitarian distribution of the efficiency gain the organ produces. This forms a primary moral obligation on the organ and its members; the moral obligation of egalitarian fairness. The obligation taints the organ's actions and it conditions its output. Egalitarian fairness is thus the dark matter of laws within the democratic state, and thus the object of fairness in its laws. The laws are fair if their substance abides to the moral *a priori* of egalitarian fairness.

Zooming out a bit, the laws are subject to a deterministic requirement that is established by the choice of a democratic system of governance. The laws in the positivistic sense are thus

not entirely subject to the free will of the legislator, he is bound by an initial choice that exists outside of the domain of the positive law. By this free will is however not entirely expelled; we still have the choice between establishing a cooperative or an egoistic social contract, or as Machiavelli noted in the early 1500s: one can become a ruler through the favour of the people, or through the coercion of them.

Having established the object of fairness in the laws of the democratic society, the next problem for the purposes of this research is to find a way to translate the abstract concept of egalitarian fairness, into actual legislative decision-making.

Tools for translating optimality into laws

The object of optimality in the laws can be assessed from differing viewpoints. A law's optimality can be considered based on their compliance with an *a priori* that dictates how the laws *ought to be*. A law's optimality can also be assessed based on the quality of its output. The former approach to optimality is deontological in character, while the latter is consequential in character. Laws can thus be considered optimal based on their compliance with what *a priori ought to be*, irrespective of the actual consequences;¹⁴³ or they can be considered optimal based on their actual consequences, irrespective of compliance with what *a priori ought to be*.¹⁴⁴

In the previous chapter a standard of optimality was developed, i.e. the standard of egalitarian fairness, which implies a Pareto efficiency condition and an egalitarian distribution condition. Thus, the laws are fair if they comply with the optimisation standard of egalitarian fairness. Having established this, the next problem becomes to find the appropriate methodology for assessing the law's compliance with this standard of optimality. The assessment can be approached based on a deontological methodology, or a consequentialist methodology.

1. Traditional tools for translating optimality

Before exploring specific methodological tools suitable for assessing the optimisation standard of egalitarian fairness, it is appropriate to critically analyse two common optimisation tools in modern legal methodology. One based on a consequentialist approach to the laws, and the other based on a deontological approach to the laws. The *law and economics* movement utilises the *cost benefit analysis* method, which centres on the economic term *welfare*, for approaching optimality in a rule. The traditional juridical methodology, based on a hierarchal approach to legislative rules, focuses on the balancing of competing moral claims implied by the rules at the top of the hierarchy. Under the traditional juridical approach, the *proportionality test* is widely used to formalise and optimise intuitive thinking about moral

¹⁴³ i.e., this was optimal because we did as we were supposed to, although the result could have been better.

¹⁴⁴ i.e., this was optimal because the result was perfect, although we arrived at it differently than we should have.

claims in the context of the laws, and thus to maximise the moral appropriateness of a given legal rule.

1.1. Optimising towards economic efficiency

The suggestion that *'things are worth doing if the benefits resulting from doing them outweigh their costs'* is at the heart of most economic approaches to law.¹⁴⁵ From this intuitively appealing starting point, a consensus on what to actually include, as costs and benefits in a particular instance, has yet to emerge. The primary analytical tool within the law and economics approach, is the so-called *cost benefit analysis* (CBA) that seeks, on its mainstream version, to monetize the variables that are accepted as costs or benefits resulting from a given action.¹⁴⁶

In standard welfare economics, costs and benefits refer to individual preferences. An individual benefits, if his preferences are fulfilled, while a cost is inflicted on him if he dislikes the consequences of an action. For analytical purposes, the reference point is the individual's actual preferences with its potential for misinformation and irrationality, as opposed to the preferences of the perfectly informed rational.¹⁴⁷ People are simply taken to like what they like, not for what they ought to like.

Rational response to incentives and quantification through price mechanisms are the key insights the discipline of economics provides into the domain of laws. Legislators can create different behavioural incentives depending on the legislative decision they choose to take. Microeconomics offer mathematically precise theories and tools to identify the value of

¹⁴⁵ See for example Amartya Sen, 'The Discipline of Cost-Benefit Analysis' (2000) 29 *The Journal of Legal Studies* 931, 934. See also Robert H. Frank 'Why is cost-benefit analysis so controversial?' (2000) 29 *The Journal of Legal Studies* 913, 913: *'The cost-benefit principle says we should install a guardrail on a dangerous stretch of mountain road if the dollar cost of doing so is less than the implicit dollar value of the injuries, deaths, and property damage thus pre-vented.'*

¹⁴⁶ See for example Matthew D. Adler and Eric A. Posner, *New Foundations of Cost-Benefit Analysis* (Harvard University Press 2006) 13. The monetization element of CBA is often also referred to as *'willingness to pay'*. See for example; Amartya Sen, 'The Discipline of Cost-Benefit Analysis' (2000) 29 *The Journal of Legal Studies* 931, 945.

¹⁴⁷ See Matthew D. Adler and Eric A. Posner, *New Foundations of Cost-Benefit Analysis* (Harvard University Press 2006) 12-13.

competing legislative options. Microeconomic quantification enables a more scientific legislative approach than the traditional intuitive based juridical method allows.¹⁴⁸

The law and economics approach is sometimes mistakenly taken as synonymous with the normative agenda of the Chicago movement in economics of the 1970s and 1980s. This has not the least been fuelled by the early pioneering work of Richard Posner in the field, which in line with the Chicago school of economics emphasised the role of wealth maximisation and the efficient use of resources as a normative objective of the legal process.¹⁴⁹ At a closer look, one will however see that law and economics, as an approach to legal studies, is nowadays methodologically diverse and capable of accommodating a range of normative claims about the laws.¹⁵⁰

Taking this evolution into account, the modern version of *cost benefit analysis* in the context of laws, can be described as a method of measurement that enables the identification of efficient policy options that fulfil a specific normative criterion.¹⁵¹ The normative criterion is sensitive towards the definition of what counts as an input. Be it an element of wealth, welfare, or something else. The CBA as a measurement method for analysing legislative options is suggested to have the following minimum features; the inputs should be quantifiable, the assessment is consequential, and both positive and negative inputs are considered.¹⁵²

¹⁴⁸ See Robert Cooter and Thomas Ulen, *Law and Economics* (6th edn, Pearson 2012) 3; Richard A. Posner, *Economic Analysis of Law* (7th edn, Aspen 2007) 4.

¹⁴⁹ Posner's normative stance has somewhat relaxed over the years. In the latest version of his widely-cited textbook on the subject, he acknowledges alternative objectives of social policy, but nonetheless emphasises the power of efficiency arguments such as; '*maximization of the value of output*', for normative policy purposes. See Richard A. Posner, *Economic Analysis of Law* (7th edn, Aspen 2007) 14 and 24. See also Ronald Dworkin's critique of the normative value of wealth as described in the early writings of Richard Posner; Ronald Dworkin, 'Is wealth a value?' (1980) 9 *The Journal of Legal Studies* 191.

¹⁵⁰ The behavioural and welfare economics branch of the law and economics movement has for example explored several advanced methods for quantifying the supposedly unquantifiable aspects of social policy. The inclusion of these aspects significantly contrasts the crude simplistic agenda of wealth maximization as advocated by the pioneers in the 1970s. For a recent example see Cass Sunstein, 'The Limits of Quantification' (2014) 102 *California Law Review* 1369.

¹⁵¹ Lewis Kornhauser suggests defending CBA as a decision procedure rather than as a moral criterion. This enables an argumentative defence for the use of CBA independently of the moral plausibility of the efficiency criterion it is used to advance in an instance. See Lewis A. Kornhauser, 'On Justifying Cost-Benefit Analysis' (2000) 29 *The Journal of Legal Studies* 1037, 1051-53.

¹⁵² Amartya Sen has suggested three '*foundational principles*' of Cost Benefit Analysis; '*explicit valuation*' that eliminates non-quantifiables as an input in policy analysis, '*consequential valuation*' that eliminates deontological arguments, and '*additive accounting*' that takes notice of both negative and positive inputs. See Amartya Sen, 'The Discipline of Cost-Benefit Analysis' (2000) 29 *The Journal of Legal Studies* 931, 935-39.

Monetisation is not included as a necessary component of a CBA although, as mentioned above, the orthodox approaches to law and economics usually utilise money as a quantifier for costs and benefits. A distinction needs to be made between what is being measured (i.e. individual preference fulfilment) and the medium of measurement (i.e. quantification through willingness to pay money).¹⁵³ Monetisation should be seen as a simplifying assumption; a particular level of preference fulfilment equates willingness to pay a particular amount of money. Willingness to pay, is however by no means a universal function of the level of preference fulfilment. Monetisation favours those with the material means to show greater willingness to pay.¹⁵⁴

The output of a cost benefit analysis is a claim of a specific level of efficiency, or in other words, a specific level of preference satisfaction. A policy should, or should not, be pursued based on the level of efficiency. The orthodox literature recognises the concept of Pareto-efficiency, and in practice the modified Kaldor-Hicks version, as the most commonly used benchmark level.¹⁵⁵ The Pareto standard implies that a policy action should be pursued if someone can be made better off as a result and no one worse off, or, on the Kaldor-Hicks version, that the total benefit can potentially be used to compensate the losers.

Even if the CBA is claimed as normatively neutral method of measurement, the plain act of measure still forces *a priori* definitions of two kinds: firstly, the medium of measurement requires a defined zero point; and secondly, the subject of measurement needs to be defined in terms of positives and negatives relative to the zero point. These *a priori* choices regarding the inputs and the medium of measurement can hardly be made without eventually regressing towards a moral criterion of what *ought to* count as good and bad in terms of public policy,

¹⁵³ Cass Sunstein lists several examples of quantifiable costs or benefits that is hard to translate into actual money in a recent article. See Cass Sunstein, 'The Limits of Quantification' (2014) 102 California Law Review 1369, 1382-85.

¹⁵⁴ This is a standard objection to the CBA method. See Robert H Frank, 'Why is Cost-Benefit Analysis So Controversial?' (2000) 29 The Journal of Legal Studies 913, 916. An explicit distinction between quantification as such, and monetisation can overcome the objection, but that leaves unsolved the problem of finding a superior alternative for interpersonal comparison of preference satisfaction (i.e. utilities).

¹⁵⁵ See Richard A. Posner, *Economic Analysis of Law* (7th edn, Aspen 2007) 14, 42-43; Robert Cooter and Thomas Ulen, *Law and Economics* (6th edn, Pearson 2012) 3.

and at which point the balance of the good and the bad stops being an overall negative and becomes an overall a positive.

The Kaldor-Hicks efficiency benchmark provides the orthodox CBA approaches with a zero point; an outcome from a cost benefit analysis is either a negative or a positive, based on the benchmark of Kaldor-Hicks efficiency. The legislator *ought to* promote legislative options that pass the test of this specific efficiency, and refrain from pursuing options that do not. This choice of a zero point has clear moral implications; Kaldor-Hicks efficiency is pursued due to its potential to achieve a specific standard of distributional fairness that *a priori* is considered to possess a desirable quality. An argument about the laws anchored in the Kaldor-Hicks standard of optimality, is thus always defeasible by an outcome that fails to reach the zero-point threshold of potential distributional fairness. Arguments on the orthodox CBA approaches are thus defeasible with reference to an *a priori* that is reflected in the choice of a zero-point for the efficiency test being used. The core moral claim of the CBA is thus reflected in the choice of a zero point of neutral consequences.

Once the orthodox CBA method is reduced to its basic moral claim, its principal weaknesses as an authoritative tool for legislative decision-making become evident. The normative force of the CBA only reaches as far as the strength of its foundational moral claim. As it happens, the Pareto/Kaldor-Hicks standards are normatively weak on the scale of moral appropriateness. In terms of distributional fairness, they fall into the category utilitarian approaches that are more concerned with total output, than with the individual's share in the output. Policy options with a potential to enhance distributional inequalities pass the Pareto test. The Kaldor-Hicks test fares worse; policy options benefiting some, at the expense of others are permissible, given that the cumulative gains exceed the cumulative losses. As an optimising tool for legislative purposes, the orthodox CBA method thus cannot be relied on in cases where distributional fairness matters.¹⁵⁶

¹⁵⁶ One can think of several strategies to overcome this problem without sacrificing the CBA. The condition of egalitarian distributional fairness could either be internalised into the CBA as an efficiency factor, or added as an external condition that is assessed separately from the CBA. Internalisation creates a problem of quantification; is it possible to monetize the importance of distributional equality? Externalisation demotes the CBA to an ancillary role; the decisive normative force would be attached to the external condition. For discussion on the moral foundations of the CBA and for an attempt to overcome the problem of CBA's moral appropriateness see

The CBA as a method for qualitatively ranking competing legislative choices has great potential in the abstract. In the orthodox practice, this potential is however hampered by three practical problems. One is the problem of quantifying things that are not easily quantifiable. In practice, this results in disproportionate weight being given to quantifiable factors, over the unquantifiable ones. A second problem is the simplifying assumption of monetisation that eschews the accuracy of the method, by neglecting the principle of decreasing marginal utility. A third problem is the method's weak normative premises (i.e. on its orthodox variants). Neglecting concerns over distributional fairness has impact on the moral appropriateness of the advice given by the CBA. In combination, these problems with the orthodox versions, severely reduces the feasibility of the CBA as a general methodology for accurately identifying optimal legislative options.

A consequentialist ethical view defines the CBA approach.¹⁵⁷ Actions are not inherently good or bad. Each action entails consequences that are weighted in terms of preferableness. Bad consequences can be disregarded, if other consequences are sufficiently good to outweigh them.¹⁵⁸ The CBA attempts to strike this balance based on a specific standard of efficiency as the tipping point between an action that should be undertaken, and an action that should be rejected. For the CBA, in its orthodox variants, the ultimate standard of optimality is either Pareto, or Kaldor-Hicks efficiency.

1.2. Optimising towards moral appropriateness

A deontological approach to ethics is the traditional rival of the consequentialist view. Deontological thinking about laws relies on predefined axiomatic duties that are assumed to

Matthew D. Adler and Eric A. Posner, *New Foundations of Cost-Benefit Analysis* (Harvard University Press 2006) 25-61 and in a more elaborate form; Matthew D Adler, *Well-Being and Fair Distribution: Beyond Cost Benefit Analysis* (Oxford University Press 2011).

¹⁵⁷ On this point see for example; Robert H Frank, 'Why is Cost-Benefit Analysis So Controversial?' (2000) 29 *The Journal of Legal Studies* 913, 929; Amartya Sen, 'The Discipline of Cost-Benefit Analysis' (2000) 29 *The Journal of Legal Studies* 931, 936-38; Lewis A. Kornhauser, 'On Justifying Cost-Benefit Analysis' (2000) 29 *The Journal of Legal Studies* 1037, 1056-57; Cass Sunstein, 'The Limits of Quantification' (2014) 102 *California Law Review* 1369, 1372.

¹⁵⁸ A standard objection to the consequentialist view exploits this point using the '*transplant*' example: Should a healthy individual be killed to save five others through the transplant of his organs? In purely utilitarian consequentialist terms, the answer seems to be affirmative. The intuition of most people would however consider such an act prohibited with reference to moral reasons.

contribute towards a desirable state of the world. One should act as if one's actions were to become a universal standard of behaviour, as Kant suggested, guided by the will to do something good.¹⁵⁹ The focus of deontology is on the action and the motive of the actor. If the action is motivated and executed based on the will to do one's duty and in accordance with one's duty, the action is morally plausible. Even if the consequences turn out to be undesirable in terms of overall preference satisfaction or on other utilitarian scales.

Deontological thinking about laws centres on axiomatic duties. Depending on allegiance to positivist or naturalist view on the concept of law, these duties either naturally exist, or can be positively invented. Either way, the result is a set of legal axioms that forms the most senior layer of a hierarchically structured legal system. In such systems, the ultimate legal argument is made based on axiomatic duties derived from the normative layer of a legal system. The functional rationale of the modern constitutional democracy is based this kind of thinking; constitutions of some sort outline the axiomatic duties that junior legal actions must comply with.¹⁶⁰

The central problem with deontological approaches to laws arises when separate axiomatic duties are at conflict. This problem is compounded by the growing list of constitutional duties, most significantly in the form of rights of various sorts. The plurality of axiomatic duties has gradually increased the complexity of the legislator's task in acting within the boundaries of permissible legislative actions.

The non-regressive moral argument is at the heart of deontological thinking. In the context of law, the argument comes in two main variants, which rely on the same circular argument: One should follow the laws because they are the laws, or; one should comply with one's moral duties because they are moral duties.¹⁶¹ In any case, the constitutional instruments usually

¹⁵⁹ Kant's formulation of the *categorical imperative* is perhaps the ultimate example of deontological thinking about ethics. See Immanuel Kant, *Groundwork for the Metaphysics of Morals* (first published 1785, Mary Gregor and Jens Timmermann eds, 2nd edn, Cambridge University Press 2012) 26-34.

¹⁶⁰ I use the term *constitution* broadly as a reference to any legal source that has a hierarchal status equivalent to (or greater than) a traditional written constitution.

¹⁶¹ Reference is made here to the positivism v. naturalism debate within legal philosophy. I caricature the most extreme positions within that debate, but it should be noted that most authors recognise a mix of both arguments in varying proportions.

describe the moral duties that subsequently become non-regressive moral-legal standards. The self-referential non-regressive nature of moral-legal standards has an important implication. Each standard, claims to be morally imperative and as a result it is legally imperative that all actions comply with the standard. Once a set of such standards exists that each lays an imperative condition on the legislator, collision is bound to happen. Within the deontological system of thought, such collisions are hard to reconcile due to the structure of the constitutional moral argument. How do we prioritize incompatible actions that each claim to be morally imperative?

Considering the difficulties of finding a solution within the deontological system, it is hardly surprising that the orthodox way of reconciling competing deontological claims is a tool based on consequential thinking.¹⁶² I am referring to the proportionality test,¹⁶³ which is used by legislators and judicial bodies in many jurisdictions to test the compliance of legislative actions with competing constitutional standards, usually of the moral kind.¹⁶⁴

The intricacies of the proportionality test are detailed and often case specific, but it does nonetheless possess an element of simplicity to which, undoubtedly, it owes a large part of its spreading success. On the face, it is a test of three (sometimes four) successive questions that each form a threshold that a contested legislative action needs to pass. Only after having been deemed i) suitable and ii) necessary towards a legitimate aim, the primary action is iii) balanced against a secondary consideration¹⁶⁵ that usually takes the form of a right that is being violated.

¹⁶² Note though, Ronald Dworkin's ambitious project of finding the legally optimal answer in deontological terms without slipping into consequential terminology. Although hugely influential on modern legal thought, he did not provide a comprehensive methodology for identifying the single optimal solution he argued for. For a good overview of his project see Ronald Dworkin, *Justice for Hedgehogs* (Harvard University Press 2011).

¹⁶³ Originating as a German principle of administrative law in the 1950s, the proportionality test has reached a status of orthodoxy within most European legal systems. Not the least through the supranational institutions and instruments of the EU and the ECHR. Robert Alexy's work on constitutional theory is standardly cited as an elegant theoretical account of the principle. In what follows, I will use Alexy's account as a main reference. See Robert Alexy, *A Theory of Constitutional Rights* (first published 1985, Julian Rivers tr, Oxford University Press 2002) 44-110.

¹⁶⁴ A.S. Sweet and J. Mathews claim that the use of the proportionality test in rights based adjudication has become so widespread that it possesses the defining features of a global constitutional principle. See Alec Stone Sweet and Jud Mathews, 'Proportionality Balancing and Global Constitutionalism' (2008) 47 *Columbia Journal of Transnational Law* 72, 74.

¹⁶⁵ The third step is sometimes referred to as proportionality in the strict or narrower sense. See Robert Alexy, 'Constitutional Rights and Proportionality' (2014) 22 *Revus* 51, 52.

In Robert Alexy's renowned version of the proportionality test, the move from the territory of the deontological over to the territory of the consequential is made based on a distinction between rules that are binary in character, and principles that are 'optimization requirements'.¹⁶⁶ He argues that constitutional rights are principles that necessarily have to be optimized in terms of the factual and the legal reality. The legal reality includes competing constitutional principles. For Alexy the act of optimization and balancing is an integral part of applying constitutional principles.¹⁶⁷

Implicit in Alexy's move is a departure from a model of thinking that considers specific actions morally imperative, to a model of thinking that considers moral imperatives contingent on specific consequences. A morally imperative action is only permissible if the consequences are optimal with regards to the legal and factual reality. The focus moves from the action per se, over to the consequences of the action.

By framing the proportionality test in consequential terms, Alexy implicitly acknowledges the transition from the deontological. He argues that the suitability test and the necessity test are in fact Pareto requirements that aim at avoiding unnecessary costs with reference to the factual reality: Suitability requires Pareto efficiency, and necessity requires Pareto optimality. The final test of proportionality in the narrow sense concerns the legal reality rather than the factual. When legal principles collide, costs cannot be avoided, which requires a decision on how to distribute the costs between the competing concerns. The actual balancing exercise is supposed solve this decisional problem based on the 'law of balancing' that uses the quantitative terms; 'degrees' and 'importance' of 'satisfaction' as an object of balancing.¹⁶⁸ For

¹⁶⁶ Alexy's distinction between the rule as a '*definitive ought*' and principles as a '*prima face ought*' draws on Ronald Dworkin's well-known distinction between rules and principles, with a slight variation though. See Ronald Dworkin, *Taking Rights Seriously* (Duckworth 1977) 24. See also Matthias Jestaedt, 'The Doctrine of Balancing – Strengths and Weaknesses' in Matthias Klatt (ed), *Institutionalized Reason: The Jurisprudence of Robert Alexy* (Oxford University Press 2012) 153-54.

¹⁶⁷ See Robert Alexy, 'Constitutional Rights and Proportionality' (2014) 22 *Revus* 51, 52.

¹⁶⁸ Alexy formulates the '*law of balancing*' in the following way; '*The greater the degree of non-satisfaction of, or detriment to, one principle, the greater must be the importance of satisfying the other.*' See Robert Alexy, *A Theory of Constitutional Rights* (first published 1985, Julian Rivers tr, Oxford University Press 2002) 102.

the purposes of formalising the practical application of the law of balancing, Alexy designed the *weight formula*.¹⁶⁹

The *weight formula* simply multiplies three variables concerning each of the two conflicting principles. The variables are: abstract importance; probability of gain/infringement; and intensity of gain/infringement.¹⁷⁰ Once a number has been calculated for both principles, the numeric value of the primary principle is divided with the numeric value of the secondary principle.¹⁷¹ If the result (x) is equal to or greater than one ($x \geq 1$) the disputed action is allowed, but if it is lesser than one ($x < 1$) the action is disproportionate.¹⁷²

Although not explicitly accredited, Alexy's formulation of the third step of the proportionality test slides into a version of the Kaldor-Hicks efficiency test; an action is allowed so long as the primary benefit is equal to, or higher, than the costs incurred. During the final step, the proportionality test relaxes the Pareto standard used in the two previous steps for identifying the optimal primary action, and allows the execution of the action if it achieves at least Kaldor-Hicks efficiency with regards to the losses incurred on a colliding principle.

We can formalise Alexy's representation of the proportionality test in the following way where action (a) has consequences (+) and (-) for the primary interest (P) and the secondary consideration (S):

¹⁶⁹ See Robert Alexy, 'Constitutional Rights and Proportionality' (2014) 22 *Revus* 51, 52-57.

¹⁷⁰ The real x factor is the *level of gain/infringement* factor. The first factor would need to use formal importance, which usually would mean that colliding constitutional principles of the same hierarchical seniority would have the same abstract value. The second factor is tricky; it might be hard to establish objective probabilities, and subjective probabilities would open a door for arbitrariness.

¹⁷¹ Alexy argues that a legal argument can't be ranked on a cardinal scale, but concedes that it might be possible to rank such arguments on a simple ordinal scale. See Robert Alexy, 'Constitutional Rights and Proportionality' (2014) 22 *Revus* 51, 55. Giovanni Sartor has shown elegantly how Alexy's proportionality test can be quantitatively formalised using non-numerical magnitudes. See Giovanni Sartor, 'The Logic of Proportionality: Reasoning with Non-Numerical Magnitudes' (2013) 14 *German Law Journal* 1419.

¹⁷² See Robert Alexy, 'Constitutional Rights and Proportionality' (2014) 22 *Revus* 51, 54-55. See also Robert Alexy, 'The Weight Formula', in Jerzy Stelmach, Bartosz Brożek, Wojciech Załuski (eds), *Studies in the Philosophy of Law - Frontiers of the Economic Analysis of Law* (Jagiellonian University Press 2007) 9-27; Robert Alexy, 'On Balancing and Subsumption – A Structural Comparison' (2003) 16 *Ratio Juris* 433.

Consequentialist representation of Alexy's proportionality test

Step 1: Suitability	Consequences of (a) must benefit (P)	Pareto efficiency
Step 2: Necessity	Consequences of (a) must benefit (P) more than other available actions	Pareto optimality
Step 3: Narrow Proportionality	The sum of (+) consequences to (P) and (-) consequences to (S) from (a) must ≥ 0	Kaldor-Hicks efficiency

I identify two main problems with the proportionality test for optimising legislative decisions:

Firstly, the results of the proportionality test are largely decided by the framing of the primary objective, not so much by the balancing of colliding objectives. Once an action has been proposed that is suitable and necessary towards an objective (i.e. Pareto optimal action) it will prevail unless a secondary consideration defeats it. Since the action is already designed to be Pareto optimal with regards to the primary objective, the violation to the secondary consideration needs to exceed the high threshold of a Pareto optimal gain. If the primary gain is (for sure) exceptionally high, the negative consequences can justifiably be severe, without failing the test of proportionality. If, however, the roles of the two interests would be reversed during the initial framing, it would be proportional to take an action that would have the exact opposite and potentially drastically different consequences.

The problem is that the proportionality test does little to balance the satisfaction of competing interest; it just makes sure that the benefit from promoting one interest, does not exceed the harm done to another interest. The problem compounds when we have situations where competing moral considerations have different consequences for different groups of people. In such situations, an action can potentially pass the proportionality test while diverting all the benefits to a limited group of people, and all the costs to another group. Essentially, this is the same distributional justice objection, which the Pareto and Kaldor-Hicks standards have difficulties overcoming under the economic approach to law.

Secondly, the proportionality test may hold certain advantages over other deontological approaches in terms of accuracy and predictability, and in providing a plausible narrative

reason for violating moral imperatives. But if the proportionality tests were to be taken for what it really is, as just another consequentialist approach, it would rank as quantitatively and qualitatively imprecise. The quantitative methods used to rank the competing considerations are intuitive based and subjective to individual assessors. There is usually no way to retrospectively verify the accuracy or correctness of the balancing exercise. It is simply a subjective opinion of a specific decision maker. The qualitative selection of inputs is also restricted to a set of consequences that have previously been recognised as relevant (usually morally relevant) by the legislator. This risk negligence of other important consequences and invites the overrepresentation of special reasons coined as moral imperatives.

As a final consideration, we may ask why the seemingly deontological approach of the proportionality test slides into the territory of the consequential, and whether that is inevitable for all deontological approaches to legislative optimisation. To answer this question, we can think about Kant's categorical imperative as a singular moral duty that encompasses everything that one morally *ought to do*. Such a singularity concept, that at the same time is universally applicable, is hardly imaginable except as a composite concept, that is composed of many context specific duties, that each provides a direction on how to act in a specific context in order to comply with the singular moral duty.¹⁷³ The existence of a composite concept, creates a potential of a clash between different context specific duties, if the context is sometimes dynamic. To resolve such clashes, a compromise must be made between competing duties that both claim to be imperative in their respective contexts. This compromise cannot be made without reference to the ensuing consequences for the superior categorical imperative, and thus the deontological approach must slide into the territory of consequential to achieve optimisation of the impact on the categorical imperative.

1.3. The benchmark for possible methodological improvements

The chapter started with a proclamation that optimality in legislation could be approached based on a deontological methodology, or a consequentialist methodology. After having reviewed orthodox variants of each methodological typology, it becomes apparent that despite

¹⁷³ If moral appropriateness is context specific, we would need to adjust our behaviour according to circumstances. If moral appropriateness is not context specific, we can irrespective of circumstances behave in the single correct manner.

claims that the cost benefit analysis is normatively neutral, it cannot escape relying on *a priori* normative choices. The opposite is true with the proportionality test; despite relying on deontological rhetoric of imperative rights and duties, it is a quantitative test of the consequential type, that uses Kaldor-Hicks efficiency as a normative benchmark. We can thus see that for approaching optimality in legislation, the orthodox approaches rely on a normative *a priori* of the deontological type, and quantitative tools of the consequentialist type to accurately balance potentially competing normative claims.

The orthodox methods that I have reviewed have their strengths and weaknesses in giving advice on legislative optimisation. The economic approach is good for finding legislative options that promote economic efficiency measured in effects that can be monetised, but less good for finding morally appropriate options. The proportionality test is good at promoting important moral values in legislation, but less good in identifying other important efficiency considerations. Both approaches are simple and practical in application, at the cost of quantitative accuracy, and both fail to identify legislative options that promote distributional fairness in the egalitarian sense. The problem of inaccurate quantification concerns reliance on intuitive approximation in a proportionality test, and the use of monetisation as a proxy in a cost benefit analysis.

The two types of methodologies for approaching optimality in legislation, described above, are thus not suitable for the purposes of identifying the optimal equilibrium point of egalitarian fairness in a legal rule, which is the task of this research. Additionally, the quantitative tools used for the described methodologies suffer from inaccuracies, which alternative tools might overcome. However, for general purposes these orthodox methodologies are the benchmark which alternative methodologies are naturally compared with.

2. Alternative tools for translating optimality

The methodological merger of the deontological and the consequential in optimising legislative decisions implies two things: firstly, the need for a normative *a priori* which becomes the locus of optimisation; and secondly, an object of measurement for the chosen quantitative medium of measurement. For the purposes of the main research question, the former has already been

defined as egalitarian fairness in a previous chapter, a definition of the latter is however still underdeveloped. In other words, we know in the abstract what constitutes ultimate optimality in terms of legislative fairness, but we are yet to articulate precisely what needs to be measured to determine the level of optimisation.

We have already seen that consequentialist approaches to optimality cannot escape making *a priori* normative choices regarding the medium of measurement; what counts as a positive, a negative, and what constitutes cumulative neutrality of positives and negatives. We have also seen that deontological approaches, dealing with hierarchically parallel imperatives, cannot escape relying on a quantitative methodology for resolving conflicts that in some way quantifies the consequences of the competing moral claims.

The object of measurement is thus also the object of the consequences. A legislative decision to design a rule in a specific way, mediates a conflict between competing stakeholders in a rule. By taking a legislative decision, the legislator decides in which proportions the competing claims should be realised. The legislator can accept a claim or reject a claim, or partially accept or reject a claim. The consequence of such a decision is subsequently liable to become the object of preference for the stakeholders involved; some may find the result preferable, while some may not, both of which can feel so to varying degrees. In social orders that are based on the idea of democracy, this feeling of preferableness towards consequences of any kind, forms the nucleus object of optimisation; a state of a fair egalitarian equilibrium needs to be reached between the preferences of the members of the relevant social contract. The object of measurement is thus the preferences of those concerned with a legislative decision. In other words, the consequence of a legislative decision is an impact on the preference fulfilment of those concerned, and the degree of this impact needs to be measured and subsequently balanced and optimised based on the *a priori* of egalitarian fairness.

As explained in previous sections, cost benefit analysis and the proportionality test are commonly used for balancing between competing preferences in legislation. The CBA methodology on its orthodox variant uses monetisation as a proxy for preferences, which can cause inaccuracies and biases of various sorts. The proportionality test is not very robust as a

quantitative methodology due to reliance on intuitive quantification that at best is a rough approximation of the underlying preferences at stake, and at worst a subjective opinion of a specific assessor. The CBA methodology is good at optimising where moral issues are not important, while the proportionality test is better at optimising where pure economic efficiency is not the main issue.

The ideal methodology for approaching optimality in legislation would overcome the quantification problems that restrain the orthodox approaches. Additionally, it would be applicable in circumstances where the object of preference is characterised as of a moral nature, in circumstances where the object of preference concerns economic factors, and in circumstances where the object of preference is a mix of moral and economic reasons. To achieve this, we would need quantitative tools that can measure and compare differing degrees of preferences, and we would need a medium of measurement that can translate inputs based both on moral and economic rationale, and which would provide a uniform output that does not distinguish between the two.

Before trying at assembling an improved methodology for optimising fairness in legislation, an introduction is warranted on the key quantitative tools and theoretical assumptions that inspire its construction. The initial assumption I want to make is to think of legislative problems, as simple decision problems. If we accept this assumption, it is possible to think of the act of legislating as a decision with posterior effects that strives for a result that satisfies *a priori* normative conditions. The orthodox modern method of modelling decisions with reference to preferences, is based on Bayesian logic and on the work of the game theorists Von Neumann and Morgenstern.¹⁷⁴ It goes by the name *expected utility theory*. An influential alternative theory, which can be viewed as complementary, is based on the work of Kahneman and Tversky¹⁷⁵ and is referred to as *prospect theory*. Sometimes the study of these kind of issues is collectively referred to as *decision theory* or *choice theory*.

¹⁷⁴ See John von Neumann and Oskar Morgenstern, *The Theory of Games and Economic Behavior* (first published 1944, 3rd edn, Oxford University Press 1954).

¹⁷⁵ See Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263; Daniel Kahneman and Amos Tversky, 'Advances in Prospect Theory: Cumulative Representation of Uncertainty' (1992) 5 *Journal of Risk and Uncertainty* 297.

2.1. The basics of decision theory

In a simplified way, decision theory is about analysing the suitability of choices towards achieving a defined agenda. The theory implies multiple choices, or else the decision at stake would be taken by default without the need for any analysis. There must also be an agenda, towards which the decision maker seeks to identify the optimal choice. A decision to pursue a specific choice, out of the set of available choices, which best optimises the agenda, is an optimal decision.¹⁷⁶

Sometimes the identification of the optimal decision is an easy process that hardly warrants a second thought. If for example, two apples need to be distributed equally between two people, most decision makers would identify the optimal choice quickly. The available decision choices are three: both apples can be given to person A, both apples to person B, or one apple to each person. Since the *a priori* requirement was equal distribution of apples, the last choice is obviously optimal. Unfortunately, decision problems can be much more difficult to solve. For example, the outcome of each choice might be uncertain, and the normative agenda might represent a complex set of conditions. At times, it can pose a considerable challenge to identify the optimal decision. Decision theory provides tools to dismantle and solve such difficult problems.

Various approaches and versions of decision theory exist. The orthodox version claims normative neutrality with regards to the substance of output of the decisional mechanism, but inevitably concedes compliance to assumptions and rationality axioms integral to the very process of deciding.¹⁷⁷ The model of deciding, judges the substantive output by its logical validity in accordance with axioms of propositional, predicate and Bayesian logic. The normative quality of the output of the decisional mechanism is thus not decided by the

¹⁷⁶ For a non-technical explanation of the basics of decision theory see Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) <<http://home.abe.kth.se/~soh/decisiontheory.pdf>> accessed 29 December 2016. See also a more mathematically sophisticated, but still rather accessible introduction to decision theory in; Nolan McCarty and Adam Meirowitz, *Political Game Theory: An Introduction* (Cambridge University Press 2014); Kenneth Binmore, *Rational Decisions* (Princeton University Press 2011); and Herbert Gintis, *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences* (Princeton University Press 2008) 2-30.

¹⁷⁷ See Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 7.

mechanism as such, but instead by the normative *a priori* that exists independently of the process of finding the optimal choice.

The basic approach to decisions problems is to compare the available options and rank them in order of preference. The most common approach is to either use *numerical* or *relational* ranking to express the order. Option A is better than B, and B is better than C, is an example of *relational* ranking. Option A has a value of 9, B has a value of 4, and C has a value of 2, on the scale between good and bad is an example of *numerical* preference ranking.¹⁷⁸ In formal language of preference logic, the value relation between different options is usually described by one of three mathematical notions: 'Better than ($>$)' signals a *strong preference*, 'equal in value to (\equiv)' signals *indifference*, and 'at least as good as (\geq)' signals a *weak preference*.¹⁷⁹

The standard approach within decision theory is to make few important assumptions with regards to the comparison of preferences. The term *preference* usually refers to either *revealed preferences* or *stated preferences* in the literature on economics. Revealed preferences use past actions as an indication of present preferences. Stated preferences are based on answers to surveys or questioners that usually register answers to hypothetical questions that are used as an indication of present preferences. For the purposes of decision theory, revealed preferences are usually considered superior, if at all available.¹⁸⁰ The use of revealed or stated preferences as a proxy for actual current preferences, requires the acceptance of several important assumptions. These assumptions are meant to translate into the language of logic the notion that choices in the past will be consistently chosen again in the present.¹⁸¹ The fundamental logic behind choice consistency is based on the assumptions of *completeness*,

¹⁷⁸ For a useful introduction to preference logic see Sven Ove Hansson and Till Grüne-Yanoff, 'Preferences' in Edward N. Zalta (ed), *The Stanford Encyclopaedia of Philosophy* (Winter 2012) <<http://plato.stanford.edu/archives/win2012/entries/preferences/>> accessed 29 December 2016.

¹⁷⁹ See Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 15.

¹⁸⁰ Binmore argues that the method of revealed preferences is superior to the method of stated preferences due to their descriptiveness and their normative neutrality; no presumptions are made about the goodness or the badness of revealed preferences, they simply are what they have been shown to be in the past. See Kenneth Binmore, *Rational Decisions* (Princeton University Press 2011) 7-12.

¹⁸¹ Again, this fundamental assumption is of course not universally true, but it does provide a proximate anchor for the analytical study of preferences, which is superior to other approaches. See Herbert Gintis, *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences* (Princeton University Press 2008) 3.

*transitivity, and independence of irrelevant alternatives.*¹⁸² These assumptions assume that the agent making the choices has rational preferences.¹⁸³

Preference relations are said to be *complete* when the preference relations between each of the preference options, that exist within a specific domain, are defined by any of the standard value notions. It should thus be clear whether there is a strong preference ($>$), a weak preference (\geq), or an indifference (\equiv) relation between each of the domain's available options. In mathematical notation¹⁸⁴ this translates into; $A > B \vee A \equiv B \vee B > A$ must be true for all preference options A and B within a specific domain.¹⁸⁵ Often the assumption of completeness fails to hold in real life circumstances, i.e. the relation between preferences is often not complete. If, for example, you go out shopping for a new sweater and know beforehand that your favourite colour is black and that you would thus prefer a black sweater to green or blue. If you do not have a formed opinion on whether green or blue would be a better secondary option, your preference relation with regards to these three sweaters is incomplete. In a real-life situation that would however not matter much since you would just buy the black one and that would be the end of it. Actual decisions are thus often made without a complete set of preference relations. In decision theory, it is however important to make the simplifying assumption of completeness to achieve choice consistency.

Transitivity of preferences means that the following preference relation is assumed to be true: if I prefer *apples (A)* to *oranges (O)* and *oranges* to *melons (M)*, then I also prefer *apples* to *melons*. In mathematical notation, this translates: if $A > O \wedge O > M \rightarrow A > M$. The same assumption is usually made with regards to indifference and weak preference relations. Intransitive preferences are problematic for choice consistency. If we would change the example above

¹⁸² See Herbert Gintis, *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences* (Princeton University Press 2008) 7.

¹⁸³ McCarty and Meirowitz define minimum, or thinly rational behaviour based on the assumptions of completeness and transitivity. See Nolan McCarty and Adam Meirowitz, *Political Game Theory: An Introduction* (Cambridge University Press 2014) 6.

¹⁸⁴ We use the mathematical term \vee for 'or', the term \wedge for 'and', and the term \rightarrow for 'then'.

¹⁸⁵ To simplify, it is also possible to only use the weak preference notion to describe all possible preference relations between two options. Then we would say that a relation between preference options x and y within a specific domain is complete, if and only if (iff) either Option x \geq Option y, or Option y \geq Option x. This simplification eliminates the need for the strong preference notion and the indifference notion to describe preference rankings. See further Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 15-16.

and say that *melons* are preferred to *apples*, we have a cyclical preference relation, which undermines consistent rational decision making by confusing the priority of preferences. This can be highlighted by the following preference relation which makes it hard to identify the fruit of primary preference: $A > O \wedge O > M \wedge M > A$. The transitivity assumption has been shown to be unrealistic in real life situations, but it still has a salient intuitive appeal, especially when the preferences have been reduced to comparable numbers. It would seem irrational that a preference with a specific numerical value could at the same time be both more and less valuable than another preference, which is expressed on the same numerical scale.¹⁸⁶ The transitivity of indifference and weak preferences is additionally subject to the Sorites Paradox, which for our purposes is only a hypothetical problem.¹⁸⁷ Despite these complications, it is nonetheless standard practice to assume transitivity when modelling decision problems within decision theory.¹⁸⁸

Preferences are *independent of irrelevant alternatives* (IIA) when their attractiveness is not influenced by preferences that are external to the relevant domain of options.¹⁸⁹ If deliberating about whether to have coffee or tea with one's breakfast, the weighing of these two options is independent of the fact that one highly values a brand of Belgian chocolate, which one sometimes buys on his occasional trips to Brussels. If, however one happens to have a box of that chocolate in the cupboard of his Florence apartment, the alternative preference might suddenly not seem as irrelevant for the preference relation of coffee and tea. The choice between coffee and tea, which often one would have indifferent preference function towards (\equiv), would in that case be influenced by the prospect of enjoying a piece of his favourite

¹⁸⁶ If we give utility values to the choices in the fruit example the intransitive formula could look like this:

$$A(10) > O(8) \wedge O(8) > M(5) \wedge M(5) > A(10)$$

The last bit (five utiles are better than ten utiles) seems irrational which partly explains the intuitive appeal of transitivity. In contrast the transitive formula appears rational since ten is better than both eight and five: $A(10) > O(8) \wedge O(8) > M(5) \rightarrow A(10) > M(5)$

¹⁸⁷ The Sorites Paradox explains that transitivity does not hold for indifference preference relations in large sets of options where tiny changes are made between each option. Each tiny change seems irrelevant, but when compiled in large sets they become significant. Because transitivity does not hold for indifference in such circumstances, it does not either hold for weak preferences because it incorporates the indifference notion. See further in Sven Ove Hansson and Till Grüne-Yanoff, 'Preferences' in Edward N. Zalta (ed), *The Stanford Encyclopaedia of Philosophy* (Winter 2012).

¹⁸⁸ See Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 17-19.

¹⁸⁹ See Herbert Gintis, *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences* (Princeton University Press 2008) 6-7. On IIA see also Kenneth Binmore, *Rational Decisions* (Princeton University Press 2011) 7-12.

chocolate with the choice he makes. As it happens, one finds the combination of Belgian chocolate and Italian coffee to be superior ($>$) to the combination of chocolate and tea. According to the IIA axiom, the relation (\equiv) between coffee and tea must not change with the introduction of the chocolate alternative. To accurately reflect the chocolate alternative, one must thus include the two combinations as additional relevant preferences within the option domain of the breakfast table. Having done that, one could continue to assume that the preference domain is independent of irrelevant alternatives. After the expansion of the choice domain the preference relation could be expressed in the following way: Coffee-Chocolate $>$ Tea-Chocolate $>$ Coffee \equiv Tea. The IIA axiom for coffee and tea holds since the preference relation remains at indifference. Abandoning the IIA axiom undermines the consistency requirement by making the preference relation between two options depend on infinite possibilities of other alternatives.

To summarise we can say that a preference relation is consistent when it is *complete, transitive, and independent of irrelevant alternatives*.¹⁹⁰ Once we have assumed consistency in preference relations, we can assume that choices in the past will be chosen again in the present, and thus the theory of revealed preferences (or alternatively stated preferences) can be used as a basis for modelling decision problems. Decisions that we would prefer to take in the present, should reflect our revealed past preferences (or alternatively our state preferences).

In many contexts, it is helpful to reduce preferences to numbers. The standard approach is to assign each preference option a certain numerical value, referred to as *utility*. Utility function is a simple mathematical device. A certain numerical value (utility) is simply assigned to each of the preference options signalling their relative value within the domain of options. The best option could, as an example, be assigned the utility 1 and the worst option the utility 0. The options in between should then be assigned a utility value between 1 and 0 depending on their value ranking within the relevant domain.¹⁹¹

¹⁹⁰ See Herbert Gintis, *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences* (Princeton University Press 2008) 7.

¹⁹¹ See Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 21-22.

The point of rational decision-making is to maximise *something*. This *something* can be anything. Once we have defined what *something* ought to be, we can rank the available options on a utility scale based on how bad or well each option contributes towards the ideal state where *something* is at its maximum. The option with the highest utility ranking contributes most towards the preferred state and is thus the *optimal decision* given the available options. Utility maximisation is the orthodox approach within decision theory.¹⁹²

2.2. Risk, uncertainty, and the decision matrix

Identification of the optimal decision based on revealed or stated preferences and a specific normative agenda, can at times be challenging in a dimension where consequences of decisions are known with certainty. Often however, there is no certainty about the real effects of decisions. The umbrella dilemma is a standard example to explain this: If I walk to work every day, I must decide whether to take an umbrella with me or not. It is a bit cumbersome to carry the umbrella around, but it is even worse to get soaking wet by a sudden shower on my way home in the afternoon. I do thus have to decide whether to take the umbrella, or not, without knowing whether I will need to use it. In this case, the optimal decision depends on the uncertain factor of whether it will rain or not. If it rains, it would be optimal to have the umbrella, but if it remains dry, it would be optimal to have no umbrella to carry. The dilemma highlights a decision problem that exists in almost any decision, the problem with uncertainty.

Most decisions are susceptible to varying degrees of uncertainty about external factors that have the potential to alter the outcome of a given action. A contemplation about the optimal decision must thus give notice to the varying externalities that may affect the outcome of a specific choice. These varying externalities are usually referred to as different *states of the world*.¹⁹³ In the example above, rain in the afternoon would be one state of the world, and no rain would be another state of the world. Each choice has potentially different outcome in each state of the world.

¹⁹² See Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 22. See also Kenneth Binmore, *Rational Decisions* (Princeton University Press 2011) 14-16, for a mathematical formulation of utility assignment and the utility maximisation axiom.

¹⁹³ Some authors also use *states of nature* for the same purposes. See Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 24-25.

The usual method for dealing with uncertainty within decision theory is with some version of Bayesian probability calculations. Before entering Bayesian probabilities, it is useful to clarify a certain terminology used in decision theory to distinguish between different levels of knowledge about the consequences of a given action. Decision under *certainty* indicates complete knowledge of the consequences of an action. Decision under *risk* refers to complete knowledge of the probabilistic outcome of a given action in the relevant states of the world. Decision under *uncertainty* means that the probabilistic outcome of an action is partially known. Decision under *ignorance* signals the least knowledge of the outcome of an action, the term is reserved for instances where the unknown is unknown.¹⁹⁴ The table below summarises the terminology:

Decision Type:	Level of knowledge about results:
<i>Certainty</i>	Everything is known
<i>Risk</i>	The unknown is known
<i>Uncertainty</i>	The unknown is partially known
<i>Ignorance</i>	The unknown is unknown

Decision problems are usually modelled into *decision matrixes*, which are simple tables where the decision options are listed in rows and the different states of the world listed in corresponding columns. The cells at the intersection of each decision alternative and the different states of the world should then contain the *consequences* of a specific *action* in a specific *state of the world*.¹⁹⁵ The decision matrix for the umbrella example could be represented in the following manner, with the *consequences* represented as utility values based on preference for being dry and not having to carry extra luggage:

¹⁹⁴ There are several ways to represent these levels of certainty during decision-making. The version described here is based on; Duncan R. Luce and Howard Raiffa, *Games and decisions: Introduction and critical survey* (Courier Corporation 1957) 13; Sven Ove Hansson, *The Ethics of Risk: Ethical analysis in an uncertain world* (Palgrave Macmillan 2013) 11-16; Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 26-28; and Kenneth Binmore, *Rational Decisions* (Princeton University Press 2011) 35.

¹⁹⁵ See Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 25-26, and Kenneth Binmore, *Rational Decisions* (Princeton University Press 2011) 2-3.

<u>Decision options</u>	<u>States of the World</u>	
	Rain	No Rain
Umbrella	Dry – Luggage (7)	Dry - Luggage (7)
No umbrella	Wet (0)	Dry (10)

If we combine the decision matrix tool and the levels of certainty about the states of the world, we can say that if we are certain that it will rain; we also know that the best decision is to bring the umbrella to the office. If we however know, based on a reliable weather forecast, that there are 30% probabilities of a rain in the afternoon, the umbrella decision would be categorised as being taken under risk. If we do not know the probabilities, but we know that there is a chance of rain, the decision about the umbrella would be taken under uncertainty. If unspecified unknown events could affect the decision about whether to take the umbrella, such as if the sprinkling system in the office would unexpectedly go off, in which case it would prove handy to have an umbrella, that would count as an additional unknown state of the world with unknown probabilities. Decision taken under the influence of such random unknown events would fall into the ignorance category.

Since almost nothing regarding the future is 100% certain, most decisions are based on an educated guess of which state of the world is likely to occur. Decisions are thus usually subject to varying degrees of uncertainty, which is intuitively integrated into the process of taking everyday decisions. Often, we cannot shy away from taking decisions that are subject to great uncertainty, we simply must do our best to estimate the probabilities and act on that. When we do so we act in the way that we think is optimal for the purposes of achieving what we consider the motive of our action. In other words, we act in a way that we expect to yield the best result. The optimal decision is thus the one that maximises expected utility. Expected utility theory, which remains the orthodox method of analysing decisions under uncertainty,¹⁹⁶ is based on this thinking.

¹⁹⁶ See Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 29.

2.3. Expected utility theory

Expected utility theory relies on inputs of two kinds: firstly, it requires information about the probabilistic denotation of each of the relevant states of the world; and secondly, information about the potential payoff for each available action in terms of preference satisfaction.¹⁹⁷ To use the Bayesian version of expected utility theory as grounds for decisions, fairly objective preferences are required and fairly objective probabilities as well. An assumption must be made that the probabilities of each state of the world are known, and that the probabilities are mathematically complete and consistent. Such decisions thus are always taken under certainty or risk, never under uncertainty or ignorance.¹⁹⁸

2.3.1. Von Neumann and Morgenstern

The modern form¹⁹⁹ of expected utility theory is based on a simple method to assign utility values to different decision options which was invented by the famed game theorist, John von Neumann.²⁰⁰ In a simplified version, the method works in the following way: Start by assigning the value zero (0) to the worst option and one (1) to the best option. To find out the utility value of a third option for a given person, he should be offered the choice of a gamble between winning the best option and the worst option, or receiving the third option for sure. Start by offering low odds of winning the best option and then gradually improve the odds until the person would be willing to take the chance in the gamble, instead of receiving the third option for sure. If the person shifts at 30% odds of winning, the utility value of the third option is 30% of 1, that is, 0.3. If the person wont shift until the odds is 90% of winning the best option, the utility value of the third option is 0.9. Von Neumann and Morgenstern defined probabilities as frequency in the long run, which enabled them to attach numerical meaning to the concept of probabilities.²⁰¹

¹⁹⁷ See Nolan McCarty and Adam Meirowitz, *Political Game Theory: An Introduction* (Cambridge University Press 2014) 27.

¹⁹⁸ See Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 37-40.

¹⁹⁹ The origins of expected utility theory are usually traced back to Daniel Bernoulli and his proposed solution to the St. Petersburg paradox in 1738. However, it was not until by the mid-20th century that the modern version gained popularity in mainstream economics.

²⁰⁰ See John von Neumann and Oskar Morgenstern, *The Theory of Games and Economic Behavior* (first published 1944, 3rd edn, Oxford University Press 1954) 15-31.

²⁰¹ See John von Neumann and Oskar Morgenstern, *The Theory of Games and Economic Behavior* (first published 1944, 3rd edn, Oxford University Press 1954) 18-19. See also Kenneth Binmore, *Rational Decisions* (Princeton University Press 2011) 36-37.

Expected utility theory, as presented by Von Neumann and Morgenstern, allows for the comparison of different consequences of different choices by enabling the assignment of specific numeric values signalling the utility of each choice, i.e. cardinal ranking. Prior to their formulation, the ordinal ranking notions of *better than - less than - indifference* were the only available means to formulate preference relations. To make the numerical representation of utilities possible, Von Neumann and Morgenstern defined four rationality axioms that needed to be satisfied by the person upon which the utility function is assigned. The Von Neumann and Morgenstern rational person must have *complete* and *transitive* preferences, *independent of irrelevant alternatives* and which satisfy the axiom of *continuity*.²⁰²

We have already defined three of the axioms. The *continuity* axiom can be explained by imagining a situation where you have three options; the best, the worst, and one in-between. According to the *continuity* axiom, it should be possible to find a probabilistic combination of the best and the worst option that would be equally attractive to the option in-between. In other words; if one has a very good probabilistic option that is combined it with a very bad probabilistic option, one will end up with a combined probabilistic option that is somewhere in-between. By mixing the probabilistic weight of the best and the worst option in correct proportions, it should be possible to produce a mix that is equally attractive to any point on the line between the two options. The continuity axiom implies that there is a tipping point on the line between the two polarising options, where the preference for the third option shifts.²⁰³

The axiom of *reduction of compound lotteries* is often included as an assumption in expected utility modelling; it adds the condition that the rational agent will only care about the compound probabilities of a multi stage process, and will thus have the same preference

²⁰² See John von Neumann and Oskar Morgenstern, *The Theory of Games and Economic Behavior* (first published 1944, 3rd edn, Oxford University Press 1954) 26-27.

²⁰³ We can also model this on a lottery: Option 1 = 100% chances of winning, option 2 = 50% chances of winning, option 3 = 0% chances of winning. Probabilities of winning with option 2 are equal to the combined probabilities of winning with option 1 and 3 $((100\% + 0\%)/2 = 50\%)$. The formal mathematical representation of the continuity axiom is as follows: A preference relation \geq for the alternative set A-C is continuous if for any A, B, C \in A-C (where $A \geq B \geq C$) there exists some probability $(p) \in [0,1]$ where $pA+(1-p)C \equiv B$.

towards an outcome of a single lottery that has corresponding probabilities to the compound probabilities of multiple lotteries.²⁰⁴

Once we have a person that satisfies the Von Neumann and Morgenstern rationality criteria, it is possible to assign numerical utility values to her preferences and thus identify the optimal action that will maximise her expected utility.²⁰⁵

The practical application of the Von Neumann and Morgenstern expected utility criteria is however limited in several important ways. Firstly, we must have objective information about the decision maker's preferences, and secondly, we must have objective information about the long-term frequency occurrence of each state of the world to enable probability assignment on each state. These two conditions mean that only decisions under certainty and risk are applicable. A third important limitation is that the Von Neumann and Morgenstern expected utility function only applies to a specific individual. Further assumptions about interpersonal comparison of revealed preferences is needed to extend the Von Neumann and Morgenstern utility function to a group of people, or whole societies.

2.3.2. Savage and Harsanyi

The limitations to the application of the Von Neumann and Morgenstern expected utility model, mentioned above, have to some extent been overcome by later amendments to the model. The statistician Leonard J. Savage solved the limitation regarding objective probabilities, which restricted applicability to decisions under risk, by using a theory of subjective probabilities as basis for utility assignment.²⁰⁶ As mentioned before, the Von Neumann and Morgenstern model assumed that probabilities are frequencies in the long run. Within statistical theory, there is a longstanding debate about how probabilities should be perceived. On one side, there are proponents of *frequentist inference* of statistical data, which

²⁰⁴ See Nolan McCarty and Adam Meirowitz, *Political Game Theory: An Introduction* (Cambridge University Press 2014) 29-33.

²⁰⁵ This is captured in the Von Neumann and Morgenstern utility representation theorem: Utility (U) can be represented as a function from choices (X) to real numbers (R); $U : X \rightarrow R$. This assigns real numbers to a consequence of an action, which reflects the preferences of the decision maker. Complying with the four Von Neumann and Morgenstern rationality axioms a person will then prefer choice A over choice B, if and only if the expected utility (Eu) of A is greater than the expected utility of B; $A > B$ iff $Eu(A) > Eu(B)$.

²⁰⁶ See Leonard J. Savage, *The Foundation of Statistics* (first published 1954, 2nd edn, Dover Publications 1972).

consider probabilities as objective evidences in the way Von Neumann and Morgenstern assumed in their model. On the other side are Bayesians, who prefer *Bayesian inference* of statistical data and which view probabilities as a subjective perception that is updated as new beliefs emerge. Importantly for our purposes, the later allows for probability beliefs about uncertain events, while the former does not. Savage worked out a mathematical theorem based the intuition that decisions about uncertain events are taken as if we knew the probabilities of the consequences that our actions will have. He inferred that we have consistent beliefs based on which decisions are taken, irrespective of certainty or uncertainty of the future states of the world. Decisions are thus always taken based on the available data and probabilities about future states of the world inferred on this basis as well.²⁰⁷

Savage's thinking centres on *a priori* beliefs about the states of the world. Based on these *a priori* beliefs, decisions are taken with *posteriori* consequences. Once these *a posteriori* consequences have emerged, they in turn form a basis for a new set of *a priori* beliefs for the taking of the next decision. This line of argument is exposed to an infinite regress attack, concerning the origins of the first *a priori* belief; how is *a priori* belief formed in the absence of an initial *a posteriori* consequence? This objection can be overcome. If we are completely ignorant about what has happened in the past, we form our *a priori* beliefs about future acts by imagining the consequences of different acts in different states of the world. By doing that we have formed *a posteriori* belief through a thought experiment. By moulding such *posteriori* beliefs into a consistent probabilistic set of beliefs about the future, we have simultaneously formed a consistent set of beliefs that becomes the *a priori* belief set, on the basis of which a decision is taken.²⁰⁸ This means that if we know nothing about what has happened in the past, we try to imagine what could possibly happen in the future and take a decision based on our best estimate thereof. If we know something about the past, we use that knowledge to inform our imagination about possible future events and take decisions about our future actions based on that.

²⁰⁷ For a simplified version of Savage's Bayesian thinking see Leonard J Savage, 'The Foundations of Statistics Reconsidered' in Jerzy Neyman (ed), *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability - Volume 1: Contributions to the Theory of Statistics* (University of California Press 1961) 575.

²⁰⁸ See Kenneth Binmore, *Rational Decisions* (Princeton University Press 2011) 129-34.

Notably, on Savage's view probabilities are not assumed; they are derived from existing preferences over acts. This limits the applicability of the theory to small worlds, that is, worlds where the potential states of the world can be imagined and definitive consistent preferences over potential acts be formed. Large worlds where the potential states of the world are beyond imagination are, by Savage's own admission, not captured by his Bayesian probability inference. The reason is that the emergence of new evidences has the potential to change the existing preference relations and thus the probability assessment, which in turn undermines choice consistency that is a prerequisite for a rational decision model.²⁰⁹ Importantly, Savage's objective probabilities only apply to individuals. There is no guarantee that two persons will make the same probabilistic assessment from their private preferences.²¹⁰ These two caveats reduce the practical applicability of Savage's probability theorem for the purposes of social decisions in large worlds involving groups of people.

John Harsanyi further advanced Savage's Bayesian approach by adding an additional assumption to allow for interpersonal comparison of utility.²¹¹ Harsanyi's thinking was based on the insight that, in Savage's model, individual differences in probabilistic assignment to different states of the world must be explained either by asymmetry in information that would lead to differences in *a posteriori* beliefs, or by asymmetries in *a priori* beliefs that would also lead to asymmetries in *a posteriori* beliefs. To counter this problem, Harsanyi proposed the assumption of the *common priori*. The *common priori* assumption, assumes that every person has the same information to begin with, and due to the rationality requirement, all will form a *common priori* belief about the current state of the world. Harsanyi argued that there was no reason to believe that if people were feed the exact same information, that they would subsequently form different subjective probability beliefs about the future states of the world, given that people acted rationally.

²⁰⁹ See Leonard J. Savage, *The Foundation of Statistics* (first published 1954, 2nd edn, Dover Publications 1972) 16-17.

²¹⁰ Savage used the more fluent term *personalistic view of probability* for what is now usually referred to as subjective view or Bayesian view. See Leonard J. Savage, *The Foundation of Statistics* (first published 1954, 2nd edn, Dover Publications 1972) 3-5.

²¹¹ Harsanyi published his ideas in three technical papers. See John Harsanyi, 'Games with Incomplete Information Played by Bayesian Players I-III' (1967-1968) 14 *Management Science* 159; 320; 486.

Robert J Aumann elaborated further about the consequences of Harsanyi's *common priori* by showing that '*people with the same priors cannot agree to disagree*' if their beliefs are *common knowledge*.²¹² By common knowledge, he meant a situation where people know each other's beliefs. The significance of Harsanyi's and Aumann's insights, is that by using their assumptions about the *common priori* and about the *common knowledge*, it becomes possible to use Savage's objective probabilities to infer utility values on uncertain decision consequences that affect groups of people.

There have been some attempts to overcome the other main limitation of Savage's decision model, the restriction to small worlds, which have so far not resulted in an orthodox method.²¹³ We have previously categorised large world situations as decision problems under ignorance, or where the unknown is unknown. Given the open contextual nature of such decisions, any consistent precision in decision making towards the optimal decision is difficult to model. The main choices seem to be to pursue a cautious strategy to such decisions, or optimistic strategy, anchored in the Bayesian decision model of Savage. The preferred risk attitude will often depend on the circumstances and the stakes of the decision. For our purposes of modelling legal decisions, Savage's model already provides sufficient tools to sketch a rudimentary framework for consistent legal decision making under risk and uncertainty. Before we proceed with doing so in the next chapter, few extra tools need to be introduced that can provide additional depth and accuracy. These tools belong to a branch of decision theory that has gained increasing scholarly attention in recent years, through the emergence of behavioural economics and its application in variety of contexts. This is the prospect theory of Kahneman and Tversky.

2.4. Prospect theory

The expected utility model of Von Neumann and Morgenstern has a strong intuitive appeal in its rational simplicity. This quickly made it immensely popular for modelling all kinds of economic situations, both normatively and descriptively. It became generally accepted that

²¹² In technical terms, he stated that if '*two people have the same priors, and their posteriors for an event A are common knowledge, then these posteriors are equal*'. See Robert J. Aumann, 'Agreeing to Disagree' (1976) 4 *Annals of Statistics* 1236.

²¹³ Hansson describes several methods in; Sven Ove Hansson, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005) 59-63.

people both *ought to* abide to the rationality axioms underpinning expected utility theory, and that people in fact largely *did so*.²¹⁴ Soon after its matriculation into mainstream economics by the mid-20th century, experiments began to test its empirical validity for descriptive purposes. Some of the results showed that the rationality axioms were consistently violated in certain decision problems.²¹⁵ In their famous paper of 1979, Kahneman and Tversky gathered these consistent violations and proposed an alternative model to give a better descriptive account of decisions under risk that could accommodate these known consistent violations of the rationality axioms.²¹⁶ The alternative model became known as *prospect theory*.

2.4.1. Empirical invalidity of the rationality assumption

As a basic premise, Kahneman and Tversky listed five widely tested and proven phenomenon that violate the traditional rationality model. Their aim was to provide a descriptive model that would incorporate these violations. The violations they listed were the following:²¹⁷

Framing effect: Rational choice theory assumes that persons have the same preferences for rationally equivalent choices. This has been shown to be false in several experiments. The presentation, or the framing, of choices consistently influences preferences.

Source dependence: Ellsberg's paradox showed that people prefer betting on known probabilities to logically equivalent unknown probabilities. Other studies have shown that people prefer risk in their area of competence to unknown areas, even if the unknown area

²¹⁴ Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263.

²¹⁵ A famous early example is Maurice Allais's paradox which showed that people prefer a certain outcome over an uncertain outcome, even if the expected utility is higher for the uncertain outcome. See Maurice Allais, 'Le Comportement de l'Homme Rationnel devant le Risque, Critique des Postulats et Axiomes de l'Ecole Americaine' (1953) 21 *Econometrica* 503. Another widely used early example is Ellsberg's paradox, which shows that people tend to avoid ambiguous choices over familiar risks, even if the expected utility of each choice predicts indifference. See Daniel Ellsberg, 'Risk, Ambiguity and the Savage Axioms' (1961) 75 *The Quarterly Journal of Economics* 643.

²¹⁶ Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263.

²¹⁷ The best short summary of these violations is found in; Daniel Kahneman and Amos Tversky, 'Advances in Prospect Theory: Cumulative Representation of Uncertainty' (1992) 5 *Journal of Risk and Uncertainty* 297, 298. The following list is based on this summary. For detailed references to the experiments that reveal the listed violations I refer to the cited paper of Kahneman and Tversky (1992).

provides a clearer risk choice. We like what we know, even if it is not necessarily rational to do so.

Nonlinear preferences: Allais's paradox showed that people tend to prefer certainty to rationally equivalent risky prospects. People even seem to like certainty over logically superior risk options. Similar effects have been measured for prospects that do not involve certainty. Preferences do thus not seem obey probabilistic linearity.

Risk seeking: The rationality model assumes that one would be indifferent between taking a risk and receiving immediately the expected outcome of that risk. Studies have however shown that people consistently prefer the prospect of a small probability of winning a large prize, to the prospect of receiving for sure the expected value of the same risk. This explains why people buy lottery tickets. The same violation is consistently measured when people have a choice between a certain loss and the prospect of a high probability of a larger loss. In such cases people tend to take the risk of a higher loss instead of accepting the certainty of a lower loss.

Loss aversion: People tend to value losses higher than corresponding gains. The prospect of losing 1000 € is perceived as much worse than the corresponding joy of gaining 1000 €. Several studies have shown a remarkably strong loss aversion affects for even mundane things as inexpensive coffee mugs. The consequence of loss aversion is that you may not be willing to sell an item you bought for 100 €, unless you receive more than 100 € in return.

2.4.2. The alternative model of prospect theory

Decisions under risk in prospect theory occur in two successive phases that are meant to reflect the intuitive process of actual decisions. The function of the initial *framing phase*²¹⁸ is to process the relevant options and the available information into a simplified set of decision options.²¹⁹ During this phase, *'the decision maker constructs a representation of the acts, the*

²¹⁸ In their 1979 paper Kahneman and Tversky use the terminology *editing phase* and *evaluation phase*, but in their 1992 paper they use the terms *framing phase* and *valuation phase* to describe the same phenomenon. For our purposes, we shall use the later terminology.

²¹⁹ See Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263, 274-5.

contingencies, and outcomes that are relevant to the decision'.²²⁰ The options are then compared and evaluated in the subsequent *valuation phase* where a detailed comparison and eventual selection occurs. Some of the violations of the rationality model can be explained by the psychological function of the framing phase in simplifying details. Low-key elements of hidden importance are thus occasionally omitted before the actual valuation and selection takes place.²²¹

Prospect theory in mathematical terms contains four functional elements that distinguish it from expected utility theory: reference point dependence, loss aversion, diminishing sensitivity, and probability weighting.²²²

The existence of the reference point is the key insight of prospect theory. In expected utility theory, utility is derived from absolute levels of the relevant value. In prospect theory, utility is derived from gains or losses in relation to a neutral reference point.²²³ During the framing phase, the relevant options are coded as either losses or gains in relation to a reference point. As the framing effects show, the actual location of the reference point is sensitive towards how the decision problem is presented. If an option is presented in relation to a given reference point, the valuation of the outcome becomes dependent on that relation. If a logically equivalent option is then presented in relation to a different reference point, the valuation of the outcome might be different. The framing phase is meant to correct for such mistakes by coordinating the relevant reference point across different options, but the framing is not always successful in doing so.

Linked to the reference point is the diminishing sensitivity to changes in absolute gains and losses. In expected utility theory gaining 10.000 € has the same utility value regardless of the current status of wealth. In prospect theory, the prospect of gaining 10.000 € derives different

²²⁰ See Daniel Kahneman and Amos Tversky, 'Advances in Prospect Theory: Cumulative Representation of Uncertainty' (1992) 5 *Journal of Risk and Uncertainty* 297, 299.

²²¹ See Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263, 274-5.

²²² See Nicholas C. Barberis, 'Thirty Years of Prospect Theory in Economics: A Review and Assessment' (2013) 27 *The Journal of Economic Perspectives* 173.

²²³ See Nicholas C. Barberis, 'Thirty Years of Prospect Theory in Economics: A Review and Assessment' (2013) 27 *The Journal of Economic Perspectives* 173, 175 and 191.

utility for those who own nothing and for those who already have plenty. To correct for this effect, prospect theory uses a value function that is concave for gains and convex for losses. This incorporates the diminishing sensitivity effect for absolute gains and losses into the utility function; *'the impact of a change diminishes with the distance from the reference point.'*²²⁴

Loss aversion also relates to the reference point and is incorporated into the value function in prospect theory. According to the expected utility model, there should be a symmetry in the utility function for a loss of 500€ and for a corresponding gain. Empirical studies have however confirmed that *'losses loom larger than gains'*.²²⁵ To capture this effect, the utility function for losses needs to be steeper than for gains. The utility function for losses and for gains relative to the reference point is thus asymmetrical.

Probability weighing is also a distinct feature of prospect theory. In expected utility theory, probabilities influence utility objectively by their true value. In prospect theory, probabilities are qualified by a weighing function, which controls how probabilities affect utility. The weighing function proposed by Kahneman and Tversky, overweighs low probabilities and underweighs high probabilities.²²⁶ This is meant to reflect the empirically observed behaviour of risk seeking at the both ends of the probability spectrum.

The formal representation of the prospect theory model, for the prospect x with the probability p and for the prospect y with the probability q , is the following, where V is the overall value of the prospect, v is the value function and π is the probability weighing function:²²⁷

$$V(x,p;y,q) = \pi(p)v(x) + \pi(q)v(y)$$

²²⁴ See Daniel Kahneman and Amos Tversky, 'Advances in Prospect Theory: Cumulative Representation of Uncertainty' (1992) 5 *Journal of Risk and Uncertainty* 297, 303.

²²⁵ See Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263, 279.

²²⁶ See Daniel Kahneman and Amos Tversky, 'Advances in Prospect Theory: Cumulative Representation of Uncertainty' (1992) 5 *Journal of Risk and Uncertainty* 297, 297-8 and; Nicholas C. Barberis, 'Thirty Years of Prospect Theory in Economics: A Review and Assessment' (2013) 27 *The Journal of Economic Perspectives* 173, 176.

²²⁷ This is the original formulation, see Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263, 275-6.

After the initial framing phase, were the different prospects are adjusted and fitted within a specific reference frame. The overall value (V) of each prospect is found through the valuation phase where the outcome of the value function (v) is multiplied with the probability weighing function (π). The prospect with the highest overall value (V) is the optimal choice.²²⁸

The value function (v) is defined as a deviation from a reference point. It is concave for gains and convex for losses, and it is steeper for losses than for gains. Value (v) is thus a function of two arguments; the current absolute level of the relevant value, and the changes in magnitude measured from the absolute level.²²⁹

The weighing function (π) has similarities to Savage's objective probabilities. They are derived from subjective valuation of preferences in a similar way. The difference is that additional to Savage's approach, each probability is multiplied with a number between 0 and 1, which indicates the weighing function. Due to the weighing function, the probabilities lose their compliance with normal probability axioms and should thus not be considered as actual probabilities or measure of degree or believe. The weighing function is scaled so that the weighing function (π) multiplied with an impossible event ($p=0$) equals 0. In the case of a certain event ($p=1$) the weighted probabilities ($\pi(p)$) are scaled to equal certainty (1). This means that events linked to impossible probabilities ($p=0$) are ignored, and the scale is normalised at the point of a certain event ($p=1$). The important element of the weighing function is its departure from the linear probability function near both ends of the probability spectrum. Small probabilities are overweighted, which corresponds to the observed tendency of people to either ignore or exaggerate the likelihood of unlikely events. The difference between large probabilities and certainty is underweighted, which corresponds to the observed tendency of people to either ignore the difference or overestimate it. The behaviour of the weighing function is thus not very predictable close to the ends of the probability spectrum.²³⁰ This in effect means that close to the two boundaries of impossibility and

²²⁸ See Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263, 275 and 280.

²²⁹ Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263, 277 and 279.

²³⁰ Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263, 280-283.

certainty on the probability scale, the weighing function shows similar diminishing sensitivity as the value function does with increased distance from the reference point. The weighing function is thus concave near impossibility ($p=0$) and convex near certainty ($p=1$).²³¹

Kahneman and Tversky amended their prospect theory in 1992 to incorporate certain criticism and to extend the functionality to multiple uncertain prospects.²³² For our purposes this development, referred to as *cumulative prospect theory*, explores details regarding the weighing function that are not important at this primitive stage of modelling. I will thus in what follows use the simpler concepts from the original 1979 version.

The key concept of interest in prospect theory and the main departure from the expected utility model is the reference point dependence. This concept is instrumental in understanding how to perceive preferences towards different outcomes. In the expected utility tradition, a measurement of the absolute levels of the preferred outcome is all that matters, while in the reference point dependent tradition, the changes in the levels of the preferred outcome measured from the current position is what matters.

2.4.3. The location of the status quo

There is one final twist in the story of prospect theory. Since its inception in 1979, it has been widely discussed and written about.²³³ One aspect of that discussion further informs the framing phase and the important question of where to locate the reference point. The contribution I am refereeing to is a series of papers written by Botond Köszegi and Matthew Rabin.²³⁴ The key insight they convey, is that the location of the reference point for human decision makers is not necessarily the current status quo, but instead the current expectations of how things will go. They take the example of a prospective painful dental procedure, which

²³¹ See Daniel Kahneman and Amos Tversky, 'Advances in Prospect Theory: Cumulative Representation of Uncertainty' (1992) 5 *Journal of Risk and Uncertainty* 297, 303.

²³² See Daniel Kahneman and Amos Tversky, 'Advances in Prospect Theory: Cumulative Representation of Uncertainty' (1992) 5 *Journal of Risk and Uncertainty* 297.

²³³ For a good overview and references to this discussion see Nicholas C. Barberis, 'Thirty Years of Prospect Theory in Economics: A Review and Assessment' (2013) 27 *The Journal of Economic Perspectives* 173.

²³⁴ See Botond Köszegi and Matthew Rabin, 'A model of reference-dependent preferences' (2006) CXXI *The Quarterly Journal of Economics* 1133; Botond Köszegi and Matthew Rabin, 'Reference-dependent Risk Attitudes' (2007) 97 *American Economic Review* 1047; Botond Köszegi and Matthew Rabin, 'Reference-dependent Consumption Plans' (2009) 99 *American Economic Review* 909.

is cancelled unexpectedly. A person might feel that she has gained something by not having to undergo the procedure. The use of the status quo reference point would not capture this gain, compared with a person who was never going to have any dental procedure in the first place. The status quo reference point would simply view both positions as identical since nothing distinguishing had yet happened.²³⁵

The key insight lies in analysing what the person expected the status quo to be. The two points do often correlate; one expects the current status quo to remain in place. Importantly that is however not always the case. One can have positive or negative expectations that the current status quo is different from what the future status quo will be. For modelling decision-making, this point is crucial. Decisions are made based on how one expects things to be, rather than based on how things are. Often one expects things to be as the currently are, but not always.

In their paper, Köszegi and Rabin apply their model to two studies that provide empirical support for their theory. The former study concerns shopping behaviour of consumers. It has been shown that if a consumer expects to buy something, he is willing to pay more for it than for something that he did not expect to buy. This indicates that the endowment effect starts to work even before actual acquisition. The explanation given by Köszegi and Rabin is that the gain/loss reference point becomes what one expects to own in the future, not what one currently owns. If you do not buy what you expect to own, it will be experienced as a loss and thus you are willing to pay more for it, than for something that you did not expect to own.²³⁶

The later study concerns working patterns of taxi drivers. In the empirical study referred to, it was shown that the working pattern of taxi drivers is affected by their target income. If they expect to receive a certain income during the day, their hours on the job are sensitive towards that target. If they reach the target early in the day, they were shown to be more likely to work less in the afternoon. They were also shown to spend longer hours on the job if the target was not achieved early. Intuitively this behaviour should not make much sense. Presumably, it

²³⁵ See Botond Köszegi and Matthew Rabin, 'A model of reference-dependent preferences' (2006) CXXI The Quarterly Journal of Economics 1133, 1141-43.

²³⁶ See Botond Köszegi and Matthew Rabin, 'A model of reference-dependent preferences' (2006) CXXI The Quarterly Journal of Economics 1133, 1146-50.

would be rational to work long hours when the income is fast, and short hours when the income is slow. The explanation inferred from Köszegi's and Rabin's theory is that the expected income forms a gain/loss reference point, instead of the current status of wealth when they start to drive. Loss aversion with reference to the reference point of expected income induces the drivers to stay out long when business is slow, and to take it easy when the target income is reached.²³⁷

In summary, prospect theory provides tools to describe actual preferences more accurately than the Von Neumann and Morgenstern expected utility model does. Prospect theory incorporates several consistent consistency violations, which categorically induce people to choose options that appear irrational.

3. Improving the traditional by using elements of the alternative

To translate the normative premises introduced in chapter two (i.e. egalitarian fairness) into legislative praxis, robust methodological tools are needed. The traditional proportionality test and the cost benefit analysis represent different methodological approaches to such optimisation, that at a closer look both depend on a normative *a priori* as a benchmark of optimisation, and both also rely on a quantitative method of measurement that use consequential inputs.

An assessment of the traditional methodologies revealed that some intricacies of their normative premises are bound to be lost in translation towards the eventual legislative output. This provides an opportunity for methodological improvements; if a cost benefit analysis and the proportionality test have limitations in their capacity to identify the optimal legislative choices, then perhaps new methodologies could provide improvements.

Towards this end, several tools and ideas from decision theory have now been discussed and introduced. These tools focus on the nucleus concept of preferences as the locus object of measurement, which serves our purposes well in finding a methodology for optimising

²³⁷ See Botond Köszegi and Matthew Rabin, 'A model of reference-dependent preferences' (2006) CXXI The Quarterly Journal of Economics 1133, 1150-55.

legislative decisions within a system of government that is based on the abstract idea of the social contract. By its very definition, a contract is all about reaching a compromise between the preferences of the contracting parties.

Through the discussion about the basics of decision theory, several ideas and conceptual tools have been introduced, that in the next chapter will be developed into a novel coherent methodology for optimising the normative premise of egalitarian fairness into the output of the legislative process. The self-proclaimed benchmark of success for this novel methodology would be to achieve uniformity in simultaneously assessing the preferableness of moral factors and economic factors in legislation. Individually the orthodox approaches of the proportionality test and a cost benefit analysis are well equipped to separately assess moral appropriateness or economic efficiency, but less equipped to assess both at the same time.

The model of fair rules

The problem of fairness in the laws is essentially a problem of creating a legislative output that achieves success in balancing the preferences of the concerned individuals, so that they individually feel that they are being treated fairly by the legislator. For the legislator to achieve methodological consistency and accuracy in conducting this balancing task, I suggest the use of the *model of fair rules*. As discussed in chapter three, alternative tools for legislative optimisation do exist in several forms; the novelty of the *model of fair rules*, compared with the tools already discussed, is its ability to uniformly assess and quantify with enhanced accuracy social preferences towards both economic factors and factors based on moral beliefs.

The key move I want to make in this chapter is to explain how the concept of egalitarian fairness in the abstract, which as previously explained encompasses a Pareto requirement and an egalitarian distribution requirement, can translated and used for optimisation purposes in actual legislative work. For these purposes, I have designed the *model of fair rules*, using methodological components and concepts from decision- and rational choice theory.

The *model of fair rules* claims to be a practical method for identifying legislative options that are optimised towards efficiency and moral appropriateness. The model is based on several assumptions about the role of the legislator that I will start by briefly outlining.²³⁸ Having identified the key assumptions, I will proceed with describing the main elements of the model, and its intended function.

1. The social contract and the role of the legislator

The idea of the social contract is my starting point for identifying the key duties of the legislator within a democratic system of governance. When acting in isolation from other persons, the individual can indulge his preferences as he sees fit. When acting within the confines of a cooperative social group, the individual must adjust his behaviour so that his preferences do not unduly burden the others. The individual has an incentive to restrain his desires if the

²³⁸ These assumptions are explained and discussed in more details in chapters two and three.

aggregated impact of cooperation on his preference satisfaction is greater than the aggregated impact from acting alone. Social contracts can form spontaneously so long as the individual benefit for cooperating outweighs the individual cost of restraining private desires. To remain spontaneously stable, the social contract must also be more efficient than competing contracts (preferably Pareto efficient), and it must distribute the benefit equally among its members (based on an egalitarian distributional principle) to avoid rebellions and division of the social group. This is a Machiavellian account of the social contract; spontaneous social cooperation, kept together by the prospect of personal gain, for as long as nothing better comes around.²³⁹

We must also be aware that stability within primitive social groups can be artificially maintained through external enforcement. Incentives for specially conditioned social behaviour can be created by a credible threat of violence and other cost inflicting strategies. Externally enforced social behaviour can deviate from what would otherwise spontaneously occur. For the purposes of this thesis, the primary interest is in the principles of spontaneously stable social contracts, not in an externally enforced equilibrium behaviour.

The rational member of the social contract must make an essential compromise between his own unrestrained preferences, and the competing preferences of others. By following two simple bargaining maxims this compromise can form a stable equilibrium solution: Firstly, the *efficiency maxim* dictates that the proposed action should have efficient consequences; and secondly the *fairness maxim* dictates that the consequences of the proposed action must be equally preferable to each member of the group.

If a cooperative action were not anticipated as efficiency enhancing, the members of the social group would not have any incentive to undertake it. If the rewards of the joint effort were not to be distributed equally, some would have a weaker incentive than others to participate in the action. Disunity undermines efficiency through lower total production of the members. This creates the incentive to abandon the group for a competing group organized on egalitarian

²³⁹ This outline of the social contract, which is discussed in more details in chapter two, draws on Machiavelli's account of the civil principality, Thomas Hobbes's *Leviathan*, and on Kenneth Binmore's game theoretical approach. See Niccolò Machiavelli, *The Prince* (first published 1513, Quentin Skinner and Russel Price eds, Cambridge University Press 1988); Thomas Hobbes, *Leviathan* (first published 1651, Richard Tuck ed, 2nd edn, Cambridge University Press 1996); Kenneth Binmore, *Natural Justice* (Oxford University Press 2005).

principles that would, due to better incentivised members, yield higher total rewards. The stable spontaneous social contract thus relies on the bargaining maxims of *efficiency* and *fairness*.

The function of the legislator in a modern democracy, can be viewed as a caricature of the primitive bargaining process that we frequently use for solving every day cooperative decision problems. In a setting of few people, bargaining can be concluded through a short discussion based on the two bargaining maxims (i.e. by finding an efficient and a fair solution). Within larger contexts, the simple process of bargaining can be simulated with the aid of simplifying assumptions.²⁴⁰ The idea of democracy can be considered as an institutionalisation of the bargaining process on a societal scale. A democratic action on a societal scale should ideally be efficient and fair with regards to the society's members. In the modern democracy, legal acts are to be considered as the product of a democratic action. Following our premises, each legal act should accordingly strive for efficiency and fairness in line with the maxims of primitive bargaining. The role of the legislator in a democracy, premised in this way, is to identify and implement legislative options that optimize this normative agenda. By doing so the legislator is simulating a societal wide bargaining process, between the competing preferences of the society's members; notably in a simplified way to ensure expedient and efficient practical decision-making.²⁴¹

I have now described the legislator as an arbitrator of conflicting preferences of the members of society. If he is premised on the idea of democracy, he should select legislative options (i.e. actions) that achieve fair and efficient preference satisfaction (i.e. consequences) of his peoples. For the completion of our caricature of the legislator, a final component is needed. I have described two maxims that should guide his decisions once he has swung into action, but I have yet to explain the rationale for acting in the first place.

²⁴⁰ The most notable simplifying assumption is the principle of the majoritarian rule, which simplifies the task of reaching bargaining consensus within large groups and thus reduces the cost of reaching decisions.

²⁴¹ Note that the simplification is usually layered; with a more detailed simulation for more important decisions. Ranging from decisions with constitutional ramifications that often require referendums or some special amendment process, to mundane administrative orders that can be unilaterally issued by plain government officials. In practical terms this averts the prohibitive cost of simulating in details a bargaining process for every action of the democratically governed state.

Sophisticated mainstream political ideologies fundamentally disagree about the optimal level of regulation and deregulation. We can differentiate between variants of Adam Smith's invisible hand that allows for a spontaneous decentralised formation of the most efficient form of social cooperation, and theories premised by the need for an active central coordinator that identifies and polices the optimal social cooperative strategy.

For the legislator that follows the maxims of the social contract, the justification for a legislative intervention could be based on either of the two maxims. If the current status quo lacks fairness or efficiency, the legislator has a valid argument to intervene. If, however the current status quo is already at a stable equilibrium, optimized with regards to fairness and efficiency, an argument for intervention would be missing. This would apply regardless of whether the optimal situation resulted from a previous legislative intervention, or came about spontaneously through social cooperation. Viewed in this way, legislative interventions have the primary purpose of maintaining an optimal efficiency level by responding to external changes, and to fairly rearrange the internal allocation of the efficiency output to fit the current mixture of preferences.²⁴²

The rationale for initiating a legislative action has now been explained, but the actual object of fairness and efficiency on each occasion is still missing. If we assume that we generally act to bring about specially targeted consequences that we prefer, we can also infer that we individually participate in cooperative actions based on the same rationale. Each cooperative action is thus aimed at satisfying an individual preference or a set of individual preferences. The rationale for participation does not need to be the same for all, and the intensity of gain that each action brings can vary between individuals. To achieve his task, the legislator thus needs to map these reasons and find a way to quantify them to correctly identify the optimal legislative solution based on the fairness and efficiency maxims.

²⁴² The discovery of new efficient techniques of satisfying preferences could count as external circumstances, which the legislator might want to implement through legislation to improve efficiency. New ideas about what is preferable within a group, such as increased tolerance towards different life choices, may similarly prompt the legislator to intervene so that the satisfaction of diverse preferences is fairly facilitated through the laws.

The role of the legislator is to take as good decisions as he can; he strives for the optimal decision. His primary preference is to excel at taking actions that bring about consequences that he personally prefers. If he is conditioned by the maxims of the social contract he will satisfy his own preferences by optimizing efficiency and fairness of the preference satisfaction of his peoples. The legislator, on this view, is like an external observer that takes his sole pleasure out of tuning to perfection the realisation of the desires and pleasures of those he observes.

In chapter three two popular methods for assessing competing legislative options were described, that each has pros and cons. The cost benefit analysis is an excellent tool for assessing the type of preferences that can be monetized and the proportionality test focuses on different types of preferences that are not easily expressed in monetary terms. Each method gives disproportional weight to the type of preferences they are good at measuring, at the cost of other potential considerations. Both neglect distributional concerns for the type of efficiency they each measure, and both can improve with regards to quantitative accuracy.

If we assess these two methods from the perspective of the legislator's role, as I have just described it, the following shortcomings emerge: i) *Methodological accuracy*; the legislator needs accurate and practical tools for identifying the best legislative options. Both methods have flaws with regards to quantification and comparison of preferences. ii) *The Efficiency maxim*; the legislator ought to make efficiency claims about preference satisfaction based on the full spectrum of the relevant consequences. Both methods make efficiency claims that only incorporate part of the potentially relevant consequences. iii) *Fairness maxim*; the legislator ought to select fair legislative options. Both methods allow legislative options that are not optimised with regards to distributional fairness in the egalitarian sense.

With the model outlined below, the hope is to provide an alternative tool for legislators in their task of identifying efficient and fair legislative options. The tool should be accurate, but still practical. The flaws I have described with the orthodox approaches serves as a benchmark for potential improvements.

2. Legislative actions as decision problems

The identification of the optimal legislative decision requires a qualitative comparison and ranking of the available legislative options. We need to establish which actions are comparatively inefficient or unfair towards achieving the legislative consequences pursued. The proposed model uses selected microeconomic tools from decision theory to achieve this comparison.

A decision matrix can be used to frame decision problems in a clear schematic way. Potential legal actions are ranked in rows of a table, against one or more potential scenarios of future events that are listed in the columns of the same table. The conceivable outcome of each act in each potential scenario is then described as a consequence in the cells at the relevant intersection of the table. By using Von Neumann's and Morgenstern's rationality assumptions²⁴³ it is possible to infer cardinal numerical values on the preference consequences of competing decision options, given that the decision is taken under conditions of probabilistic certainty. This is the classic expected utility inference method. Based on further rationality assumptions argued for successively by Savage,²⁴⁴ Harsanyi²⁴⁵ and Aumann²⁴⁶ the inference of cardinal numerical utility rankings, can be extended to decision options for groups of people under decision conditions of risk and uncertainty.²⁴⁷

²⁴³ Von Neumann's and Morgenstern's method is conditioned on the acceptance of four assumptions about the rationality of the decision maker. Their rational decision maker must have *complete* and *transitive* preferences, that are *independent of irrelevant alternatives* and which satisfy the axiom of *continuity*. See John von Neumann and Oskar Morgenstern, *The Theory of Games and Economic Behavior* (first published 1944, 3rd edn, Oxford University Press 1954) 15-31.

²⁴⁴ Savage argued for the Bayesian inference of probabilities (subjective probabilities) over frequentist inference of probabilities (objective probabilities), which can be used to expand the utility inference model to probabilistic situations of risk and uncertainty. See Leonard J. Savage, *The Foundation of Statistics* (first published 1954, 2nd edn, Dover Publications 1972).

²⁴⁵ Harsanyi suggested the assumption of the common priori so that Bayesian probability inference could be extended to groups of people. His argument was that by assuming the same priori beliefs and knowledge, different rational people should make the same probabilistic inference. This significantly enables interpersonal comparison of utility. See John Harsanyi, 'Games with Incomplete Information Played by Bayesian Players I-III' (1967-1968) 14 *Management Science* (I)159; (II) 320; (III) 486.

²⁴⁶ Aumann elaborated further on the consequences of Harsanyi's common priori by showing that 'people with the same priors cannot agree to disagree' if their beliefs are common knowledge. See Robert J. Aumann, 'Agreeing to Disagree' (1976) 4 *Annals of Statistics* 1236, 1236.

²⁴⁷ These assumptions are explained and discussed in more details in chapter three.

The prospect theory of Tversky and Kahneman²⁴⁸ relaxes some of the rationality assumptions to accommodate for certain consistent consistency violations, which are not caught by the expected utility model. This comes at the cost of mathematical complexity, but the benefit is a more accurate description of preferences and thus a more accurate basis for finding the optimal legislative decision. The key insights for the present purposes are the empirical observations of a reference point dependence and loss aversion. The former implies that people tend to assess preference gains and losses in relation to a status quo reference point;²⁴⁹ instead of with reference to absolute levels as normally would be assumed to be rational. The latter implies that losses are conceived as weighing more than corresponding gains; instead of having symmetrical effects on preferences, as would be assumed in the rational model.

3. The elements of the model of fair rules

To enable the legislator to compare and identify the best legislative options, in terms of our fairness concept, he would need to: i) establish a legislative plan; ii) map the relevant stakeholders and their preference function towards the legislative agenda; iii) rank in order of importance the desires of the different stakeholders; iv) describe in cardinal numerical terms the probable consequences of each potential legislative action on the preferences of the stakeholders involved; and v) process the gathered information to extract the optimally fair legislative option. The basic assumptions and the tools of decision theory, outlined above and in more details in chapter three, provide the means to achieve these tasks.

3.1. The plan of the legislator

Every legislative action must start with an idea of what ought to be achieved through the action. Starting from the premise that the ultimate objective of any legislative action is to improve society's equilibrium of fairness and efficiency, we would need a more specific

²⁴⁸ See Daniel Kahneman and Amos Tversky, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263; Daniel Kahneman and Amos Tversky, 'Advances in Prospect Theory: Cumulative Representation of Uncertainty' (1992) 5 *Journal of Risk and Uncertainty* 297.

²⁴⁹ Studies have further shown that people perceive the status quo reference point in terms of their expectations of how things will go, rather in terms of what it ends up being. In that way, the factual reality may not change, but nonetheless people can experience a sensation of loss or gain if they had expected something to happen. See further on this point; Botond Köszegi and Matthew Rabin, 'A model of reference-dependent preferences' (2006) CXXI *The Quarterly Journal of Economics* 1133; Botond Köszegi and Matthew Rabin, 'Reference-dependent Risk Attitudes' (2007) 97 *American Economic Review* 1047; Botond Köszegi and Matthew Rabin, 'Reference-dependent Consumption Plans' (2009) 99 *American Economic Review* 909.

guidance on how to achieve such equilibrium in specific cases. To these ends, the legislator benefits from the deliberations of the political process that devises legislative plans based on policy objectives that seek to convey and contextualise the meaning of the social contract. Based on such legislative plans, the legislator designs legislation with the intention of implementing the plans and thus the policy objectives.²⁵⁰

If legislative plans are viewed as derivative from the fundamental rationale of the social contract, the legislative plans must also abide to the bargaining maxims of fairness and efficiency, and the initiation of a new plan can be justified based on the need to restore an equilibrium that satisfies these maxims within a specific context. In a broad context, we want society to be both fair and efficient from the point of view of its members. In a narrow context, we would also want a chocolate cake to be split fairly and efficiently among the members of a household. The same rationale applies to different kinds of legislation that, for example, deals with a specific taxation, or a specific crime, or the organisation of a specific institution; all legislation should be designed in a way that ensures fairness and efficiency for the relevant stakeholders involved, using their preferences as an ultimate object of optimisation.

For our purposes, a legislative action seeks to implement a legislative plan. The stakeholders in a legislation first need to reach an agreement on the legislative plan, before reaching an agreement on specific legislative actions to implement the plan. The object of optimisation for the former are the preferences of the stakeholders towards a specific legislative policy context, while the object of optimisation for the latter are the preferences of the stakeholders towards the implantation of the mutually agreed legislative plan.²⁵¹

While agreeing on the object of the legislator's plan, the relevant stakeholders in a legislation can hold diverging preferences towards the specific legislative actions undertaken to implement the plan, given that an implementation action that archives perfect optimality with

²⁵⁰ I.e. the hierarchical relation is as follows, starting from the general, ending with the specific: social contract > policy objective > legislative plan > legislative action.

²⁵¹ In parliamentary context, these processes often occur simultaneously; the political compromise on public policy plan is reached at the same time as the technical implementation is decided. The distinction between disagreeing on policy, and disagreeing on the best way to implement a policy, is thus often blurred in parliamentary practice.

absolute certainty is not available. Probabilistic uncertainty of the consequences of specific actions, the practical unattainability of complete optimisation success, and differently situated stakeholders, creates space for disagreement about the favoured legislative action to undertake.²⁵²

3.2. Stakeholders and the reduction of moral and efficiency claims

The stakeholders should be defined based on their prior circumstances during the legislative process with regards to the legislative plan at stake. By inference from Harsanyi's idea about the *common priori*, it can be assumed that stakeholders that do not share *common priori*, are liable to hold different preferences towards the legislative actions that are available for the purposes of implementing the legislative plan, unless the perfectly optimal action is available with absolute certainty, in which case there would not be any reason for disagreement about which action to take. Based on the *common priori* it should be possible to assume what each type of a stakeholder ought to rationally want with regards to the legislative plan at hand, and thus categorise them based on their preference function.

It is important to note that the preferences of different groups of stakeholders are not necessarily equally important. Ronald Dworkin famously talked about rights as trumps, to emphasise that certain preferences of individuals should weigh more than certain collective preferences.²⁵³ Much of the rights literature relies on Kantian thinking about categorical imperatives; rights are imperative reasons for action that override non-rights reasons. The categorisation of reasons for actions, into rights and non-rights is helpful as a simplifying assumption in some cases. Increasingly however, this binary categorisation fails to account for the varying degrees of importance undeniably attached to different rights. The gradual inflation of the list of rights also highlights a problem at the margins of what counts as a right. Often the grounds for including or excluding a reason are not obvious when approached by deontological rationale.

²⁵² See further on the unattainability of perfect procedural accuracy in chapter 5.

²⁵³ See Ronald Dworkin, 'Rights as Trumps' in Jeremy Waldron (ed), *Theories of Rights* (Oxford University Press 1984) 153-167.

For our purposes, we need to let go of the binary rights/non-rights ranking of people's reasons for action. Instead, we should recognise a range of different preferences towards the probable consequences of distinct reasons for actions, which objective importance can be attached too based on the intensity of the preference feeling. Thinking of rights as simplified proxies for a specific preference, offers the possibility of more descriptive accuracy than the binary on-off deontological scale of rights imperatives can ever achieve. A specific right thus signals a preference towards specific consequences, and the intensity of the preference can be measured and compared with other rights on a cardinal scale of importance.

By reducing rights to mere preferable consequences, we gain categorical unity of the legislator's subject of valuation, since efficiency arguments are already traditionally valued based on a consequential methodology. Every potential argument for or against a specific legislative action, including both moral and efficiency arguments, can thus be reduced to a preference towards specific consequences. This does not necessarily eliminate the essence of the deontological idea of the moral imperative. By attaching a high preference value to the consequences that individual rights promote, this imperative importance can be conveyed through the language of preferences.

The shift from a deontological conception of rights, over to a consequential conception, has an intuitive appeal. It is challenging from the outset to explain the idea of multiple imperative rights in strict deontological terms. How do multiple imperative rights make any sense in social organs where compromises are integral to the organs function? It feels more reasonable to think of rights as strong preferences towards some substantive consequences that promote something essential to the human condition. For example, most people have a strong preference towards remaining alive. This preference can be branded as the right to life.

By the nature of things, which exists beyond the conscious human condition, there cannot however be any imperative right to life; on the contrary, the nature of things requires that we must all eventually die, irrespective of what we want. How can it then be imperative in the deontological sense, that we all have the right to life, if it is at the same time imperative in the naturalistic sense that we must all die? The short answer is that there is no such thing as an

imperative right to life in the naturalistic sense, which means that a right to life must be a human artefact. If it is a human artefact, it must also be the expression of something that someone wants, or alternatively the expression of a compromise between the wills of several individuals, both of which leads inevitably to the nucleus concept of preferences. Even if we must all die in the naturalistic sense, most of us would nonetheless prefer to live well for as long as possible, and we would seek to cooperate with others on the mutual fulfilment of that preference.

By reference to the consequential shift, it is easier to explain the possibility of reconciling conflicting moral claims. The need to violate a moral imperative, as deontological thinking would require, is eliminated, and replaced by a negotiation of conflicting preferences. The role of marginal non-rights in relation to marginal rights can also be explained in a more satisfying manner based on consequential thinking about preferences. By eliminating the categorical rights distinction, a uniform set of intra compatible preferences appears, that can be ranked in a cardinal order of importance. This however only extends to stakeholders with the same *common priori* and thus the same preference function. To rank and compare the preferences of different types of stakeholders that do not share *common priori*, an additional device is needed that achieves interpersonal comparison and importance ranking of what each stakeholder type wants.

3.3. Interpersonal comparison of stakeholder claims through a preference index

As mentioned above, branding some preferences as rights is the deontological approach to interpersonally compare the importance of the preference claims of differing types of stakeholders. The one that can claim a right should prevail against those who cannot claim a right. The rudimentary binary scale is however insensitive to the small difference that can be among claims at the opposite margins of what counts as a right, and the binary scale has difficulties with ranking competing claims that belong to the same category on the binary scale of rights and non-rights. Interpersonal comparison based on rights also ignores the importance of efficiency considerations, when pitted against moral considerations. The traditional consequential method of reducing claims to a monetary value and thus achieving interpersonal comparison of competing claims through the universal medium of money faces the inverse

problem of having difficulties with incorporating moral claims due to their lack of practical utility that can be translated by the medium of money.

An initial step in solving these methodological difficulties is the utilisation of preferences as a universal medium that incorporates both moral claims for specific consequences and economic efficiency claims for specific consequences. Once a unity has been established with regards to the medium of measurement for the subjective preferences of likeminded stakeholders, a method is needed to enable objective comparison of the importance of the subjective preferences of different types of stakeholders towards a specific issue. If one group of stakeholders prefers a different solution to a problem than another group of stakeholders prefers, we need a way to assess from an objective point of view, which group has a more important claim. The method I suggest is a *preference index* that interpersonally compares and ranks on a cardinal scale, the objective importance of competing preference claims of different types of stakeholders.

The *preference index* can be used as a sort of a price ratio for competing interests. In the same way as we can establish based on market prices how many units of silver would need to be exchanged for a unit of gold, we should be able to establish a price ratio between a preference unit of stakeholder A, and a preference unit of stakeholder B. We need to overcome two practical obstacles to build the *preference index*: first, we need a system to measure magnitudes of preferences; and second, we need to a way to compare the importance of similar magnitudes of different preferences.

A simple device can be used to solve the problem of magnitudes. In the same way as magnitudes of different precious metals can be described in comparable terms through the metric system, the same is possible for describing magnitudes of different preferences. A numerical value for the maximum fulfilment of a specific preference in consequential terms is simply assigned, and another numerical value for the complete deprivation of the same preference, i.e. the best and the worst consequences in relation to this specific preference.²⁵⁴

²⁵⁴ The exact point of complete deprivation or saturation of a specific preference can create an interpretive problem for some preferences, in specific preferences for the acquisition of specific resources. For example, the preference of unrestrained expression is saturated and deprived at uncontroversial points (total prohibition or

If the assigned numbers are zero (0) and one (1), different preference consequences can be assigned a magnitude value between 0 and 1, depending on the specific intensity of preference fulfilment. The specific level can be found through the normal utility inference method of Von Neumann and Morgenstern. In this way, we can talk about 10% fulfilment of a specific preference in relation to 10% of another preference. The advantage this brings is like the advantage of being able to talk about 100 grams of silver and 100 grams of gold. A uniform medium for comparing magnitudes (e.g. grams), enables the comparison of an objective value in terms of a price per magnitude unit (e.g. price per gram).

An objective value of the preferences of different types of stakeholders can be achieved (i.e. the objective price of each preference unit), either by reference to a universally applicable importance value, or by reference to a case specific importance value. The former method is easier to apply due to its nature of generalisation, but loses in return some descriptive accuracy that can be achieved through a custom-made standard of importance valuation. We can also think of a mixed comparison approach that takes some of the case specific details into consideration, but nonetheless tries to achieve results that can be applied universally to a range of circumstances.

The key to achieving objective comparison can be explained based on the *veil of ignorance* thought experiment of John Rawls.²⁵⁵ We must imagine that we are deciding the level of importance for the competing preferences from behind the veil of ignorance; the decision on the importance weighing will apply to us, but we do not yet know which type of stakeholder we are. Towards the end of achieving objectivity of importance inference, it is not sufficient to make a simple survey of what people prefer. Such results would be exposed to a bias in favour of the actual position of the respondents. We thus need a survey that measures preference importance from a state where the respondent is oblivious about his actual position.

total liberty), while the preference for acquiring money is not as obviously saturated at a specific point. One way to solve this would be to refer to the average lifetime use of a specific resource as a saturation point and the average use over some specific short timespan (a day, a month, or a year) as the point of scarcity. If we again refer to the preference for money, it would be saturated at the amount an average person needs during his lifetime, and deprived if the amount would only last the average person for one day. The preference for other types of resources could be bracketed in a similar way.

²⁵⁵ See John Rawls, *Theory of Justice* (Harvard University Press 1971).

For the simpler comparison of generalised importance, I propose the following method that focuses on absolute importance values, as opposed to the more nuanced importance values that are relative to the current status quo. If we were comparing two preferences (p_1) and (p_2) we would first have to determine which is more important by a simple survey using comparable magnitudes of both. Once an importance relation in the ordinal sense has been established, we use the Von Neumann and Morgenstern utility inference method to find the location of the inferior preference on a scale between the superior preference (1) and a neutral consequence (0). This would give a cardinal ranking of importance for the stakeholder preferences being assessed. This method could provide a good general indication of preference importance, but is exposed to inaccuracies in case specific circumstances since it ignores the variable of reference point dependency.

The more complex comparison method of subjective importance seeks to incorporate the reference point element from prospect theory for achieving more descriptive accuracy in case specific circumstances.²⁵⁶ I propose a method that exploits this point based on the intuition that a subjective level of preferableness, attached to different types of preferences in a given situation, can be inferred from a comparison of the effects on each type from a positive and a negative magnitude change measured from the status quo.

The method can be sketched as follows: Comparing two preferences²⁵⁷ (p^1) and (p^2), their range is bracketed by describing a deprivation scenario ($-p^n$) and a saturation scenario ($+p^n$) for both. This provides four variables ($-p^1$, $+p^1$, $-p^2$, and $+p^2$) that can be bundled into a best case ($+p^{1+2}$) and a worst case ($-p^{1+2}$) scenario. The reference points (Q^{p^1} and Q^{p^2}) form the fifth and the sixth variables that can also be bundled into a general status quo ($Q^{p^1+p^2}$). If we think of $-p^{1+2}$ as the worst possible scenario and $+p^{1+2}$ as the best possible scenario, each of the other

²⁵⁶ The reference point dependence is the key insight of Kahneman's and Tversky's prospect theory and it has been advanced by the insight of the expected status quo as opposed to the current status quo. See Botond Köszegi and Matthew Rabin, 'A model of reference-dependent preferences' (2006) CXXI The Quarterly Journal of Economics 1133; Botond Köszegi and Matthew Rabin, 'Reference-dependent Risk Attitudes' (2007) 97 American Economic Review 1047; Botond Köszegi and Matthew Rabin, 'Reference-dependent Consumption Plans' (2009) 99 American Economic Review 909.

²⁵⁷ The method is not restricted to the comparison of two alternatives. In theory, any number of preferences could be compared and ranked using this method.

points should fall on a line between these two extremes. The relative preferableness of each point could then be extracted and located on the line by using Von Neumann's and Morgenstern's utility inference method. That is how a given magnitude of a given preference can be compared with a corresponding magnitude of another preference to determine their relative impact on the scale between the worst and the best possible consequences.

The use of a reference point adds a small nuance. We are not measuring the impact of a given magnitude in absolute terms; we are measuring the impact of a given change, counting from the currently deposited magnitude. To achieve this, we need to locate the bundled reference point on the overall scale between the best and the worst. The location may not be in the middle, thus creating a problem of unsymmetrical effects of corresponding magnitudes of positive and negative consequences. To overcome this, we need to split the scale into a negative and a positive subscale, breaking at the location of the bundled reference point. The split enables an assessment of the impact of a given loss or a gain magnitude for a specific preference, relative to the bundled status quo, and the bundled best or the worst-case scenario.

To measure how preferable a specific preference type is compared with another type we simply drop measuring points on the defined scales of negative and positive consequences, using the same magnitude for both types (e.g. 20% increase). The number of measuring points affects the accuracy of the description. Dropping only one measuring point risks neglecting different effects of losses and gains for different types of preferences.²⁵⁸ Using measuring points on both the positive and the negative side of the status quo requires symmetry in terms of distance from the reference point, or alternatively a special weighing to ensure that the outcome for gains and losses weigh equally in the aggregate result.

By using the location of the bundled status quo on the overall scale, the registered impact on the two subscales can be translated into a loss or a gain on the overall scale relative to the current level of preference satisfaction. Comparison of the effects of corresponding

²⁵⁸ The key element in some types of preferences may be that we care a lot about maintaining at least the current magnitude, but we would not care so much about an increase. For other types this effect can be vice versa.

magnitudes of different preferences shows the relative impact of each, and thus the relative preferableness of each. By using Von Neumann's and Morgenstern's utility inference method the difference in impact can be expressed in numbers.

A simple start to applying this method would be to use the magnitudes already defined during the bracketing of the total range for each preference as measuring points on the subscales; status quo is zero, total deprivation is a 100% loss on the negative scale, and total saturation is a 100% gain on the positive scale. Assuming completeness, the aggregate preferableness of a bundle of preferences equals the sum of its parts. Each part, i.e. each preference type, can however carry different weight in the bundle. Using two preference types (p^1 and p^2) as an example and the magnitude of a 100% gain as a measuring point, the sum of a 100% gain for p^1 and p^2 , should equal a corresponding magnitude for a bundle of the two. A given magnitude of p^1 can however potentially weigh more, or less, in the bundle p^{1+2} than an equivalent magnitude of p^2 .²⁵⁹ The difference in weight indicates a difference in relative preferableness of the two preference types. The weight of each relative to the bundle and the status quo can be extracted in numerical terms by applying the utility inference method of Von Neumann and Morgenstern. This can be repeated for any number of measuring points to increase accuracy. Subsequently these values on the subscales are translated into values on the overall scale. The aggregate weight of the measuring points for each of the preference types on the overall scale provides a comparison of impact from which relative preferableness of the different types can be inferred. We can call the aggregate result for each preference type a *preference index* and the ratio between different types a *preference ratio*.²⁶⁰

By establishing a *preference index* for each of the relevant stakeholders in a specific legislative plan, the legislator has means to compare the relative importance of differing stakeholder preference claims. He can see that a preference claim of stakeholder A is, for example, three times more important than the competing preference claim of stakeholder B. Depending on whether the insight about reference point dependence is utilised, this comparison refers to

²⁵⁹ Think of a bundle of precious metals worth 1000 €. The bundle contains 100 g silver worth 100 € and 100g gold worth 900 €.

²⁶⁰ We can think of this as an exchange rate between different preferences: How many units of one preference would equal a single unit of another. The function would be like exchange rates of different currencies.

absolute levels of preference satisfaction (simple method), or the change in preference satisfaction measured from the current status quo (complex method).

3.4. Actual stakeholder preferences registered into payoff matrices

Having provided the means to interpersonally compare preference utility with the *preference index*, the next step is to describe how the different types of stakeholders feel about individual legislative actions that are suggested for the purposes of realising the legislative plan. The *preference index* required an objective assessment perspective, due to the contaminating potential of information regarding actual prior circumstances. The description of actual preferences is not subject to the same requirement. On the contrary, the legislator should make a point of describing the subjective point of view. Each subjective point of view represents a separate preference function. Each person sharing common priori for a specific legislative plan also shares a preference function for that specific plan.

By listing potential legislative actions in the rows of a matrix, and the potential states of the world in the columns of the same matrix, the payoff in terms of preference satisfaction can be described in the cells of the matrix to represent how satisfied a stakeholder is with each proposed legislative action, given a specific state of the world scenario. The precise utility of each action could be inferred in a numerical cardinal format by using Von Neumann's and Morgenstern's inference methodology. A separate payoff matrix could be assembled for each stakeholder type to get an overview of how the proposed legislative actions are received by individual stakeholder types.

Once information about how strongly different stakeholders prefer proposed legislative actions has been gathered (i.e. payoff matrices), and information about how important from an objective point of view the preferences of different stakeholders are (i.e. preference indexes), information about the preference utility of individual legislative options is within a derivative reach.

3.5. Identifying the fair and efficient legislative action

If we follow Tversky's and Kahneman's abstraction of decision making under risk,²⁶¹ we can say that the steps taken to establish the *preference index* and the payoff matrices belonging to the framing phase of the legislative decision making, in which the relevant information is gathered, processed, and often simplified to enable a subsequent valuation of the available decision options during the valuation phase, before the eventual selection of the optimal option.

For the purposes of the *model of fair rules*, we are working towards identifying legislative options that promote the normative anchor of the social contract, which was discussed in detail in chapter two. The rationale of the social contract can be summarised by saying that; individually we seek to be treated fairly vis-à-vis the others by claiming our share of the collective output, and as members of the collective we seek efficiency to maximise the joint output from which the individual share is derived. The maxims of fairness and efficiency are thus the normative anchors of the social contract.

When the legislator wields the power to negotiate the details of the social contract through his legislative plans and the legislative actions to implement the plans, he must carefully consider the duality of the contract's rationale; his legislative decisions must promote the general good, while simultaneously making sure that the right incentives from the individual point of view are in place. Different valuation methods are needed to assess these different optimisation aspects. Let us start with efficiency and then consider the fairness criteria.

Through the information gathered by constructing the matrices, we know what makes each stakeholder type tick. We know how each of the types perceives different legislative options in terms of magnitudes of preference satisfaction. The magnitudes represent a subjective view seen from the perspective of each stakeholder type. To make the subjective information useful for assessing efficiency, we need to be able to compare and quantify the different subjective perspectives, i.e. we need to be able to interpersonally compare and quantify the magnitudes of preference satisfaction for different types of stakeholders. A subjective representation of magnitudes can be turned into an objective representation by factoring in the relative

²⁶¹ See discussion in chapter three.

importance of each of the subjective perspective through the *preference index*. After adjustment, the numbers in each of the matrixes would represent the change in magnitude of preference satisfaction, considering the objective importance of the preference function of that specific stakeholder type. After adjustment, the numbers in different matrixes can be meaningfully compared in terms of objective magnitudes of preference fulfilment.²⁶²

The size, or the relative size, of the stakeholder segment under assessment is of key importance to determine the overall efficiency of each legislative option under consideration. The size could either be determined in real numbers, or as a percentage of the whole population of relevant stakeholders. For the purposes of manageability, it is useful to assume completeness of the diverging preference functions. This means that all members of the population must be assigned to one of the defined stakeholder types and not more.²⁶³

The total efficiency of each decision option is derived from three variables: the applicable subjective preference value, the relevant *preference index*, and the stakeholder segment size. The *preference index* (PI) is used to translate subjective preference values (SV) into objective preference values (OV), and the population segment size (P%) determines the number of individuals producing the said objective value of preference satisfaction. These variables enable the extraction of the total output of each stakeholder segment, and through aggregation with other stakeholder segments, the total output of all relevant stakeholders, for each of the legislative options that are being compared. We can call these numbers in their final format objective utility (OU). The legislative option with the highest objective utility score accordingly ranks as the most efficient legislative option.

Valuation of the fairness criteria requires an alternation of perspectives. Fairness should be identified subjectively from the perspective of the individuals involved; contrasting efficiency

²⁶² Think of two coin collectors: One collects gold coins and prefers them over other coins, the other silver coins and prefers them over other coins. Their subjective preferences are different due to their commitment to their respective collections. A way to represent a general preference view over these two subjective preferences, would be to find out the general exchange index of collections of gold and silver coins. By knowing the exchange index, we could perhaps establish that a certain magnitude of gold coins, one coin for example, is worth five coins of silver. The subjective preferences tell us what each want, but the objective preference tells us that a gold coin is five times more preferable than a silver coin.

²⁶³ All stakeholders S^n in the set X must also be members of one of the subsets; X^a , X^b , or X^c .

that is best approached objectively on utilitarian terms. An individual demand for equitable treatment vis-à-vis the others, constitutes the essence of the fairness criteria. The optimal legislative option should affect the preference fulfilment of each individual by an equitable magnitude. This should however not be taken to mean that each has a claim to receiving the same magnitude of his subjective preference fulfilment. We have already established that a subjective preference of a person has an objective value when compared with the preferences of others. Some hold objectively important preferences, while others hold less important ones. When making a claim for an equitable share, the relative importance of the claim in the objective sense, will determine the magnitude at which the individual feels that he is being treated fairly in the subjective sense.²⁶⁴

During the valuation of the fairness criteria, the legislator needs to translate the subjective preference values described in the matrices, into subjective values adjusted for relative importance. To achieve this, the legislator can use the *preference index*. The preference index shows the relative importance of a single magnitude unit of the distinct types of preferences involved. We can use this information to establish a ratio between the relevant preferences. The ratio indicates how much, in terms of magnitudes, is needed of each preference to achieve an equitable outcome.²⁶⁵

In practice, the legislator would need to be able to see, when comparing the same decision option across the matrices of different stakeholders, whether they perceive the consequences equitably. In terms of clarity, it would thus be ideal if the legislator would simply need to find numerical pairs in corresponding consequence cells of the different decision matrices. The

²⁶⁴ Think of two persons with different tastes. One has an expensive taste when it comes to wardrobe composition, while the other is happy with cheaper options. Imagine they work comparable jobs at the same company that annually allocated a small common budget for buying cloths for wearing at work. To fulfil their private (subjective) preferences for business attire to the same magnitude level, one would need Armani suits, while the other would be just as happy with Zara suits. Objectively speaking the Armani suit costs at least twice as much as the Zara suit. Given this difference in objective value, the one with the expensive taste would need to relax his claim to a fair share of the cloth budget. A claim of receiving the same magnitude of suits would not be perceived as fair by anyone, not even by the guy with the expensive taste. A fair magnitude ratio might be one Armani suit for you, for every two Zara suits for me.

²⁶⁵ If for example the preference ratio between two preferences P^1 and P^2 is $1 : 0.75$, we know that a single magnitude of P^1 is equally preferable to 0.75 magnitudes of P^2 . An equitable legislative option would satisfy the preference satisfaction of the population in these proportions; i.e. for every fulfilled magnitude of P^1 a 0.75 magnitude of P^2 would also need to be fulfilled.

pairs would indicate an optimisation in terms of equitable preference fulfilment between the different population segments and simultaneously indicate a solution to the fairness criteria. To achieve this clarity of representation for the purposes of valuating fairness, the numbers representing the subjective preferences and the *preference index* would need to be divided, instead of multiplied as when the most efficient option is identified. The result could be referred to as objective fairness score (OF) as opposed to objective utility (OU) for measuring efficiency of specific legislative options. Notably, the size of the stakeholder segment is irrelevant when determining fairness; fairness is perceived individually, while efficiency collectively concerns the entire organ at stake.

We can summarise the identification of the fair legislative options and the efficient legislative options in the following way where (SV) is subjective preference value, (PI) is *preference index*, and (S) is a stakeholder:

Inferring efficiency: $SV \cdot PI \cdot S\% = OU$

Inferring fairness: $SV/PI = OF$

The legislator now has information on the objective total utility of each legislative option, and he has information about how each legislative option is perceived in terms of objective fairness from the point of view of each stakeholder. To satisfy the maxims of the social contract, the legislator would simply need to identify the legislative options that have the most equitable distribution among stakeholders of the objective fairness score,²⁶⁶ and pick the one that has the highest total objective utility. This means that first, the fair legislative options are identified and then the most efficient of them is selected.

Through the method, I have now described,²⁶⁷ *the model of fair rules*, it should be possible to identify legislative options that achieve optimality in terms of fairness and efficiency. The fairness criterion ensures moral plausibility of the model's recommendation, while the

²⁶⁶ In theory, the preferableness of a specific legislative option should rank equally across different groups factoring in the objective value of their preferences. In practice, it could become necessary to allow a small range of variation in order not to exclude options that result in only marginal inequality of preference satisfaction.

²⁶⁷ The model is described here in abstract terms; for its practical application, a reference is made to chapter eight.

efficiency criteria ensures maximisation of overall preference satisfaction within the boundaries of the morally plausible. By using quantitative methods inspired by insights from expected utility theory and more generally from decision theory, the optimisation recommendation achieves more accuracy than the intuitive based proportionality test, and overcomes the problems associated with monetisation that symptomize the orthodox variants of cost benefit analysis.

4. The utility of a model for identifying fair rules

The orthodox approaches to legislative optimisation, the cost benefit analysis, and the proportionality test, depend on different modes of thinking about the role of the legislator. The deontological approach focuses on moral imperatives that should be obeyed irrespective of material consequences. The rigidity of the morally imperative makes the moral approach to laws unattainable in practice in its pure form, thus the relaxed standard of the proportional. The proportionality test, on the variant of Robert Alexy, slides into consequential methodology without adopting proper tools of quantification and without considering the full range of relevant consequences. The result is qualitatively and quantitatively inadequate recommendation, hedged on a moral standard equivalent to the Kaldor-Hicks efficiency test.

The economic approach to law is faced with an inverse problem. While claiming immunity from the morally imperative by focusing on measuring consequences, the regression to an ultimate moral choice is inescapable. In the case of Kaldor-Hicks or Pareto, the moral choice is inadequate. By using monetisation as a quantification tool, the cost benefit analysis also risks neglecting qualitative considerations that are not easily quantified in terms of money, such as moral imperatives, and invites a quantitative problem resulting from the material fact that people do not have equal means to show willingness to pay. The result is a recommendation that; neglects important qualitative concerns; is quantitatively eschewed in favour of those with material means; and based on a utilitarian moral standard.

The *model of fair rules*, consequential in approach, reduces morality in the context of laws to a single imperative demand; the laws ought to be fair. For the purposes of the model, fairness should be understood as a quantifying term in relation to specific desires of the people subject

to the laws. The legislator ought to fulfil the desires of his peoples to the greatest extent possible (i.e. efficiently), but he must do so in quantities perceived as fair from the individual point of view (i.e. fairly).

The model derives its rationale and normative force from the social contract vision of Niccolò Machiavelli and Thomas Hobbes, as explained in the modern terms of game theory, devote of John Rawls's reliance on Kantian thinking. The reduction of morality in the laws to a single quantitative term of egalitarian fairness, allows the reduction of every conceivable reason with relevance for legislative decisions, under the term of a preference. All that matters for the legislator is to examine what his people prefers, and make the best effort to optimise the realisation of these preferences in accordance with the fairness and efficiency maxims of the social contract. Inspired by microeconomic tools and decision theory thinking, the *model of fair rules* provides the legislator with a methodology for solving this task with practical ease and quantitative accuracy.

Having outlined in the abstract how rules can be optimised towards a concept of fairness based on the maxims of the social contract, we can start to think about how these insights can assist in approaching the concept of procedural fairness in the context of EU competition law. I assume that the general concept of fairness in the laws described in the preceding chapters is also applicable in the special case of procedural law; more specifically in the context of EU competition procedure. To apply the *model of fair rules* in this special context, we need to define the applicable inputs for these special purposes. The task of the next chapter is to examine the legislative plan of procedural legislation in general, which in the normative sense, should reveal the normative reasons that become the object of preference for the stakeholders in a procedural regulation, and in the instrumental sense should reveal the functional structure of the instrument of a procedural regulation. Understanding the instrumental and normative fundamentals about procedural rules is essential to enabling the design of a fair procedure based on the *model of fair rules*.

The essence of a procedure

Flosi asks Eyjólfur: 'Do you see any defence against these charges?'

'None,' says Eyjólfur.

'What council is now to be taken?' Flosi asks.

'Each available alternative seems problematic,' says Eyjólfur, 'but I shall nonetheless give you few advices.'²⁶⁸

The quoted conversation supposedly took place during the trial of Flosi Þórðarson in the summer of 1011 or 1012 at the annual legislative and judicial gathering of the Icelandic Commonwealth (930 AD -1262 AD). Flosi was being tried for having led a band of armed men that burned the farmhouse of the great lawyer Njáll Bergþórsson the year before, atrociously killing him and his extended family in the fire. The lawyer, Eyjólfur Bölverksson, went on to instruct the defendant Flosi on how, despite the complete absence of viable substantive legal defences, he could nonetheless machinate a defence on procedural grounds that might suffice to disrupt the case against him.

The conceptual distinction between the substantive and the procedural reflected in the conversation highlights a tension that modern legal systems struggle with in the same way as the courts of the Icelandic Commonwealth did in the early 11th century. While the procedural is supposed to facilitate the substantive, strict procedural formalism can nonetheless distort the substantive ends pursued.

Law enforcement as a process is inherently fragile due to the fine line that separates a procedure that facilitates the substantive, and a procedure that distorts the substantive. In an ideal world, we would want neutral procedures that perfectly facilitate the ideal substance. Unfortunately, the reality is not always ideal; the substance is corrupted at times, and procedures can distort and corrupt (for good or for bad) instead of facilitating. Paradoxically a

²⁶⁸ See *Brennu-Njáls saga* (~1270-1290) ch 141 <http://www.sagadb.org/brennu-njals_saga> accessed 3 January 2017. Loose translation into English from the original Icelandic text by HLK.

procedure can do justice by distorting, instead of facilitating, unjust substance. Symmetrically a procedure can do injustice by distorting, instead of facilitating, just substance.

Procedural fairness, or the synonymous term procedural justice, is the universal solution to the problem of the ideal procedure. A procedure is perceived as desirable if it is fair and conversely undesirable if it is unjust. The consensus on the solution to the problem of the ideal procedure unfortunately only extends to the vague terminology of procedural fairness. Interpretation of the practical implications from requiring procedures to be fair have led to disagreements about the level of emphasis on different considerations for designing the ideal procedural rule. The main disagreement is between the two orthodoxies for thinking normatively about problems of law. Should the requirement of procedural fairness be understood in consequential or deontological terms?

The focus of this chapter centres on two themes: firstly, to examine to what extent law enforcement procedures can be treated as a uniform legislative category in terms of the functional ends pursued through different types of such procedures; and secondly, to identify the normative rationale on which procedures are based. The approach is based on the hypothesis that functional consistency across different procedures, correlates with a consistent normative rationale.

1. The procedure as an instrument of law

Viewed instrumentally, a procedure simply refers to a specific mean aimed at bringing about a specific end. In the context of law, the substantive refers to the ends pursued, and the procedural refers to the means that facilitate the pursued ends. The difference between a mean and an end is often subtle in the context of law. A decision to pursue a specific end often implies a specific process that is so obvious and uncontroversial that it does not need to be spelled out explicitly.²⁶⁹ Similarly, a decision to follow a specific process can imply the

²⁶⁹ 90 km/hour speed limit at a specific road could be used as an example of a substantive end pursued, that implies a specific process or a mean of driving to achieve; if one drives carefully below the speed limit, one will achieve the end of staying below the articulated limit. The process is however so obvious that we do not need a specific procedural rule to facilitate the substantive end.

pursuance of some obvious consequences that likewise are taken for granted.²⁷⁰ Procedural rules in this way sometimes acquire quasi-substantive character, and substantive rules in the same way become quasi-procedural.²⁷¹

The blurred boundary between the procedural and the substantive in legal praxis requires a motivation for treating them as categorically separate instruments of law. In the absence of a specific reason, the optimal legislative instrument would arguable often be a synthesis of a mean specifically designed to achieve a specific end, or the other way around. Many, and even most laws are designed without a clear distinction between rules that serve as ends and rules that serve as means. In many cases, the imposition of such distinction would be superfluous; an end usually implies a specific mean, and a mean usually implies a specific end, which eliminates the purpose of making a distinction between the two.²⁷²

The need for a distinction between the substantive and the procedural arises once a certain level of complication is passed, and the importance of the procedural gradually increases with increased level of complication. If there is only one way to do things, or when only one thing will result from doing things, a distinction between an action and a consequence is unimportant. If, however several actions can lead to the same consequence, or if a variety of consequences can be achieved through a specific action, it becomes useful to distinguish between the mean and the end.

²⁷⁰90 km/hour speed limit could also be considered as a rule of procedure to achieve the end of careful driving. The process implies the end pursued so that the later does not need to be spelled out explicitly.

²⁷¹Solum discusses the blurred line between the procedural and the substantive in more details. See Lawrence B. Solum, 'Procedural Justice' (2004) 78 Southern California Law Review 181, 215-24. Sunstein also discusses the difficulty of engineering procedures that do not paternalistically affect the substantive ends. See Cass R. Sunstein, 'The Storrs Lectures: Behavioural Economics and Paternalism' (2013) 122 Yale Law Journal 1826, 1855-58.

²⁷²In philosophy reference is sometimes made to *Ockham's razor* as a rule of thumb. That implies that the simple is usually superior to the complicated, unless there is a specific advantage to the complication. Albert Einstein is often quoted for having said that *everything should be made as simple as possible, but not simpler*; echoing Ockham's razor and often referred to as Einstein's razor. This quote is believed to be derived from his remarks in a 1933 lecture: *'The basic concepts and laws which are not logically further reducible constitute the indispensable and not rationally deducible part of the theory. It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience.'* See the published version of the lecture; Albert Einstein, 'On the Method of Theoretical Physics' (1934) 1 Philosophy of Science 163, 165.

A separation of legal instruments for describing the substantive ends pursued, and the procedural means for achieving them, often occurs over the problem of enforcement. Usually the substantive criterion is explicit, but it can be hard to determine whether it applies to a specific factual event or not. The classic solution to such problems of legal application is an enforcement procedure, which determines the applicability of the substantive criteria to the factual situation at hand. These kinds of enforcement procedures are at the core of the category of legal instruments that legal professionals usually refer to as procedural rules, or procedural regulation.

Civil procedures, criminal procedures, and administrative procedures represent three archetypes of law enforcement procedures. Each of them exists in some form in most if not all modern legal systems in the western legal tradition.²⁷³ By looking for correlations in the macro function of these three archetypes of the procedural instrument, we should get closer to identifying their functional essence. The following analysis will consider the principal agents involved in each procedural type, and take note of their historical sources. An examination of the principal agents involved, and the historical pedigree of each procedural type, will provide several explanations for variations in the microstructure of modern law enforcement procedures, but will also reveal correlations in the macrostructure that is common to all of them; across jurisdictions, and across different forms of government through the history of the western civilisation.

1.1. Civil procedures as an instrument

Civil procedures deal with disputes of a private nature in which a public body is enlisted to decide on the rights and duties of the parties involved, and to enable their enforcement. Viewed in broad terms, the principal agents are the private parties, and the role of the state is usually limited to the role of dispute resolution. The stakes of the parties are their private interests, but the incentive of the state is the maintenance of an orderly functioning society.

²⁷³ I use the term *western legal tradition* to cover the Romanistic, German, Nordic, and the Anglo-American legal families as defined by Zweigert and Kötz. See Konrad Zweigert and Hein Kötz, *Introduction to Comparative Law* (Tony Weir tr, 3rdedn, Oxford University Press 1998) 63-74.

The key interests at stake and the principal agents holding them have historically influenced the particularities in the design of procedures in general and, along with few accidents of history, account for a great deal of the differences and the similarities in the modern design of different types of law enforcement procedures. The perception of a law enforcement procedure correlates with how the state and the state's power is perceived. In modern times, we perceive the state based on the social contract ideas of the enlightenment, and this perception serves as a background for how we perceive the role of the state in a law enforcement procedure; the state is a mean serving the ultimate end, i.e. sovereign citizen. Historically this has not always been the case, and even within the boundaries of the society of the social contract differences exist, as to how active or passive the role of the state should be. This can be demonstrated by a short review of the history of law enforcement procedures.

The gradual development of the civil procedure in classical Roman Law highlights the link between a trial procedure and the perceived role of the individual in society. Initially, during the times of the Roman Republic, trial procedures were presided over by a public official, *the praetor*, which was primarily concerned with the form and the correct presentation of the competing claims before a panel of judges consisting of private citizens (i.e. *legis actio* and formulary procedures). Later, as the power of the state gradually consolidated in the persona of the emperor, the trial procedures took a simplified form (i.e. *cognitio* procedure). The praetor, acting on behalf of the sovereign emperor, could under that procedure both prepare cases for trial and decide their outcome. The *cognitio* procedure was initially used by the commanders of the Roman armies in the provinces, but gradually became the standard in Rome as well.²⁷⁴ A shift in the perception of the sources of sovereign power enabled a differently designed and rationalised enforcement procedure.²⁷⁵

²⁷⁴ For details on the development of Roman civil procedure see Ernest Metzger, 'An Outline of Roman Civil Procedure' (2013) 9 Roman Legal Tradition 1, 17-29. See also George Mousourakis, *Roman Law and the Origins of the Civil Law Tradition* (Springer 2015) 154-57.

²⁷⁵ The change in the design of procedure followed a general shift in legal authority after the establishment of the Empire. The legislative role of the Roman senate gradually diminished and eventually seized during the early *principate* (27 BC – AD 284) in line with the increasing power of the emperor, which had become absolute during the *dominate* (AD 284 – AD 585). See Brad Inwood and Fred D. Miller, 'Law in Roman Philosophy' in Fred D. Miller and Carrie-Ann Biondi (eds) *A Treatise of Legal Philosophy and General Jurisprudence – Vol 6 – A History of the Philosophy of Law from the Ancient Greeks to the Scholastics* (Springer 2007) 136-39.

The Constantinople based Roman Emperor Justinian (527-565) sought to revive the best elements of Roman imperial law through the *Code* (529), the *Digest* (533), and the *Institutes* (533), in what later became known as the *Corpus Iuris Civilis*. The work elevated Roman imperial law in its sophisticated third century form, rather than in the contemporary degenerated sixth century form. After the collapse of Roman rule in the west, its legal legacy survived in the east through the Byzantine Empire. During the Italian Renaissance, an interest in Greek scholarship was revived in the west through Byzantine influences. Copies of Justinian's *Corpus Iuris Civilis* were discovered in Florence and Bologna, where they were subsequently studied. From there the *Corpus* spread as a subject of academic study to other leading universities in Western Europe.²⁷⁶

The subject of procedure was not systematically addressed in the *Corpus*, but in conjunction with canons from Gratian's *Decretum* and other canonical texts, legal scholars of the Catholic Church shaped a new Romano-canonical procedure that became the standard procedure in a Europe-wide system of ecclesiastic courts. Through influences from the *Corpus*, as interpreted by scholars in the Middle Ages, these procedures resembled the Roman *cognitio* procedure.²⁷⁷ As the European states grew stronger, more centralised, and increasingly complicated to govern, the monarchs sought advice in their law-making efforts to legal scholars that were already familiar with Roman and canon law that offered advanced legal tools superior to the Germanic customary and tribal law systems in place. By then the population of much of Western Europe was already familiar with the Romano-canonical law procedure through the Roman Catholic Church, which became the model procedure for much of these legislative projects.²⁷⁸ Louis XIV of France completed one such project in 1667 with the issuance of the

²⁷⁶ See Thomas M. Banchich, John Marenbon and Charles J. Reid, 'The Revival of Roman Law and Canon Law' in Fred D. Miller and Carrie-Ann Biondi (eds) *A Treatise of Legal Philosophy and General Jurisprudence – Vol 6 – A History of the Philosophy of Law from the Ancient Greeks to the Scholastics* (Springer 2007) 251-57; George Mousourakis, *Roman Law and the Origins of the Civil Law Tradition* (Springer 2015) 233-54.

²⁷⁷ See George Mousourakis, *Roman Law and the Origins of the Civil Law Tradition* (Springer 2015) 253.

²⁷⁸ See C.H. van Rhee, 'Civil Procedure: a European *Ius Commune*?' (2000) 8 *European Review of Private Law* 589, 594-95. See also R.C. van Caenegem, 'History of European Civil Procedure' (1973) in Mauro Cappelletti (ed), *International Encyclopaedia of Comparative Law – Volume XVI - Civil Procedure* (Martinus Nijhoff Publishers 1987) 11-17.

Grande Ordonnance de Procédure Civile, which later became the model for Napoleon's *Code de Procédure Civile* in 1806.²⁷⁹

Napoleon's civil procedure should be viewed in the context of the politics at the time; following the collapse of the monarchy of Louis XVI in 1789, and the failed reform attempts of the revolutionaries in the 1790s. The new procedure formed sort of a political equilibrium between contemporary progressive and conservative political agendas. It maintained the basic institutional structures established by the Ordinance of 1667, which satisfied the needs of the old establishment for continuity. By making several procedural reforms, most notably by increasing the transparency of the process and by providing the parties with greater control over the procedure, the reformed procedure also came to terms with the demands of the influential class of progressive bourgeoisie revolutionaries. Coincidentally Napoleon's civil procedure was subsequently adopted in large parts of the European continent following the conquest of the *Grande Armée*, and became along with the French *Code Civil* the default European standard for legislative reforms in the 19th century, despite Napoleon's ultimate military defeat.²⁸⁰

The Austrian *Zivilprozessordnung* of 1895 introduced new ideas into civil procedure following the rise of the working class as a political force that forecasted the eventual evolution of the 20th century welfare state. Through this reform, its author Franz Klein advanced the idea of the *sozialfunktion* of civil procedure in which not only the private parties to the dispute have a stake in its resolution, but also the whole society due to the inefficiency of disputes. This approach to civil procedure became the standard model for 20th century reforms in continental Europe.²⁸¹ Socialist civil procedure as practiced in the Soviet Union and its European satellite states during a part of the 20th century took this social function very seriously; allowing the

²⁷⁹ See C.H. van Rhee, 'Introduction' in C.H. van Rhee (ed) *European Traditions in Civil Procedure* (Intersentia 2005) 5-6.

²⁸⁰ For general on the influence of Napoleon's civil procedure code see C.H. van Rhee, 'The Influence of the French Code de Procédure Civile (1806) in 19th Century Europe' in: L. Cadiet and G. Canivet (eds) *De la Commémoration d'un code à l'autre: 200 ans de procédure civile en France* (LexisNexis Litec 2006) 129-65. See also R.C. van Caenegem, 'History of European Civil Procedure' (1973) in Mauro Cappelletti (ed), *International Encyclopaedia of Comparative Law – Volume XVI - Civil Procedure* (Martinus Nijhoff Publishers 1987) 87-103.

²⁸¹ See C.H. van Rhee, 'Introduction' in C.H. van Rhee (ed) *European Traditions in Civil Procedure* (Intersentia 2005) 11-14.

state to initiate civil claims on behalf of plaintiffs that possibly did not even have the desire to take legal action in support of their rightful claims.²⁸²

The common-law variant of the civil procedure developed differently; most significantly it never became subject to Romanisation to the same extent as the continental version of Germanic customary law. The Romans developed a procedure that fitted a perception of absolute imperial power that resonated with the post renaissance continental European monarchs, who perceived themselves politically in a similarly absolute way. The English common law system however managed to evolve from its primitive Germanic tribal origins into a comprehensive system of law that fitted a monarchy restricted by a politically strong aristocratic class. The relative absence of an absolute monarch power made the Roman imperial procedure less attractive, and supported an evolution of a common law civil procedure in which the parties retained more control and the role of the state was less proactive.²⁸³

As we can see through this short historical overview, the role of the state in civil procedures, very much depends on how the state and the state's power is perceived in the larger political context. If there are radical changes in how sovereign power is perceived politically, it tends to disseminate into the design of trial procedures, and the role of different actors in the trial.

Functionally, civil procedures consist of several procedural steps. The common-law system distinguishes between a pre-trial stage and a trial stage, while the procedure in most civil law systems consist of several preliminary hearings that cumulate in a separate main hearing.²⁸⁴ A procedure is initiated by a lawsuit that serves the purpose of informing a party of a claim

²⁸²See Mauro Cappelletti and Bryant G. Garth, 'Introduction – Policies, Trends and Ideas in Civil Procedure' in Mauro Cappelletti (ed), *International Encyclopaedia of Comparative Law – Volume XVI - Civil Procedure* (Martinus Nijhoff Publishers 1987) 16.

²⁸³ Many reasons and coincidences contributed to the result that the Germanic common law system survived in England, while the continental monarchs resorted to the more advanced Roman law system. Part of it was the political coincident that the Tudor monarchs resisted the adoption of Roman law on the grounds of religious politics and opposition to the Pope. It has also been pointed out that while Roman law was studied at Oxford and Cambridge, the practicing lawyers and judges were trained at the Inns of Court in London. See further George Mousourakis, *Roman Law and the Origins of the Civil Law Tradition* (Springer 2015) 272-77.

²⁸⁴ See Mauro Cappelletti and Bryant G. Garth, 'Introduction – Policies, Trends and Ideas in Civil Procedure' in Mauro Cappelletti (ed), *International Encyclopaedia of Comparative Law – Volume XVI - Civil Procedure* (Martinus Nijhoff Publishers 1987) 8-11.

against him. After having been duly summoned the receiving party, can either admit liability and thus avoid further judicial proceedings, or give notice of his intentions to contest the claims raised against him through further proceedings. Sometimes the parties request provisional remedies at the initial stage of a procedure. The purpose being to secure the effectiveness of remedies sought, which may be at a risk of depreciating during a procedure to determine the claim's validity. After the parties, have outlined their main claims and the remedies sought through a written pleading, the preliminary stage of the trial starts.

During the preliminary stage of the trial, the main legal and factual elements of the dispute are framed and the supporting evidences gathered and sorted out. The approach to this task differs between the common-law systems (e.g. discovery procedure, led by the parties) and the civil law systems (e.g. preparatory hearings and investigation under the control of the judge), but the purpose of framing the decision problem is common to both types of legal systems. At the main hearing, or the actual trial, the parties present their arguments and the supporting evidences in front of the deciding judges or juries. Subsequently the judges decide which facts have been established and which laws apply to the situation at hand.

In the US, a jury sometimes decides on the establishment of facts. The common-law system, especially the jury based US system, is more rigid in terms of how facts can be established. Elaborate rules of evidence facilitate a clear framing of the competing arguments to enable laypersons to assess easily the merits of the factual claims. In contrast, the judge in a civil law system usually has liberty to consider, or dismiss, various sorts of evidences irrespective of formalities; the presumption being that due to professional qualifications, the judge will be able to weigh correctly the relevance of various evidences without the aid of a detailed prescription.

The aggregated instrumental function of the typical civil procedure as described above is to facilitate the enforcement of claims by creating a venue where opposing arguments can be represented on equal grounds in front of a neutral arbitrator. The essential function is to enable the framing of the competing claims so that the arbitrator, undisturbed by irrelevant

rhetorical or factual noise, can assess and decide, on the factual and the legal merits of the substantive claims.

1.2. Criminal procedures as an instrument

The composition of the principal agents involved in a typical criminal procedure is different from a typical civil procedure. While the civil procedure is primarily concerned with correcting or compensating for tortious behaviour in the relations of private individuals, the criminal procedure is concerned with criminal behaviour directed against society.²⁸⁵ The perception of the harm and the enforcement objective involved increases the role of the state's organs in the procedure and lessens the control of the private parties involved, in comparison with a typical civil procedure. The principal agents in a criminal procedure are thus the state and the private individual accused, while the potential private victim of the crime has more of an auxiliary role.

The historical evolution of the criminal procedure is largely parallel with the evolution of the civil procedure.²⁸⁶ The shifting trends in the location and the concentration of sovereign power have had similar influences on the design of criminal procedures. Criminal law in the modern era has expanded over many categories of acts that in previous times would have been deemed a matter of private concern, best dealt with through a civil procedure. In early Roman history, for example, only acts such as treason and murder that seriously jeopardised the community were considered criminal. The list of acts that could be the object of a crime grew during the late Roman Republic parallel to the concentration of centralised state authority. Eventually special criminal jury courts were created and their procedure based on the Roman civil procedure. The jury courts survived into the early Roman Empire, but as the power of the emperor consolidated, the role of the jury diminished until they had been completely replaced

²⁸⁵ The distinction in modern law between tort law and criminal law can be traced back to the distinction in Roman law between *delictum* and *crimen* acts, which in the former involved harmful acts against other individuals, but in the later against the state or society at large. See George Mousourakis, *Roman Law and the Origins of the Civil Law Tradition* (Springer 2015) 159-60.

²⁸⁶ For a detailed historical analysis of the evolution of the European criminal procedure see A. Esmein, *A History of Continental Criminal Procedure with Special Reference to France* (John Simpson tr; Little, Brown, and Company 1913).

by a *cognitio extraordinaria* procedure presided over by the emperor and his deputies during the Dominate period.²⁸⁷

Criminal procedures of various sorts developed during the period of weak central authority in Western Europe after the fall of the Roman Empire, some influenced by Germanic tribal justice, while others retained some of the Roman imperial character, especially where influenced by the Romano-canonical procedure of the ecclesiastic courts. The emergence of the political entity of the secular nation state on the model of the Peace of Westphalia brought about consolidation of sovereign power once again, which quickly brought about ambitious legislative programs.²⁸⁸ The Romano-canonical procedure, which was influenced by the Roman imperial *cognitio* procedure, became the model for the criminal procedure ordinance of Louis XIV of France that came to effect in 1670; three years after the corresponding civil procedure ordinance became effective.²⁸⁹ The perception of sovereign power by the rulers of the late Roman Empire, as translated through the procedural regime, resonated with the monarchs of the emerging European nation states of 1600s and 1700s. Rulers that were able to consolidate effectively the power of the state under their own private rule thus favoured the Roman imperial inspired inquisitorial approach to criminal procedure.

The French Revolution and Napoleon's subsequent accent to power in Europe brought about procedural reforms in France that were spread across Europe during the wars of the early 1800s. Napoleons *Code d'instruction criminelle* of 1808 attempted to strike a balance between the inquisitorial criminal procedure of the old monarchy, and the English inspired accusatorial type of procedure favoured during the revolutionary period of the 1790s. The result was a reformed inquisitorial procedure with some accusatorial procedural elements.²⁹⁰

²⁸⁷ See George Mousourakis, *Roman Law and the Origins of the Civil Law Tradition* (Springer 2015) 159-91.

²⁸⁸ Henry Kissinger discusses the emergence of the sovereign state as the centre of gravity and an ultimate reference point following the Peace of Westphalia of 1648. See Henry Kissinger, *World Order* (Allen Lane, London 2014)

²⁸⁹ See further A. Esmein, *A History of Continental Criminal Procedure with Special Reference to France* (John Simpson tr; Little, Brown, and Company 1913) 183-286.

²⁹⁰ See J.R. Spencer, 'Introduction' in Mirielle Delmas-Marty and J.R. Spencer (eds), *European Criminal Procedures* (Cambridge University Press 2002).

The continental procedure is still under the influence of Napoleon's pragmatic legislative approach; his criminal procedural regime retained the inheritance from imperial Rome by preserving the central role of the existing institutional structures of the state, while simultaneously elevating the influence and role of the involved private parties in the criminal procedure in line with the enlightenment trends.²⁹¹

The format of a modern criminal procedure centres on the formal accusation made by the state against the defendant. Prior to the accusation, the authorities become aware of a potential criminal breach of the laws, either through a notification from a citizen, or through their independent observation. Following awareness of a potential breach, the state initiates an investigation to verify the facts and to identify the probable suspect. The exact institutional structure responsible for the investigation phase varies between legal systems, but the purpose is universally to enable a decision on whether to make a formal accusation. The threshold for making a formal accusation normally requires that the investigation has shown that a trial on the accusation will probably lead to a guilty verdict. The exact formulation of the standard varies slightly between different legal systems, but the probabilities of a trial outcome should lean in favour of confirming the accusation.

After the formal accusation, the trial is institutionally designed so that the defendant has an opportunity to challenge the accusation factually and legally, and the state must likewise show that the accusation is factually and legally accurate. The key element in the institutional design of the criminal trial is the neutral decision maker. In the civil law tradition, the neutral decision maker has an active investigatory role, while the common-law tradition relies more on the parties to feed information into a passive decision body in the form of a non-specialist jury or a professional trial judge. Much of the procedural details in each tradition are designed to accommodate the strengths and the weaknesses of the respective decision-making structure, and thus to facilitate a factually and legally correct decisional outcome.

²⁹¹ Sarah Summers discusses the effects of the Napoleon's *Code d'instruction criminelle* of 1808 on the development of criminal procedure. See SJ Summers, *Fair Trials: The European Criminal Procedural Tradition and the European Court of Human Rights* (Hart Publishing, Oxford 2007)

The instrumental essence of criminal procedures relates to the two focal points: the formal accusation, and the judgement. Leading up to these focal points, there are framing processes that are aimed at facilitating factually and legally correct decisions at the decision points. The framing processes are quite different in terms of institutional structure. The former process leading up to the formal accusation empowers the state with comprehensive investigatory tools to extract the necessary evidences to reveal the factual truth behind the suspected criminal activity. During this phase, the role of the suspect is passive but nonetheless subject to certain procedural guarantees that curb the extensive investigatory powers of the state. During the actual trial process, the more rigidly regulated framing phase aims at providing a level field for the accused to represent factual and legal objections to the accusation.

1.3. Administrative procedures as an instrument

Administrative enforcement procedures include procedures from various fields of law that have the commonality of involving a public organ as one of the principal agents acting in the virtue of itself as a guardian of specified public interests. The role of the public organ can either be passive in that it receives and processes complaints from the public about its own administration of public power, or it can take an active enforcement role in upholding some specific public interests.

Administrative procedures can be placed at the far end of a scale in relation to civil and criminal procedures, viewed from the principal agents involved. The scale I have in mind gradually slides from a civil procedure with two private parties as the principal agents, over to a criminal procedure with the state and the accused in principle role and the victim in auxiliary role, over to the other end where in the administrative procedure the principal agent is the state and the citizen. The focus of the scale moves gradually from the relation of the private parties in the civil procedure, over to the relation of the state and the citizen in the administrative procedure. The focus of the criminal procedure is mixed between these two extremes. Sometimes with focus on an individual victim of a crime, sometimes with focus on the collective interests of the community, and often with a simultaneous focus on both.

Administrative procedures do not have the same uniformity across different substantive fields as civil and criminal procedures have within their respective spheres. Administrative procedures within the same legal system can involve various types of decision bodies depending on the subject matter, and various procedural designs. The uniformity of the category of administrative procedures is thus not derived from an institutional design; the uniformity derives from adherence to certain principles of administrative conduct. This code of administrative conduct is taken to apply in abstract over different fields and is referred to under different, but essentially synonymous labels depending on the specific legal system. In the EU administrative system, there are *principles of good governance*,²⁹² in the UK, there is *natural justice*,²⁹³ and in continental Europe, there are the principles of the *rechtsstaat*.²⁹⁴

The principles of modern administrative procedure evolved out of bureaucratic decision mechanisms that evolved parallel to the rise of strong centralised states assisted by a system of elaborate public administration in the 1600s and the 1700s. The bureaucracy became the link between the sovereign ruler and the citizens. By delegating decision power to the bureaucracy, the ruler could focus on selected issues of great importance, while leaving smaller routine issues in the hands of administrative officials. The rise of bureaucracy created a need for consistency and foreseeability. From the ruler's perspective, the delegation of actual power could be minimised if the officials were obliged to act consistently within the general framework laid out by the sovereign. From the perspective of the citizens, the increased capacity of the ruler to enforce his will, through delegation of power, could be more easily tolerated if the obligations were consistent and predictable. In this way *rechtsstaat* compatible equilibriums of bureaucratic conduct, could naturally evolve in the pre-democratic monarchies of Europe.²⁹⁵

²⁹² See for example; Carol Harlow and Richard Rawlings, *Process and Procedure in EU Administration* (Hart Publishing 2014) 40-46.

²⁹³ See for example; H.W.R. Wade and C.F. Forsyth, *Administrative Law* (9thedn, Oxford University Press 2004) 439-449.

²⁹⁴ See for example; Albertjan Tollenaar and Ko De Ridder, 'Administrative justice from a Continental European Perspective' in Michael Adler (ed), *Administrative Justice in Context* (Hart Publishing 2009) 304-306.

²⁹⁵ Wade and Forsyth argue that the principles of natural justice are efficiency enhancing in the context of administrative procedure, both from the perspective of the citizens and from the perspective of the administrative officials. See H.W.R. Wade and C.F. Forsyth, *Administrative Law* (9thedn, Oxford University Press 2004) 440. Tollenaar and De Ridder also argue that the *rechtsstaat* principles can be explained based on incentives for both the citizens and the monarch rulers of the pre-democratic states of Europe. See Albertjan Tollenaar and Ko

The concept of the sovereign state, on the model of the Peace of Westphalia, predated enlightenment ideas about democracy. The sovereign state was thus initially synonymous with the sovereign monarch ruler. The idea of the sovereign state, however, survived the democratic revolution of the 20th century by replacing the monarch with a body of democratically elected representatives. The ultimate reference point within the legal systems, conceptually remained as the will of the sovereign ruler, but the identity of the ruler seized being the monarch and became the public. For the administrative procedure, the ascendance of democracy simply meant that a new master was being served. The interests of the state were still of primary importance, but the state was no longer synonymous with the persona of the monarch, but was instead the collective of the public.²⁹⁶

As mentioned above, the category of administrative procedures is instrumentally diverse and thus draws its commonality from adherence to a common set of procedural maxims that are pursued through various institutional designs. In the UK system, the principle maxims are minimalistic in character and have been identified as those underpinning the concept of *natural justice*; an impartial official should make administrative decisions in the UK, and those subject to the decision should have the chance to be heard.²⁹⁷ The *rechtsstaat* maxims are more elaborate and focus on the theme of legality; the state should always act based on an adequate legal source.²⁹⁸

A good example of the *rechtsstaat* approach is the right to good administration in Article 41 of the EU's Charter of Fundamental Rights.²⁹⁹ The article of the Charter lays out several maxims of a good administrative procedure: EU administration should act fairly, impartially, and within

De Ridder, 'Administrative justice from a Continental European Perspective' in Michael Adler (ed), *Administrative Justice in Context* (Hart Publishing 2009) 304-306.

²⁹⁶ Tollenaar and De Ridder argue that the French revolution can be regarded as a revolution against the abuse of power by the elites through the administrative system, and that the revolution marks a shift towards the promotion of public interest in continental administrative procedure. See Albertjan Tollenaar and Ko De Ridder, 'Administrative justice from a Continental European Perspective' in Michael Adler (ed), *Administrative Justice in Context* (Hart Publishing 2009) 316-17.

²⁹⁷ See H.W.R. Wade and C.F. Forsyth, *Administrative Law* (9thedn, Oxford University Press 2004) 440-41.

²⁹⁸ See Albertjan Tollenaar and Ko De Ridder, 'Administrative justice from a Continental European Perspective' in Michael Adler (ed), *Administrative Justice in Context* (Hart Publishing 2009) 309.

²⁹⁹ See Charter of Fundamental Rights of the European Union [2010] OJ C 83/391.

a reasonable period. The EU citizen has the right to be heard, have access to his file, and he is entitled to a decision motivated by adequate reasons. The EU citizen also has a right to compensation for damages caused by faulty administrative actions, and he has the right to communicate with the administration in any of the official EU languages. The EU has additionally issued a white paper describing five principles of good governance relevant for all levels of government with the EU Member States for the purposes of underpinning '*democracy and the rule of law in the Member States*'. These five principles are: openness, participation, accountability, effectiveness, and coherence. These principles are meant to compliment the already established EU law principles of subsidiarity and proportionality.³⁰⁰

The *natural justice* commitment dictates that decisions should be taken based on the relevant facts and the applicable laws, and thus not based on biased or arbitrary assessment of those factors. We can call this a legality requirement. The *natural justice* commitment also dictates that during the framing phase, prior to the actual decision, the subject of the pending decision should have the opportunity to give an input to clarify the factual and the legal situation and thus increase the quality of the framing phase. The *rechtsstaat* administrative procedural maxims, as they are presented in Article 41 of EU's Charter of Fundamental Rights, can also be linked to either: the quality of the framing phase (right to be heard, access to file, and communication in a preferred language); or to the legality of the actual decision (fair, impartial, based on reasons). The requirements of a reasonable procedural time and of a compensation for damages caused by faulty administrative actions are of a different character. They relate to the economic factors of the procedure, not to the quality of the decision or quality of the framing.

As an instrument, the administrative procedure is typically simpler and thus less costly than the typical trial procedure in civil and criminal cases. The state is usually granted with extensive investigatory and decisional powers, which are slightly curtailed by a general condition of legality and by the granting of the right of the citizen to provide a statement on the issue.

³⁰⁰ See A White Paper on European Governance [2001] OJ C 287/07-08.

Instrumentally the administrative procedure does much of the same work as criminal and civil trial procedures do; it frames the decision problem through an investigation where the subject of the decision should have the opportunity to provide an input, and eventually a decision on rights and duties is taken based on a factual and a legal assessment of the situation at hand.

1.4. The instrumental essence of a law enforcement procedure

The instrumental structure of the three types of law enforcement procedures analysed, reveal uniform salient features. Each of them contains one or more phases of framing, where information on the factual and the legal is gathered and organised. Each of them also contains one or more decision phases, which strive for objectively reaching decisions based on the factual and the legal evidences identified as relevant during the framing phases. The framing phase and the decision phase form the instrumental macrostructure of a civil procedure, a criminal procedure, and an administrative procedure.

Each of the three procedural categories contains a specific set of procedural maxims that form the instrumental macrostructure of each category, and at the level of individual procedural systems, the instrumental microstructure is revealed through specific regulatory provisions. If this analysis holds, it should be possible to identify the instrumental function of any procedural maxim in relation to either the framing or the decision phase of the procedure's essential function. The same should hold for every provision with a procedural character within individual procedural systems; functionally each rule should relate either, to the framing task of a procedure, or to the decision task, or sometimes to both.

The differences in institutional design of the three types of law enforcement procedures can, to a degree, be explained based on their pedigree in history and with reference to the shifting historical trends in how state power is perceived and exercised. Despite all the accidents, successes and coincidences of history, law enforcement procedures have generally followed the instrumental rationale that I have just identified with the three types of modern law enforcement procedures. At least since the times of the Roman Republic, law enforcement procedures have generally followed the pattern of an initial framing phase followed by a subsequent decision phase. The alterations in institutional design have been less about the

essence of the instrumental task of the procedure as such, and more about how and by whom it was best achieved. In a sense, this instrumental structure of a procedure should not be surprising; it correlates with and mimics how the human mind takes decisions involving high stakes. According to Kahneman, such decisions are taken based on deliberations about the situation at hand in which we try to identify the relevant information and to organise that information comprehensively to allow for a subsequent informed decision.³⁰¹ It would thus not be an unlikely hypothesis, that law enforcement procedures initially emerged in antiquity as an attempt to standardise complicated decision-making, based on how the human mind normally deals with such problems.

Institutional differences between different categories of law enforcement procedures, both within the same legal system and across different systems of law, can be explained with reference to the principal agents involved and how they are perceived. If the issue at stake is perceived as primarily concerning specific private individuals, the institutional design of the civil procedure is usually employed. If, however the issue is perceived as being between the authorities of the state and a private individual, the institutional design of the administrative procedure is often employed. The issues that are processed through a criminal procedure often involve instances of a private individual's breach against another individual, which through its severity becomes the issue of the state as well, or against collective state interest deemed too severe to channel through a simplified administrative procedure. Each of these procedural techniques seeks functionally to perform the instrumental task of framing and deciding, but due to differences in how the principal agents and the issue of concern is perceived, varying categories of procedures are considered suitable for each. The perception of the issue and the principal agents involved can change over time; an issue that was considered a matter of private concern becomes of public concern, and an administrative issue becomes a criminal issue. Despite changes in perception, the instrumental task of framing and deciding however remains the same. Institutional designs can be altered and issues moved between procedural tracks, but the core instrumental task always remains the same.

³⁰¹ Kahneman distinguishes between the process for taking instant automatic decisions, and decisions that are taken slowly based on a deliberate framing-decision process. See further Daniel Kahneman, *Thinking fast and slow* (Farrar, Straus and Giroux 2011) 19-109.

We can draw from this that although distinct types of law enforcement procedures serve the same instrumental purpose of facilitating the framing and the taking of decisions, they can be differently suited to the task depending on the nature of the principal agents and the substantive interests involved.³⁰²

A comparison of adjudicative decision making, with other modes of democratic decision-making, reveals an important functional element in a law enforcement procedure. Voting and negotiation are based on the subjective view of the decision maker; votes are casted based on subjective preferences, and settlements are reached based on the subjective preferences of the negotiating parties. By contrast, we expect adjudicative decisions to be based on objective considerations; judges should reach a conclusion based on a process of rational inference from the available information on the factual and the legal. The casting of votes can be sufficiently explained with reference to individual opinions, but court decisions need anchorage in an objective truth.³⁰³

It is also possible to view these different decision modes as two successive phases of democratic decision-making. Initially issues are settled contractually based on negotiation of preferences or competitively by voting between competing preferences. Subsequently the contract or the result of the voting becomes an objective object of consensus that forms an objective reference point in adjudicative decision making along with factual issues that can also be objectively verified. At the subsequent stage, subjective opinions are defeated through practical reasoning based on facts about what has already been decided. The three types of law enforcement procedures that we have been discussing fall within the subsequent sphere of decision modality. The instrumental objective of a civil, criminal, and administrative procedure is to frame the legal and the factual reality to facilitate an objective deduction of a rational decision. The object of rational inference being the laws as previously decided and an objective assessment of factual events.

³⁰² On this point see Paul Craig, *Administrative Law* (7th edn, Sweet & Maxwell 2012) 380-81.

³⁰³ Lon Fuller made this distinction in a draft paper, posthumously published. His focus is on different modes of participation for three different decision mechanisms for taking decisions, settling disputes, and defining relations between people. The three decision modes he identifies are; adjudication, contract, and voting. See Lon L. Fuller and Kenneth I. Winston, 'Forms and Limits of Adjudication' (1978) 92 *Harvard Law Review* 353, 363-72.

The appropriateness of a procedural instrument for a specific type of a law enforcement decision depends on the consequences pursued and the effectiveness of each instrument in the circumstances of application. Jerry Mashaw gives example of this, with his distinction between three instruments of administrative procedures that each is based on a different functional rationale. The *bureaucratic justice* instrument follows a technocratic decisional pattern of cost efficiency, consistency, and objective rigidity of application. The *professional treatment* instrument defers decisional discretion to professional experts for assessing each case based on subjective individual needs. The *moral judgement* instrument incorporates the traditional accusatorial model by deferring control over the procedure to the parties involved.³⁰⁴

Building on Mashaw's modelling of different procedural instruments, Robert Kagan suggests that the preferred decisional instrument is, to an extent, contingent upon the political trust, or mistrust that each institutional design enjoys. The level of trust and mistrust of each design can depend upon complex arbitrary factors. High faith in expertise may incline political elites to defer decisional powers to experts. Upon negative impression of such deference, the political elites may resort to greater control through rigid bureaucratic standards. If the high level of political distrust is also towards the bureaucracy, the political elite may prefer the adversarial legalistic procedure that minimizes the procedural control of administrative agencies. Kagan argues that political mistrust facilitates more '*formal, bureaucratic and adversarial*' procedural modes, which pressures individual officials towards '*legalistic style of applying rules and deciding cases*'. Depending on whether the political distrust is deserved, the resulting pressures can have either negative or positive effects on the decisional output.³⁰⁵

The framing and the decisional tasks have been identified as the essential instrumental features of a law enforcement procedure. The essential work a procedure should do is to organise and extract the relevant information to enable the rational deduction of the

³⁰⁴ See Jerry L. Mashaw, *Bureaucratic Justice: Managing Social Security Disability Claims* (Yale University Press 1983) 23-34. Michael Adler has developed Mashaw's framework further by adding additional categories of 'managerial, consumerist and market' administrative decision models. See Michael Adler, 'Understanding and Analysing Administrative Justice' in Michael Adler (ed), *Administrative Justice in Context* (Hart Publishing 2009) 147-51. See also Michael Adler, 'A Socio-Legal Approach to Administrative Justice' (2003) 25 *Law and Policy* 323; Michael Adler, 'Fairness in Context' (2006) 33 *Journal of Law and Society* 615.

³⁰⁵ See Robert A. Kagan, 'The Organisation of Administrative Justice Systems: The Role of Political Mistrust' in Michael Adler (ed), *Administrative Justice in Context* (Hart Publishing 2009) 161-180.

objectively best available enforcement decision. Both essential tasks can be carried out successfully or unsuccessfully to varying degrees. We have also seen that different designs of the framing and the decisional phases can be differently suited for taking a specific type of a decision. This hints at an underlying normative element that determines the degree to which the instrumental function succeeds.

2. The normativity of a procedure

The practical dilemma of normativity in the design of decision procedures can be explained with reference to decision problems in two popular games.

The game of chess takes place in a fixed environment after the initial decision on who plays the white pieces. The potential outcomes of the game are only three: white wins; black wins; or a draw. The rules of the game are fixed, the possibilities for making the next moves are fixed, and both players have complete information about the factual situation on the chessboard. If one were to assess all potential moves, only three moves ahead it would be about billion possibilities to assess. If one had the capability to assess each possible option within one millionth of a second (i.e. microsecond), it would take 16 minutes to assess a combination of three future moves.³⁰⁶ Obviously, such computational capacity is beyond any human mind, and so far, computers have not yet been able to solve chess by computing a combination of moves that will provide the optimal result with certainty. In respond to the impossibility of assessing all the available information, the top chess players apply decision techniques that approximate the optimal moves by simplifying the decision problem, and chess computer programs are designed to apply similar strategies.

The game of football also takes place in a fixed environment with fixed rules and the possibilities of permissible actions within the game are fixed by the rules of soccer. In professional games, a small team of referees ensures compliance with the rules of the game through a simple enforcement procedure. The primary referee assesses factual occurrences during the game, sometimes with the help of the assistant referees, and unilaterally decides how the rules should be applied to individual instances. Anyone who has ever played a game

³⁰⁶ See Claude E. Shannon, 'Programming a Computer for Playing Chess' (1950) 41 Philosophical Magazine 314.

of football, or watched it on television, knows that the referee sometimes gets things wrong. He can apply the rules wrongly, or he can make a factual oversight. If the referee were to increase the correctness of his decisions, he could do so by pausing the match during any moment of controversy and set up a trial on the pitch. He could allow the teams to argue their case and perhaps call in bystanders as witnesses. With the help of modern technology, a recording of each instance could also be reviewed during the pitch trial to help getting the facts straight. This kind of a decision procedure would undoubtedly reduce the risk of referee errors, but would in turn interrupt the flow of the game and diminish its entertainment value.

The simple game of chess highlights an information capacity problem that applies to the taking of almost any decision. The complete set of information is usually so vast that even the most powerful computers in the world struggle with its processing and need to resort to techniques of simplification for approximating the correct decision. The game of football highlights a problem with another type of decision constraints. Correct referee decisions may be important, but as the regular football spectator knows, it is not the only thing of importance. A quick and efficient decision procedure that occasionally makes errors in many practical cases may be preferable to a more cumbersome procedure that makes fewer errors.

The practical dilemma of normativity in a decision procedure is thus that even if all the relevant information is available, it may still be impossible in practice to assess it for the purposes of reaching the correct decision, and even if it is possible to process all the relevant information, it may nonetheless be undesirable to do so for practical reasons.

The quest for the normative foundations of a law enforcement procedure is the quest for the essential element that a law enforcement procedure ought to be pursuing. This quest can be approached through the traditionally contrasting deontological and consequentialist tracks of thinking.

2.1. The deontological view

The deontological approach to explaining law enforcement procedure, presupposes that there is some essential procedure to a procedure, i.e. some specific process is imperative to the

proper process of law enforcement. To facilitate the deontological approach, a distinction needs to be made between substantive justice that refers to the outcome of a procedure, and procedural justice that refers to how the procedure proceeds.³⁰⁷ The consequentialist approach, in contrast, does not need this distinction.

John Rawls and Ronald Dworkin elegantly illustrate of how the normativity of procedural law can be approached deontologically.³⁰⁸

John Rawls explained the role of the deontological argument through a comprehensive distinction between a procedure in the substantive sense, and in the *pure procedural sense*.³⁰⁹ Procedure in the substantive sense, on Rawls view, relies on an independent *a priori* criterion of what is being pursued through the procedure. The normative task of the procedure thus becomes to facilitate and achieve this outcome. If the pursued end is defined clearly and it is also possible to create a procedure that perfectly facilitates the end, there is a possibility to achieve a *perfect procedural justice* in the substantive sense. Rawls, however, argues that procedural justice in the substantive sense is often flawed due to the practical difficulties of building a procedure that flawlessly facilitates the desired end.

In case the idealised *perfect procedural justice* cannot be achieved, Rawls talks about *imperfect procedural justice* as the common practical alternative. He refers to the criminal trial as an example of such an imperfection. In the criminal trial, the criteria for the desirable outcome exists *a priori*, but the trial procedure gives no guarantee that the result will not lead to a false outcome.

Due to the practical difficulties of achieving the perfect, or near perfect procedural justice, Rawls introduces the notion of *pure procedural justice*. *Pure procedural justice*, in contrast with the terms of *perfect* and *imperfect procedural justice*, does not use the outcome of the procedure as a reference point for its success. The reference point of success instead becomes

³⁰⁷ See Michael Adler, 'Understanding and Analysing Administrative Justice' in Michael Adler (ed), *Administrative Justice in Context* (Hart Publishing 2009) 131-33.

³⁰⁸ For an extensive literature review on this issue see Lawrence B. Solum, 'Procedural Justice' (2004) 78 *Southern California Law Review* 181.

³⁰⁹ See John Rawls, *A Theory of Justice* (Harvard University Press 1971) 83-90.

the purity of the procedure to the procedure. If the procedure about how a procedure ought to be conducted is followed in a pure enough fashion, it becomes a sufficient condition for finding a procedure procedurally just.

Ronald Dworkin illustrates how the deontological approach can be employed to provide a rationale to civil and criminal procedures.³¹⁰ His argument starts by identifying a tension at the foundation of the well-known procedural maxim from criminal procedure: *'no one should be convicted for a crime that he did not commit'*. He asks whether it follows from this maxim that each citizen *'has a right to the most accurate procedures possible to test his guilt or innocence, no matter how expensive these procedures might be to the community as a whole?'*³¹¹ The identified tension, is between the ideal state of perfect accuracy of a procedure, and the practical reality of society's limited resources to invest in individual procedures and procedural instances.

Dworkin dismisses the use of a cost benefit analysis for finding a reasonable compromise between the accuracy and the cost of a procedure. In support of that dismissal, he refers to a distinction between *bare harm* that people feel through the imposition of a law enforcement penalty, and the additional *moral harm* that people feel when penalised unjustly. Dworkin claims that the *'injustice factor in a mistaken punishment will escape the net of any utilitarian calculation'* no matter how sophisticated.³¹² On his view, *moral harm* should be considered as an objective notion that occurs also, and especially, when no one knows or cares about it. This supposedly makes *moral harm* an impossible object of utilitarian quantification. Any efforts to utilise cost benefit analysis to strike the right balance between procedural accuracy and procedural costs are thus bound to result in a morally suboptimal result. Importantly, Dworkin recognises the unattainability of committing to the ideal of the perfectly accurate procedure

³¹⁰ See Ronald Dworkin, 'Principle, Policy, Procedure' in Ronald Dworkin, *A Matter of Principle* (Clarendon Press 1986) 72-103.

³¹¹ See Ronald Dworkin, 'Principle, Policy, Procedure' in Ronald Dworkin, *A Matter of Principle* (Clarendon Press 1986) 72.

³¹² See Ronald Dworkin, 'Principle, Policy, Procedure' in Ronald Dworkin, *A Matter of Principle* (Clarendon Press 1986) 80-81.

in world of scarce resources, where multiple types of potential moral harms exist in addition to the prospect of being unjustly convicted for a crime.³¹³

The practical impossibility of achieving the perfectly accurate procedure, and the difficulties of making a morally acceptable balance between accuracy and cost efficiency, lead Dworkin towards an alternative approach. Inspired by Rawls's pure procedural justice argument, he proposes two principles of *'fair play in government'*: firstly that *'any political decision must treat all citizens as equals'*; and secondly that *'a later enforcement of that decision is not a fresh political decision that must also be equal in its impact'*.³¹⁴ The principles defer the balancing problem to the political process, which Dworkin believes is well suited for deciding the relative importance of different moral harms.³¹⁵ From the two principles, Dworkin infers that people have two rights with regards to a criminal procedure: the procedure should *'attach the correct importance to the risk of moral harm'*; and the procedure should ensure *'consistent weighting of the importance of moral harm'*. These two rights, Dworkin claims, provide *'a middle ground between the denial of all procedural rights'* implicit in the utilitarian approaches, and *'the acceptance of a grand right to supreme accuracy'*.³¹⁶

Dworkin and Rawls both stress the unattainability of achieving the procedure that perfectly facilitates the consequences pursued through the procedure, and they both dismiss the alternative option of imperfectly facilitating the pursued consequences. Rawls on grounds of impracticality and Dworkin on grounds of an unacceptable moral outcome. The solution of both is to shift the focus away from the consequences of the procedure and over to the process of the procedure. Rawls stresses that the process of the procedure needs to be somehow *pure* while Dworkin suggest that the process needs to be *fair* and *equal*. Rawls does not elaborate further on what the pureness requirement implies in practical terms, but Dworkin elaborates further and identifies two procedural rights implicit in the fairness requirement.

³¹³ See Ronald Dworkin, 'Principle, Policy, Procedure' in Ronald Dworkin, *A Matter of Principle* (Clarendon Press 1986) 84.

³¹⁴ See Ronald Dworkin, 'Principle, Policy, Procedure' in Ronald Dworkin, *A Matter of Principle* (Clarendon Press 1986) 84-85.

³¹⁵ See Ronald Dworkin, 'Principle, Policy, Procedure' in Ronald Dworkin, *A Matter of Principle* (Clarendon Press 1986) 87.

³¹⁶ See Ronald Dworkin, 'Principle, Policy, Procedure' in Ronald Dworkin, *A Matter of Principle* (Clarendon Press 1986) 89.

The structure of the deontological argument to the normative requirements of a procedure, as represented by Rawls and Dworkin, reveals an inherent weakness to the deontological argument. The problem of normativity is about how a procedure ideally ought to be. Rawls and Dworkin evade the problem by ignoring the question of preferable consequences and focus instead on actions that are imperative to the performance of the ideal procedure. Thus, the ideal way to the original procedure becomes to follow an ideal procedure. This of course begs the question of how the secondary procedure then ideally ought to be. The typical deontological response could be either of two alternatives: either regress the problem down yet another level by saying that there is an ideal procedure, to the ideal procedure, to the original procedure; or stop the regression and identify a categorical imperative that represents an ultimate norm.

Categorical imperatives usually suffer from a lack of practical meaning, unless attached to specific practical consequences. If we perceive Rawls's *pure procedural justice* as a categorical imperative, this lack of meaning becomes evident. Dworkin's categorical imperative can be identified in the two equality principles of fair play in government, which he develops into two procedural rights that instruct how the consequences of a procedure should be balanced between the two primary consequences that affect the design of any procedure. Dworkin solves the inherent vagueness of the categorically imperative by reduction to the balancing of practical consequences. However, by doing that his deontological approach slides into the territory of the consequential, the orthodox variant of which he so strongly despised (i.e. the methodology of cost benefit analysis).

Rawls and Dworkin both recognise specific ideal consequences to a procedure. In Rawls's perfect procedural justice, the *a priori* criterion for the right result is perfectly achieved through the ideal procedure. Dworkin's ideal procedure achieves its mission with perfect accuracy. Both also agree that the ideal procedure, in terms of facilitating its purpose accurately, is often unattainable. Dworkin cites limited resources as a reason and Rawls the lack of a '*feasible procedure*'. They do thus both agree that in terms of the consequences of a procedure, a

balance needs to be reached between the accuracy of facilitating the procedure's purpose, and the cost of achieving accuracy.

The employment of a categorical imperative to dodge the balancing problem is not convincing; the imperative only becomes meaningful once it can be claimed as a right, and a claim for a right is only meaningful in practice if it is a claim for some specific consequences. In the case of procedural rights, a claim for a specific right becomes a claim for consequences that are balanced in a specific way. This specific balance in turn becomes the object of the right at stake, which in turn is the object of the categorical imperative at stake. The categorical imperative is thus just a proxy for a specific balance of the primary consequences at stake.

In the case of procedural justice, Rawls and Dworkin have identified procedural accuracy and procedural cost as the primary consequential concerns. Their inventions of *pure procedural justice* and the *two principles of government fair play*, however only become comprehensible and relevant in practice if seen as proxies for the ideal balance between these primary consequential concerns. Any rights deriving from the overarching principles of procedural justice thus become balancing maxims aimed at facilitating the ideal balance of the primary consequences at stake.

2.2. The consequentialist view

By focusing on the consequences pursued through a procedure, the normative foundations can be approached more directly. We have already established that the essential instrumental function of law enforcement procedures is to frame the factual and the legal to facilitate the taking of a correct enforcement decision. Different procedural elements have different tasks that relate either to the facilitation of the framing task of the procedure, or the decisional task of the procedure. We have seen that various techniques can be employed for both the framing task and the decisional task, but to assess which technique is the most appropriate for a given task we need to know which outcome would be considered optimal. The object of optimality forms the normative agenda of a procedure. A procedure ought to strive for optimality, however defined.

The quality of the outcome of a law enforcement procedure stands in relation to the substance of the substantive law it aims to facilitate. If a procedure succeeds in facilitating the outcome pursued by the substantive law, it has performed optimally. If a procedure is not successful in facilitating the substantive outcome desired, its performance is suboptimal. In simplified terms, a law enforcement procedure succeeds when it gets things legally and factually right, but fails when it gets things legally or factually wrong.

Assuming law enforcement procedures succeed most of the time in facilitating the correct substantive outcome, an absolute certainty thereof can nonetheless be hard, or impossible to establish. The plainest circumstance can easily become subject to an insurmountable factual uncertainty due to a simple variance in the perception of factual events. What a procedure can hope to do in such circumstances, is to facilitate the substantive outcome pursued with optimistic chances of being right. A procedure with high probabilities of getting things right is thus closer to optimality than a procedure with low probabilities of getting things right, other things being equal.

Scarce enforcement resources further compound an occasional impossibility of establishing the factual situation in a procedural scenario. Skilled investigators and qualified decision institutions can, if given sufficient time, increase the probabilities of reaching a correct decision. The time of such professionals and institutions is however, a scarce resource that is not in supply without an investment in professional training and institutional maintenance. The substantive law that a procedure aims to facilitate, applies to an undefined number of instances that can be estimated by economic planners from year to year. Given that the resources available to economic planners are limited, the resources available for facilitating a substantive legislative plan are also limited. This strains the quantity of resources that can be spent on getting each procedural instance right.

In a world of unlimited resources, hordes of investigators and highly trained decision-makers can be hired to ensure the best available procedure in terms of facilitating the correct substantive result. In this world, the procedures that give equivalent chances of a correct result are considered equivalent in terms of the level of optimisation, notwithstanding different cost

levels. In a world of limited resources, the intensive use of resources on a procedural instance means that less will be available to optimise other instances. The implication that follows is that procedures with analogous probabilities of a correct result can additionally be ranked in terms of optimality based on their level of resource drainage. Thus, in a world of unlimited resources only the accuracy of the result matters, but in a world of limited resources both accuracy of the result and the efficiency with which the accuracy is achieved matter for determining optimality.

The leading consequential approaches, on the issue of normativity in law enforcement procedures, implicitly assume a world of limited resources and thus identify accuracy (i.e. cost of error) and procedural costs (i.e. cost of administering the procedure) as the primary considerations in procedural design.³¹⁷ Using these approaches, the optimal procedural design can be identified by finding the efficient balance between error costs and administrative costs, using the cost benefit analysis tool.³¹⁸

The maximisation of procedural accuracy relative to procedural costs provides a plausible optimisation standard for pursuance in procedural design. At least plausible enough, to rival the alternative approach of using deontological procedural maxims. In a world of scarce resources, the use of imperative procedural standards becomes problematic due to the absence of a specific target of optimisation. If a compromise must be made between the imperative standard of accuracy and the practical concern of procedural cost, the deontological approach struggles with identifying where the optimal balance should be struck between the two, while the consequential approach can refer to the point of maximum efficiency.

³¹⁷ Richard Posner identifies the objectives of civil and criminal procedure from an economical perspective as being the minimisation of the cost of error and the cost of operating the procedural system. See Richard A. Posner, *Economic Analysis of Law* (6thedn, Aspen Publishers 2003) 563. Gordon Tullock mentions that there 'has been an assumption that we want to lower cost and raise accuracy' of procedures. See Gordon Tullock, *Trials on Trial: Pure Theory of Legal Procedure* (Columbia University Press 1980) 70. Robert Cooter and Thomas Ulen assume that 'the economic objective of procedural law is to minimize the sum of administrative cost and error cost.' See Robert Cooter and Thomas Ulen, *Law & Economics* (6thedn, Pearson 2012) 385.

³¹⁸ See for example; Robert Cooter and Thomas Ulen, *Law & Economics* (6thedn, Pearson 2012) 384-86. See also a similar approach based on the so-called 'Hand formula'; Richard A. Posner, *Economic Analysis of Law* (6thedn, Aspen Publishers 2003) 563-64.

In a world of scarce resources, procedural accuracy must be considered as a quantitative term regardless of whether it is used in the deontological sense or the consequential sense. In the absolute sense, accuracy only exists in a world of unlimited resources. Procedural standards in the deontological sense thus essentially communicate a degree of accuracy, not accuracy per se. The quantity of accuracy can be expressed informally on an ordinal scale through natural language (e.g. proof beyond reasonable doubt), or formally on a cardinal scale through the language of probabilities (e.g. 90% changes of an accurate result). While the consequential approaches can refer to a cost benefit analysis in support of their choice of preferable accuracy level, given the availability of a specific set of resources, the deontological approaches have difficulty with explaining why, for example, the accuracy standard of *beyond all reasonable doubt* should be used and not what seems to be the morally superior standard of *beyond all doubt*.

The deontological approaches face another deep problem of normativity that does not affect the consequential approaches. We have seen that they have problem with explaining based on which normative considerations the trade-off between procedural accuracy and procedural costs should be made. In addition, they have difficulties with defining optimal accuracy once the possibility of absolute accuracy is removed. Given that perfect accuracy cannot be achieved due to material restraints, each procedure that ends with a decision implies a probabilistic possibility of a decision error that can take either of two forms: a false positive and a false negative (i.e. type one and type two error). Logically the only way to avoid making type one error is to always reach a negative decision, and conversely the only way to avoid making type two error is to always take a positive decision. The dilemma of each decision maker, in a world where errors will be made, becomes to balance the risk of making an error of each kind.

If the objective baseline probabilities that a positive decision is correct, are equal to the probabilities that that a negative decision is correct, the risk of making a type one error and a type two error is equal. If, however, using the same baseline probabilities, a procedure requires that to reach a positive decision the probabilities of that being the correct decision need to be higher than 50%, the risk of making type one error disappear, but the risk of making type two

error increase by the same margin. The baseline risk of making an error is unchanged at 50 %, but the procedural requirement eliminates false positives by increasing false negatives.

If, however the baseline probabilities change so that the chances of a positive decision being correct are 75% and the chances that a negative decision is correct are 25%, the choice of the allocation of risk between type one and two errors becomes more urgent. To reduce the number of overall errors the logical strategy would be to always take a positive decision and thus purposely make type one error in one out of four decisions, while three out of four would be correct. If, however a procedure would require that a positive decision should only be reached if the baseline probabilities exceed 90% the decision maker would need to reach a negative decision on each occasion. Thus, type one errors would be eliminated, but the number of overall errors would rise to three out of four decisions being wrong, all of which would be type two errors.

In criminal procedure, a common procedural maxim dictates that people should not be found guilty unless there are high probabilities (i.e. more than 50%) of guilt. This signals the decision maker that he should emphasise on not making type one errors, if in doubt the default option should be a negative decision. The logical result from this procedural design is increased number of total errors, but the risk of making type one errors is less than making type two errors. An increase in the sheer numerical volume of errors is however often justified. A reduction in the risk of making few serious mistakes by increasing the chances of making more numerous minor blunders could be considered as preferable, both from a moral and from a cost benefit perspective. The allocation of risk between type one errors and type two errors requires the weighing of importance of each type. The consequential approaches can simply identify the point of the lowest aggregated error cost from both types as the optimal risk balance. The deontological approaches typically refer to intuitive standards. Such as the well know maxim from criminal procedure, that it is far worse to convict an innocent person than it is to acquit a guilty person, which implies that the error risk profile should be tuned to minimize the risk of type one errors.³¹⁹

³¹⁹ Larry Laudan argues that the criminal trial is strongly biased in favour of minimizing type one errors; *'While trials, in theory, are designed to find out the truth about an alleged crime, the vast majority of procedural and evidence rules are designed to protect innocent defendants from wrongful conviction. Almost invariably, the rules*

Dworkin criticised the use of cost benefit analysis to identify the optimal error risk profile due to the immeasurability of the moral harm associated with type one errors.³²⁰ The alternative of using intuitive standards for determining the error risk profile can however also lead to counterintuitive consequences. It has for example been pointed out that while the procedural maxim, that people should be allowed to state their case during a proceeding, can decrease the chances that a type one error is committed and thus increase accuracy, this maxim also invites a guilty person to mislead and confuse the proceeding and thus increase the chances of type two errors.³²¹ It has also been suggested that adherence to the criminal procedural standard of *proof beyond reasonable doubt* has the empirically verified consequence in the United States, that *'in any given year, one is six-times more likely to be the victim of a violent crime committed by someone falsely acquitted than one is to be the victim of a false conviction for a violent crime.'*³²²

The consequential approach to law enforcement procedure relies on defining cost and benefits of different procedural options. Given that the instrumental task of a law enforcement procedure is to facilitate the substance of substantive law, the ultimate objective becomes to do so accurately. In a world of limited resources procedural errors will however occur which adds an efficiency consideration into the quest for optimality. The efficiency consideration exists at two levels: on one hand, a balance needs to be reached between achievable accuracy considering the available resources (e.g. the compromise a chess grandmaster needs to make when he makes a move), and on the other hand, an efficient balance needs to be reached

designed to shield the innocent from conviction and related harms have the unintended but undeniable consequence of preventing many of the guilty from receiving their just deserts.' See Larry Laudan, 'The Rules of Trial, Political Morality, and the Cost of Error: Is Proof Beyond Reasonable Harm Doing More Harm than Good?' in Leslie Green and Brian Leiter (eds), *Oxford Studies in Philosophy of Law: Volume I* (Oxford University Press 2011) 196.

³²⁰ See Ronald Dworkin, 'Principle, Policy, Procedure' in Ronald Dworkin, *A Matter of Principle* (Clarendon Press 1986) 81.

³²¹ Gordon Tullock points out that in an adversarial proceeding *'a great deal of the resources are put in by someone who is attempting to mislead.'* See Gordon Tullock, *Trials on Trial: Pure Theory of Legal Procedure* (Columbia University Press 1980) 96.

³²² See Larry Laudan, 'The Rules of Trial, Political Morality, and the Cost of Error: Is Proof Beyond Reasonable Harm Doing More Harm than Good?' in Leslie Green and Brian Leiter (eds), *Oxford Studies in Philosophy of Law: Volume I* (Oxford University Press 2011) 202.

between the types of errors the material restraints impose (e.g. the acceptance of small referee errors in a football match to avoid interrupting the normal flow of the game).

2.3. The normative essence of a procedure

Both of the traditional approaches to explaining the normative essence of a procedure agree that the ultimate *ought* of a procedure is to facilitate the prescribed criteria for desirable consequences. The perfect procedure executes that task with perfect precision. Both approaches however also recognise that perfection is hard to achieve in a world of limited resources. The approaches disagree in how a compromise ought to be reached in the balancing between the material restraints imposed by practical circumstances, and the primary agenda of a procedure to accurately execute its facilitating task.

Using cost benefit analysis for locating the optimal balance between accuracy and cost efficiency comes with the usual caveats of over emphasising on the monetarily quantifiable at the cost of intangible variables, which Dworkin refers to as the element of moral harm. Using procedural maxims of the deontological type to strike the balance has the advantage of being able to include the element of moral harm, but comes with the risk of neglecting the important consideration of cost efficiency, which can result in wildly counterintuitive consequences.

Before concluding on the issue of normative essence there is one more consideration that deserves attention and that we can use as an example to illustrate the implication of reducing the normative essence of a law enforcement procedure to the consideration of accuracy as restricted by material restraints, with which it executes its instrumental function. In his extensive review of the literature on procedural fairness, Lawrence Solum concludes that accuracy and cost efficiency are important normative elements in a procedure, but adds that participation of the affected parties in a civil procedure should be included as an independent imperative consideration.³²³ Solum relies on a thesis of participation legitimacy derived from Jürgen Habermas's ideas about communicative action.³²⁴

³²³ Solum restricts his claim to civil procedure and explicitly excludes criminal procedure. See Lawrence B. Solum, 'Procedural Justice' (2004) 78 Southern California Law Review 181, 240 and 320-21.

³²⁴ See Lawrence B. Solum, 'Procedural Justice' (2004) 78 Southern California Law Review 181, 267-73.

Working from the premise that participation in a procedure is a function of its legitimacy, Solum refuses the type of reduction to accuracy I have just suggested.³²⁵ This type of an argument is however impaired by the fact that participation is by no means a necessary condition, nor a sufficient condition for the legitimacy of a procedure, unless a priori criteria dictates it should be so. If that is the case, then participation becomes just another substantive element that the procedure tries to facilitate accurately. The argument can be viewed as a variant of the *there is a procedure to the procedure argument* which diverts attention from the target of explanation by creating a new supplementary target, in his case the concept of legitimacy.

If we push the participation argument a bit further and compare how it fits with the instrumental macro function of a law enforcement procedure, that I have identified as being to frame and decide based on the principle of rational inference, the essentiality of participation becomes even less plausible. Lon Fuller's distinction of the different modes of participation in different types of social decisions captures this point neatly. He emphasised that the expectations we hold towards adjudicatory process, is a decision based on rational inference from the available facts. In such decisions, participation is limited to the presentation of relevant arguments and facts that have the potential of facilitating the rational inference. This mode of participation, Fuller contrasts with voting in elections, where the decision maker is by no means expected to cast his vote based on any rational inference, in which case participation per se becomes an independent consideration.³²⁶

The value of participation in a voting procedure is categorically different from the value of participating in an adjudicatory law enforcement procedure. In the former, the object of the procedure is to know the preference of the voter and thus participation becomes a central purpose and a main instrumental function. A law enforcement procedure can by contrast be conducted through proxies, and is in fact usually conducted in that way with the aid of professional litigators and prosecutors. The revelation of the subjective opinions of the parties is not the objective of the law enforcement procedure; the instrumental objective is to make

³²⁵ See Lawrence B. Solum, 'Procedural Justice' (2004) 78 Southern California Law Review 181, 291-95.

³²⁶ See Lon L. Fuller and Kenneth I. Winston, 'Forms and Limits of Adjudication' (1978) 92 Harvard Law Review 353, 363-64.

a rational inference from the relevant arguments and facts. The value of party participation in a law enforcement procedure is thus limited to the extent that it assists, or potentially assists, with this instrumental objective.

Participation per se can easily frustrate the process of rational inference and for that reason, procedures are often designed to limit participation with such potential. In the context of law enforcement procedures, participation thus only has a value if it potentially facilitates an accurate result based on rational inference. In that case, it can however also be reduced to an accuracy consideration, by which it loses any claim to a status as an independent value, which it may hold in a procedure with a different instrumental purpose such as a voting procedure.

3. A procedure is ...

The elusive concept of procedural fairness casts its shadow over the essence of any law enforcement procedure. Somehow, a procedure cannot claim to be properly designed unless it comfortably fits with the notion of being procedurally fair, or just. By examining the macro structure of the three main types of law enforcement procedures, a functional correlation appears; each procedure has a farming function and each has a decision function, which operate successively in accordance with a principle of rational inference. By examining the normative essence of such procedures, the key component of accuracy appears. Given that enforcement resources are scarce, a balance needs to be reached between the ultimately optimal and the practically possible. This problem of balancing occurs at two levels: first, a balance needs to be reached on how much should be spent on accuracy; and second, a balance needs to be reached on the type of errors conceded. The optimal solution to this balancing exercise leads to the concept of procedural fairness; if the balance is right, the procedure passes the test of being fair, or just.

The interpretation of what it entails to find a fair balance between the primary procedural considerations of accuracy and cost efficiency diverges at the familiar crossroad of the deontological and the consequential. Just as in other domains, neither approach can give the other approach a decisive blow in terms of providing a superior mould for balancing. The normal objections to each approach apply within the domain of procedural fairness, just as in

other domains where moral and efficiency considerations are at odds. Using Dworkin's terminology, the deontological approaches on average give greater weight to *moral harm* in its balancing, while conversely the consequential approaches on average give greater weight to *bare harm* in its balancing. In different terms, the argument boils down to a disagreement about which scale to use in measuring the optimal balance between procedural accuracy and procedural costs. The deontological scale tends to give imperative value to the reduction of type one errors, while the consequential scale tends to disregard or discount the intangible moral factor especially associated with the reduction of such errors compared with the reduction of type two errors.

Instrumentally a law enforcement procedure consists of a framing phase and a decision phase. Normatively a law enforcement procedure aims at accuracy and cost efficiency within the parameters of the two instrumental phases. Based on these observations about the essential instrumental and normative elements of a law enforcement procedure it can be inferred that instrumentally each procedural rule has either a framing or a decision function, and that normatively each procedural rule seeks to optimise the balance of accuracy and cost efficiency. The task of the architect of a procedural rule is frustrated by the lack of a practical standard of a balancing result, usually referred to as the standard of procedural fairness or procedural justice.

Architects of law enforcement procedures are usually guided by lexically superior procedural maxims that command how a procedure preferably ought to be. Article 41 (*right to good administration*) and Article 47 (*right to an effective remedy and a fair trial*) of the EU Charter of Fundamental Rights are examples of a collection of such procedural maxims. If my presentation of the essential elements of a procedure holds, the function of each of the procedural maxims articulated in these two Charter articles should be either to facilitate the framing or the decision phase of a procedure, and the normative purpose of each of the maxims should be to facilitate the fair balance between accuracy and cost efficiency of the procedure. The tables below show how this measures up:

Article 41 of the EU Charter of Fundamental Rights

- Right to good administration -

1. Every person has the right to have his or her affairs handled impartially, fairly and within a reasonable time by the institutions, bodies, offices and agencies of the Union.

2. This right includes:

(a) the right of every person to be heard, before any individual measure which would affect him or her adversely is taken;

(b) the right of every person to have access to his or her file, while respecting the legitimate interests of confidentiality and of professional and business secrecy;

(c) the obligation of the administration to give reasons for its decisions.

3. Every person has the right to have the Union make good any damage caused by its institutions or by its servants in the performance of their duties, in accordance with the general principles common to the laws of the Member States.

4. Every person may write to the institutions of the Union in one of the languages of the Treaties and must have an answer in the same language.

The substantive elements of Article 41 and their rationale	Functional Task	Normative Task
<i>Right to fair case handling</i> <u>Rationale</u> > substantively vague balancing rule that affects both functional tasks	Framing and decisional	Accuracy and efficiency
<i>Right to impartial case handling</i> <u>Rationale</u> > decisions should be taken based on a rational inference from objective facts	Framing and decisional	Accuracy
<i>Right to expeditious case handling</i> <u>Rationale</u> > reduce the cost of uncertainty for the citizen	Framing and decisional	Efficiency
<i>Right to be heard prior to adverse decisions</i> <u>Rationale</u> > chance to provide relevant input	Framing	Accuracy
<i>Right to have access to his file</i> <u>Rationale</u> > chance to verify the quality of the case file which puts an accuracy pressure on the framing phase and pressures decision makers to make a rational inference during the decisional phase	Framing and decisional	Accuracy
<i>Right to hear the rationale of a decision</i> <u>Rationale</u> > pressures decision makers to make a rational inference	Decisional	Accuracy
<i>Right to damages for faulty administrative actions</i> <u>Rationale</u> > pressures decision makers to make a rational inference to avoid liability and reduces the cost of the citizen	Decisional	Accuracy and efficiency
<i>Right to communicate in a preferred language</i> <u>Rationale</u> > reduces the cost of the citizen in providing input during the framing phase.	Framing	Accuracy and efficiency

Article 47 of the EU Charter of Fundamental Rights

- Right to an effective remedy and to a fair trial -

Everyone whose rights and freedoms guaranteed by the law of the Union are violated has the right to an effective remedy before a tribunal in compliance with the conditions laid down in this Article.

Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law. Everyone shall have the possibility of being advised, defended and represented.

Legal aid shall be made available to those who lack sufficient resources in so far as such aid is necessary to ensure effective access to justice.

The substantive elements of Article 47 and their rationale	Functional Task	Normative Task
<i>Right to a fair hearing</i> Rationale > substantively vague balancing rule	Framing	Accuracy / Efficiency
<i>Right to a public hearing</i> Rationale > the public's sense of justice pressures for an adequate quality standard in preparing decisions	Framing	Accuracy
<i>Right to a hearing within a reasonable time</i> Rationale > reduce the cost of uncertainty for the accused	Framing	Efficiency
<i>Right to a hearing by an independent tribunal</i> Rationale > decisions should be taken based on a rational inference from objective facts	Decisional	Accuracy
<i>Right to a hearing by an impartial tribunal</i> Rationale > decisions should be taken based on rational an inference from objective facts	Decisional	Accuracy
<i>Right to a hearing by a tribunal established a priori by law</i> Rationale > decisions should be taken based on a rational inference from objective facts	Decisional	Accuracy
<i>Right to a possibility of being advised</i> Rationale > assists the decision maker in establishing the truth	Framing	Accuracy
<i>Right to a possibility of being defended</i> Rationale > assists the decision maker in establishing the truth	Framing	Accuracy
<i>Right to a possibility of being represented</i> Rationale > assists the decision maker in establishing the truth	Framing	Accuracy
<i>Right to a legal aid in cases of necessity for accessing effective justice</i> Rationale > assists the decision maker in establishing the truth	Framing	Accuracy

The procedural maxims described here above are represented as imperative commands; a procedure ought to proceed in this or that way to qualify as fair. However, as we have seen,

these commands are auxiliary to either of the two instrumental functions of a procedure, and the action recommended through the command pursues the end of optimising these functions in accordance with the primary normative objectives of a procedure. Each of the procedural maxims can thus be considered as an action proxy for the optimal consequences pursued through the procedure. Compliance with an action command in the form of a procedural maxim does thus not guarantee an optimal consequence; the procedural maxim simply recommends an action that has, according to empirically untested conventional wisdom, favourable probabilities of facilitating the optimal consequence pursued.

The procedural maxims articulated in primary legislation with constitutional function provide the architects of law enforcement procedures with rough boundaries of the permissible, and an approximate guidance of the optimal, in the design of the more detailed aspects of a specific procedure. At the first level, a decision needs to be made about the baseline accuracy of the procedure, i.e. the level of acceptable errors. Should a specific level of accuracy be treated as a constant irrespective of costs, or should the available resources be treated as a constant towards which the accuracy level should be adjusted? At the second level, a decision needs to be taken on how the errors conceded through the first decision should be distributed between type one errors and type two errors. The procedural maxims often give guidance about the preference of reducing one type over the other type, but the exact probabilistic distribution is still usually at the discretion of the architect of the procedure.

The architect of a law enforcement procedure can resort to deontological tools and consequential tools to balance a procedure with regards to the first level balancing. The outcome of using each is however likely to be biased in favour of either cost-benefit considerations, or moral considerations, depending on which tool is used. The result would be either an over investment in the accuracy of procedures with moral implications, or under investment in procedures with moral implications. The architect of the procedure can resort to the same types of tools for the second level balancing of the types of errors conceded, again risking an over- or under emphasis on errors that have moral implications.

In view of the limitations of the orthodox balancing tools, I suggest the use of the *model of fair rules*, developed in chapter four, for achieving the optimal balance between an adequate accuracy and an adequate cost of a law enforcement procedure. Once the primary consequences pursued through a procedure have been defined in the terminology of preferences, the model can be applied with ease for identifying the optimum equilibrium point of these preferences. This equilibrium point is, as we established in previous chapters, the fair solution to the balancing dilemma. The fair solution is in turn the point of optimal procedural fairness.

Procedural fairness understood in the terms of the *model of fair rules* is not a fixed term. The exact meaning is context sensitive. People's preferences towards consequences, and a specific balance of consequences, differ depending on circumstances and context. This means that universally applicable procedural maxims aimed at facilitating procedural fairness are of little practical guidance unless applied to a specific context based on a rigorous method of balancing. In the remaining chapters, I will demonstrate how the model of fair rules can be applied in the context of EU antitrust procedure, to solve problems of balancing towards the fair procedural solution.

I began this chapter with a reference to a procedural aspect of the 11th century arson trial of Flosi Þórðarson. The result of the trial was that Flosi's council, Eyjólfur Bölverksson, succeeded in ruining the case by using procedural tricks that lead to a mistrial. Outraged by this outcome, the leading council of the opposing side, instantly upon hearing the result, grabbed a spear, and killed one of Flosi's supporters. A battle between the two camps broke out at the trial site leaving several men dead, before a settlement on monetary compensations was reached at the behest of the leading chieftains of the Icelandic Commonwealth.

The morality of abusing procedural rules is expressed in the story through the fate of the great lawyer Eyjólfur Bölverksson. He enters the story as one of the wisest lawyers of the Commonwealth, but is also known for being notoriously fond of money. His weakness leads him to accept a precious gold bracelet as a payment for representing Flosi. Accepting a payment for such job was already a bad omen; knowledge was at the time considered a

common good that should not be sold to the highest bidder. After having used his superior legal skills for achieving injustice, he was spotted at the trial site during the ensuing battle by Kári Sölmundarson, the heroic lone survivor of Flosi's arson attack:

Þorgeir said: 'Look, Eyjólfur Bölverksson is over there, if you intend to pay him for the bracelet'. Kári responded: 'Sure, I would not mind reimbursing him'. Then Kári grabbed a nearby spear and threw at Eyjólfur. It struck him in the torso and went through him. He instantly fell dead.³²⁷

The disgraceful trial defence devised by Eyjólfur Bölverksson, barred his son from receiving any compensation for his killing during the subsequent peace settlement.

Through an artful use of the narrative of the understated, an implicit understanding of the essence of a procedure as the facilitator of an accurate result appears in the actions of the protagonists. Thus, in line with the brute 11th century Commonwealth justice, Eyjólfur could be justifiably killed for facilitating injustice through his cunning trial tactics.

³²⁷ See *Brennu-Njáls saga* (~1270-1290) ch 145 <http://www.sagadb.org/brennu-njals_saga> accessed 3 January 2017. Loose translation to English from the original Icelandic text by HLK.

The norm and the process of EU competition law

A procedure stands in relation to a substantive objective of some sort. EU competition law procedures stand in relation to the substantive objective of EU's competition law regime. Procedures are auxiliary to a substantive objective and have as a primary purpose the objective's realisation. For the present purposes, the objective of EU's competition law is a fact that relates to the problem of procedural fairness. The solution to the problem of procedural fairness, affects how the substantive objective is realised. The stakeholders in a competition law procedure can disagree about what the objective of the competition law regime ought to be, and to what extent the objective ought to reflect their subjective interests over the interests of others. However, the disagreement we are currently interested in relates to the procedural aspect, i.e. *how* the objective ought to be realised, not *what* it ought to be.³²⁸ To proceed to the question of *how*, it is useful to establish first *what* the *how* relates to.

The process, or the act, of competition is the key concept in establishing the objective of competition law; both universally for all systems of law, and specifically with reference to specific systems of law. If there were, a single ultimate objective that all competition law regimes ought to be pursuing it would necessarily, in the strict sense, have to be built on the premise that there is something intrinsic about the process of competition that makes it worth pursuing independently of any other policy objectives. The process of competition can alternatively be viewed as a facility for pursuing ulterior policy objectives. If perceived in that way, the prospect of a universal objective of competition law is contingent on the universality of that specific policy.

Against the background of these fundamental considerations about the normativity of competition law, I will now proceed with recapping the historical development of competition law in Europe, which informs the proper understanding of the current consensus on the objective of EU competition law.

³²⁸ Reference is made to the discussion in chapter five on procedures, regarding the difference between a substantive provision and a procedural provision, and why it is sometimes useful to explicitly legislate a procedure, while at other times it might be superfluous.

1. Recap of the historical development

The history of European competition law is to an extent intertwined with the history of antitrust law in the United States. The enactment of the Sherman Act in 1890 in the US is usually taken as the starting point of modern competition law enforcement.³²⁹ Designed to combat the utilisation of trusts to cartelise and monopolise certain important industries, the two main pillars of the Sherman Act have their equivalent in the two main competition law provisions of the EU Treaties, which appeared already in the original 1957 Treaty of Rome.³³⁰ Section 1 of the Sherman Act prohibits the formation of cartels in a similar way as Article 101 TFEU, and section 2 of the Sherman act is hedged against market monopolisation and abuse of dominance in a similar way as Article 102 TFEU.³³¹

The fundamental provisions of the two competition law systems appear in a similar abstract format by identifying the problem of the cartel and the problem of the monopoly. However, the exact policy recommendations derived from the abstract wording of these provisions, have differed both across these two jurisdictions, and over time the policy recommendation within each system has changed and evolved. The nature and the structure of the political entities the competition law regimes have been pledged to serve can be named among the reasons for disparity in practical application.³³² The common market of the European Community of 1957 was not as homogeneous in composition as the market supervised by the US federal government in 1890. Since then both entities have evolved and changed in nature and structure, which again has prompted an evolution in the practical application of the anti-cartel and anti-monopoly provisions.

³²⁹ For an overview of the US competition law regime see for example; Maher M. Dabbah, *International and Comparative Competition Law* (Cambridge University Press 2010) 227-274.

³³⁰ The Paris Treaty of 1951 on the establishment on the European Coal and Steel Community also included provisions aimed against anti-competitive market behaviour, i.e. Articles 65 and 66. In the Rome Treaty of 1957 the main competition provisions appeared in Articles 85 and 86, which later became Articles 81 and 82. After the Lisbon Treaty of 2007, which entered force in 2009, these competition provisions have again been rebranded as Articles 101 and 102 TFEU.

³³¹ Maher M. Dabbah, *International and Comparative Competition Law* (Cambridge University Press 2010) 227-28 and 238-43.

³³² This should not be surprising given the previous analysis of the history of procedures in chapter five, where procedure's sensitivity towards the constitutional nature of the organ they serve was discussed.

The Sherman Act of 1890 can be viewed as a reaction against specific problems caused by trusts that sought to cartelise and monopolise important industries. The competition law provisions of the Treaty of Rome were of a different nature. The establishment of the common market was the key objective of the Treaty of Rome and the inclusion of the competition law provisions therein along with the free movement provisions created the *modus operandi* of that primary objective.³³³ The initial objective of the EU competition law regime can thus be seen as proactive towards the purpose of outlawing cartel and monopoly activity capable of threatening the success of the common market, while the initial objective of the Sherman Act was reactive against a particular way of doing business that had unpopular repercussions among the US electorate.³³⁴

By the mid-20th century, the US antitrust regime had developed an intellectual paradigm that built on quasi-economic thinking and the legal interpretive methodology of searching for the legislator's intent. This intellectual paradigm became known as the Harvard school of thought. The key intellectual feature of the Harvard antitrust doctrine was the hypothesis that an inference regarding the economic performance of the market could be drawn from how it was structured (i.e. number, size and market power of firms) and how the firms conducted themselves on the market (i.e. how they set prices).³³⁵ The claim was that efficiency (usually allocative, productive and dynamic efficiency) was at a maximum in a market where the firms set the prices in accordance with consumer demand. Such conduct was in turn considered more likely where the number of firms on the market, their size, and entry barriers prevented them from deciding prices unilaterally or in collusion with other firms. State intervention was

³³³ See David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 *Harvard International Law Journal* 97, 102.

³³⁴ Dabbah discusses different hypothesis for the initial inspiration of the Sherman Act and claims that competition law was one of the central issues of the 1912 presidential elections, eventually won by Woodrow Wilson. It is hardly a coincident that shortly thereafter important improvements were made to the US antitrust regime with the enactment of the Clayton Act of 1914 and the Federal Trade Commission Act of 1914. This may suggest that the origins of US competition law are modestly rooted in easily digestible anti-wealth concentration populist agenda, rather than elitist economic or moral agenda. The label 'antitrust' suggests the same. See further Maher M. Dabbah, *International and Comparative Competition Law* (Cambridge University Press 2010) 228-29. Gerber suggests that the Sherman Act was simply a 'shot in the dark' in respond to 'populist political pressures' to curb the power of certain big businesses. See David Gerber, *Global Competition: Law, Markets, and Globalization* (Oxford University Press 2010) 123

³³⁵ Monti refers to the intellectual paradigm as the 'Structure-Conduct-Performance' paradigm. See Giorgio Monti, *EC Competition Law* (Cambridge University Press 2007) 57-59. See also F. M. Scherer and D. R. Ross, *Industrial Market Structure and Economic Performance* (3rd edn, Houghton Mifflin 1990) 4.

thus considered justified to curb the concentration of market power and to protect a desirably diversified structure of the market. This narrative of economic reasoning fitted smoothly with the circumstances, which prompted the US legislator to enact the Sherman Act and later the Clayton Act. This fit made the paradigm more palatable to judicial and administrative decision-makers trained in legal methodology, familiar with the legislative intent interpretive doctrine.³³⁶

The initial intellectual underpinnings of the competition law provisions of the Rome Treaty reflected an alternative view on how the instrument of anti-cartel and anti-monopoly legislation could be utilised. At the instrumental level, the actual provisions sought inspiration in the US experience with antitrust legislation. For historical reasons, the US were at the time in a unique position to influence the organisation of the economy in the allied controlled part of Europe during the post-war restructuring phase of the early 1950s. This is reflected in the antitrust provisions of the Paris Treaty of 1951, which were drafted under strong influences of US antitrust thinking.³³⁷

The initial impact of US antitrust thinking through the provisions of the Treaty of Paris failed to materialise in the Treaty of Rome.³³⁸ An effort had been made by Jean Monnet, the chief architect of the Treaty of Paris, to conceal the US origins of the provisions, which invited the European negotiators to attach their own purpose to the provisions. This purpose fitted the current European problems and thus differed in important ways from the initial purpose of the Sherman Act and how competition law was understood in the US through the reigning intellectual paradigm of the Harvard school. The negotiators of the Treaty of Rome were negotiating the integration of their economies to yield a predicted economic efficiency gain.

³³⁶ Maher M. Dabbah, *International and Comparative Competition Law* (Cambridge University Press 2010) 251-53.

³³⁷ Jean Monnet allegedly tried to hide the American influences over the drafting of the competition law provisions of the Treaty of Paris. He enlisted his friend Robert Bowie, a professor of antitrust law at Harvard, to make the initial draft of the competition law provisions and conferred closely with the US occupation authorities during the process. The draft was subsequently redrafted to fit European legal vocabulary and traditions by the Frenchman Maurice Lagrange, which had the effect of masking the US pedigree of the basic ideas. See further David Gerber, *Law and Competition in the Twentieth-Century Europe: Protecting Prometheus* (Oxford University Press 2001) 336-42.

³³⁸ On this point see David Gerber, *Law and Competition in the Twentieth-Century Europe: Protecting Prometheus* (Oxford University Press 2001) 342.

The competition provisions were considered as ancillary to the integration purpose and where viewed through the same lens as the primary purpose.

The rise of German ordoliberal thinking provided the negotiators and the European politicians at the time with an intellectual paradigm on the basis of which the rationality of the integration program could be explained: efficiency would be achieved through economic integration based on principles of economic freedom, i.e. the economic constitution. The competition law provisions formed, along with the free movement provisions, the core maxims of the new economic constitution of the European Economic Community established by the Treaty of Rome.³³⁹ At the surface level, the actual instrument of competition law in its abstract form in the Treaty of Rome was not that different from the US model, but at the deeper theoretical level, the intellectual reasons supporting the models differed. These differences eventually yielded different enforcement policies and priorities during the initial period of European competition law enforcement; enforcement in Europe became geared towards facilitating a unified market, while US enforcement followed the Harvard school paradigm of intervening to protect a certain market structure.³⁴⁰

The significance of the competition law provisions of the Treaty of Rome was not fully appreciated by the original members of the EC. The assumption seems to have been that the provisions were mere policy objectives rather than the robust enforcement tool it later became. Due to the abstractness of the initial Treaty provisions on competition, the margin for operationalising them through secondary legislation was considerable. The Member States,

³³⁹ Parallel to the negotiation of the Rome Treaty a debate was taking place in Germany about the enactment of a national competition law regime. This debate was much influenced by the ordoliberal thinkers identified with the Freiburg School of economic and political thought. Eventually the German competition law regime entered force on the same day as the Treaty of Rome on 1 January 1958. See Andreas Weitbrecht, 'From Freiburg to Chicago and beyond – the first 50 years of European Competition law' (2008) 29(2) *European Competition Law Review* 81, 82.

³⁴⁰ See David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 *Harvard International Law Journal* 97, 102. Leucht and Marquis argue that US antitrust ideas continued to influence the principle actors of EEC competition policy during the 1960s and 1970s; key personnel in DG IV were to varying degrees exposed and influenced by connections to the US antitrust intelligentsia, while the ECJ was exposed to arguments from US antitrust thinking through the adjudicative process. From their case study, they do however conclude that despite exposure to US based antitrust ideas, the ECJ generally reached its conclusions based on its own intra EEC rationale. See Brigitte Leucht and Mel Marquis, 'American Influences on EEC Competition Law' in Kiran Klaus Patel and Heike Schweitzer, *The Historical Foundations of EU Competition Law* (Oxford University Press 2013) 125-61.

acting through the Council on the assumption that the competition provisions were unimportant, delegated to the Commission the task of designing the institutional model through which the competition law regime would be enforced. Not surprisingly, the Commission suggested that the enforcement powers should be centralised within the realm of the Commission in what later was adopted as Regulation 17.³⁴¹ Initially, it was debated whether the enforcement regime should be of an administrative character and subordinated to industrial policy, or alternatively whether it should be more independent and juridical in character modelled on the ordoliberal principles of the economic constitution. By the late 1960s the later approach favoured by Germany had gained the status of orthodoxy, at the cost of the former advocated by France.³⁴²

During the foundational period in the 1960s, the Commission and the Court of Justice collaborated on the advancement of the fundamentals of competition policy. The Commission gradually started to use the enforcement powers invested in DG IV (now DG Competition) through Regulation 17. The Court of Justice at the same time used its revision powers to confirm the Commission's approach to the enforcement and to reinforce further advancement, through sweeping general statements in its judgements that served the purpose of advancing active competition law enforcement doctrine.³⁴³ The stagnation of the integration process on the political front, following the French *empty chair crisis* (1965) and the *Luxemburg compromise* (1970), elevated the importance of this collaboration.³⁴⁴ On one hand the Commission had, lost momentum in introducing new legislative proposal for integration purposes due to lack of political support in the Council, but retained the prerogative of using

³⁴¹ Council Regulation No 17 (EEC): First Regulation Implementing Articles 85 and 86 of the Treaty [1962] OJ No. 013, 21.02.1962.

³⁴² See David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 Harvard International Law Journal 97, 103-07; and David Gerber, *Law and Competition in the Twentieth-Century Europe: Protecting Prometheus* (Oxford University Press 2001) 343.

³⁴³ See David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 Harvard International Law Journal 97, 110-11. See also David Gerber, *Law and Competition in the Twentieth-Century Europe: Protecting Prometheus* (Oxford University Press 2001) 351-52.

³⁴⁴ Weiler suggest that the process of constitutionalisation during the foundational period, spearheaded by the Court of Justice between 1963 and the early 1970s through the creation constitutional doctrines such as *direct effect* and *supremacy* of EU law, resulted in a political backlash. Subsequently, the Member States sought to regain control over the integration process that by then had mutated into an apparatus of hard law making that perhaps was not initially foreseen or intended by the national governments. The result was an equilibrium; the integration process continued, but now under tighter control by the Member States. See J.H.H. Weiler, 'The Transformation of Europe' (1991) 100 The Yale Law Journal 2403, 2426-29.

the tools already in place. On the other hand, the Court was able, in the absence of the political actors, to assert a position of intellectual leadership over the integration process through the bold use of teleological interpretation techniques.³⁴⁵

Following the economic stagnation of the 1970s and a subsequent move in the politics towards the protection of national interests *'under the shadow of the veto'*, as professor Weiler put it,³⁴⁶ the role of the Court as the primary engine of integration continued and remained so into the 1980s. The competition policy during that time, supported by a strong supranational enforcement regime established during the foundational period, was an important vehicle for the Court in advancing the integration cause at times when the other central agents in the system established by the Rome Treaty were entrenched on maintaining the status quo.³⁴⁷ The competition policy was one of few areas where the Commission could act proactively during the 1970s, largely due to the extensive powers granted to it through Regulation 17. Consequently, competition law enforcement became the primary venue for advancing the integration cause and the central laboratory for the Commission and the Court to test and extend the limits of their supranational powers. By the mid-1970s, the Court subtly began to request more rigorous reasoning from the Commission, especially with regards to providing economic data in support of reasoning based on maxims from economics.³⁴⁸

The institutional restraints during the foundational period shaped the focus of substantive enforcement. In a political environment where an industrial policy favouring national champions was supreme, and where European industry at large faced tough competition from US producers, it would have been risky for unestablished institutions to pursue aggressively

³⁴⁵ Gerber suggest that the Court of Justice could succeed in this leadership role sheltered from political scrutiny due to the traditional aura of objectivity and neutrality that senior courts enjoy. See David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 Harvard International Law Journal 97, 110-11. See also David Gerber, *Law and Competition in the Twentieth-Century Europe: Protecting Prometheus* (Oxford University Press 2001) 353.

³⁴⁶ *'[...] although the language of the Luxemburg Accord suggested its invocation only when asserting a vital national interest, its significance rested in the fact that practically all decision-making was conducted under the shadow of the veto and resulted in general consensus politics.'* See J.H.H. Weiler, 'The Transformation of Europe' (1991) 100 The Yale Law Journal 2403, 2460.

³⁴⁷ David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 Harvard International Law Journal 97, 115-16.

³⁴⁸ David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 Harvard International Law Journal 97, 118-19.

horizontal restraints or abuse of dominance that would have infiltrated vital national interest of the Member States.³⁴⁹ The initial enforcement focus was thus on vertical restraints; an policy of enforcement that could be rationalised with reference to the integration objective without penetrating key spheres of national industrial policy. The leading case in this regard was the *Consten and Grundig* case (1966)³⁵⁰ which concerned a restraint on cross-border trade through an exclusive distribution deal covering the territory of a whole country.³⁵¹

By the mid-seventies, the foundations had been laid and the Commission and the Court, by then established institutions, became more comfortable with extending the substantive scope of the enforcement policy in competition law. While the focus was still on vertical restraints, the Commission increasingly turned its attention to the provision on abuse of dominance (now Article 102 TFEU) and to a lesser degree on horizontal restraints. The Court could maintain the momentum of its efforts to further the integration project by establishing new principles within a new substantive field.³⁵² As before the interpretive method was teleological; the Treaty provisions should be interpreted towards the objective of market integration. This methodology was in contrast with the more traditional interpretive method of seeking the legislative intent in support of a literal analysis of the Treaty provisions. The *Continental Can Case* (1973)³⁵³ is an example of the extreme use of the teleological interpretive method. In that case, the Commission, with the support of the Court, applied Article 86 of the Treaty of Rome (now Article 102 TFEU) on a merger, despite a lack of explicit legal grounds in the Treaty text, and despite a recorded legislative intent of not granting the Commission with powers to regulate mergers. The integration imperative was found to be a superior objective to the adherence to traditional legal methods.³⁵⁴

³⁴⁹ David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 *Harvard International Law Journal* 97, 111-13.

³⁵⁰ Joined Cases 56 and 58/64 *Consten and Grundig v Commission* [1966] ECLI:EU:C:1966:41.

³⁵¹ Andreas Weitbrecht, 'From Freiburg to Chicago and beyond – the first 50 years of European Competition law' (2008) 29(2) *European Competition Law Review* 81, 83. See also David Gerber, *Law and Competition in the Twentieth-Century Europe: Protecting Prometheus* (Oxford University Press 2001) 354-56.

³⁵² David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 *Harvard International Law Journal* 97, 120-22.

³⁵³ Case 6/72 *Europemballage Corporation and Continental Can Company v Commission* [1973] ECLI:EU:C:1973:22.

³⁵⁴ David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 *Harvard International Law Journal* 97, 116-17.

The bold use of teleological interpretation facilitated a heterodox rationale for establishing legality. Instead of treating traditional legal sources, i.e. black letter law and legislative intent, as the ultimate reference points of legality in competition law, the ultimate reference point became the integration imperative. The integration imperative was hardly a traditional legal concept capable of forming a reference point in legal argumentation, but by assuming such position the traditional sources were relegated from the position of expressing ends that needed not to be regressed any further, into being means of furthering the end of market integration. This shift in rationale for establishing legality facilitated the use of arguments based in economics theory in competition law adjudication. The integration imperative relied on ideas in economic theory for its intellectual foundations, and thus an argument about legality in competition law needed to be compatible with such rationale.

The political climate regarding the integration project changed during the mid-1980s. The Single European Act of 1986, which signalled a political commitment to consolidate the internal market into the *Single Market* by 1992, broke the political stalemate of the consensus politics following the *Luxemburg compromise* of 1970. The decision procedure in the Council was relaxed with regards to internal market issues and most decisions were now to be taken based on a qualified majority. Decisions were thus increasingly taken '*under the shadow of the vote*' instead of '*under the shadow of the veto*'.³⁵⁵ Following renewed legislative efforts to complete the internal market, the importance of the Court with regards to the integration effort diminished. The politics had regained the initiative and the Commission was given the leading role in preparing and orchestrating the technical details of completing the political objective of the Single Market. The effects of these changed dynamics can be seen in the more cautious approach taken by the Court in its competition law decisions starting from the mid-1980s.³⁵⁶

³⁵⁵ Weiler argues that the most important provision of the Single European Act was the departure from a unanimous decisions procedure regarding internal market issues. The Member States, eager to establish the Single Market, were willing to facilitate its creation by relaxing the decision procedure in the Council. Due to the prerogative of the Commission to choose the legal basis of legislative proposals, this in effect meant that most decisions were taken under the relaxed procedure and the impact was a changed institutional culture that now operated under the possibility that a failure to reach consensus would result in a majority decision, instead of a status quo stalemate. See J.H.H. Weiler, 'The Transformation of Europe' (1991) 100 *The Yale Law Journal* 2403, 2456-61.

³⁵⁶ David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 *Harvard International Law Journal* 97, 126-30.

The establishment and the completion of the Single Market by the mid-1990s, to an extent exhausted the primary objective of EC's competition policy. The urgency of the integration imperative that had fuelled innovation in competition law diminished and the enforcement acquired a more traditional methodological flare; the key legal maxims were already in place for the main substantive fields of competition law, and the traditional juridical modes of legal argumentation could increasingly be used to infer reasons for legality and illegality.³⁵⁷

A key element in the system however remained; the basic principles of competition law, created through teleological reasoning based on the integration imperative, were still in place and retained their economics based flare. Due to their pedigree in economics based thought, a formalistic application of these principles had the potential of maturing an enforcement policy at odds with the economic rationale of the founding maxims. In response, the Court increasingly asked for more effects based analysis of factual evidences, which in turn increased the role of arguments based on economic evidences and a rationale compatible with efficiency theories from economics.³⁵⁸

Due to superior resources, the Commission was better equipped to lead the way towards what became the *more economic approach* to competition law from the mid-1990s. The economic approach at times requires the intensive use of specialist resources, which the Commission was gradually able to acquire. The role of the Court became more superficial. It lacked the resources to scrutinise seriously economic evidences and rationale, and thus focused instead on the less resources demanding issues of form and procedure. The Merger Regulation of 1989³⁵⁹ and the establishment of the Court of First Instance (now the General Court), also in 1989, further strengthened the Commission's institutional role in the Union's competition

³⁵⁷ David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 Harvard International Law Journal 97, 126-27.

³⁵⁸ As noted by Anne Witt, the volume of the Commission's decisions has increased significantly from the 1990s: '*This trend is not specific to merger review. The Commission's decision practice shows a similar development in cases investigated pursuant to Articles 101 and 102 TFEU. One side effect of this commitment is that it has considerably increased the length of the Commission's prohibition decisions. The Ryanair prohibition from 2007 exceeded 500 pages. By comparison, the Commission's very first merger prohibition in 1991, which also investigated a concentration in the air transport sector, filled less than 40 pages.*' See: Anne C. Witt, 'From Airtours to Ryanair: is the more economic approach to EU merger law really about more economics?' (2012) 49 Common Market Law Review 217, 232.

³⁵⁹ Council Regulation (EEC) 4064/89 on the control of concentrations between undertakings [1989] OJ L395.

policy and at the same time diminished the role of the Court of Justice. The merger policy was politically important and almost exclusively enforced by the Commission, while the establishment of the Court of First Instance meant a reduced jurisdiction for the Court of Justice and thus a lesser institutional role for the judiciary in its most authoritative capacity.³⁶⁰

In the mid-1990s, the Commission began internal work on redefining the role and objective of competition policy in the expanding Union, partly in response to developments in US antitrust law that was increasingly applied along the lines of the Chicago school of economics, inspired by the critiques of Posner and Bork on the Harvard paradigm in the 1970s.³⁶¹ Around the same time, the institutional processes established by Regulation 17 in 1962 were reaching capacity restraints by an ever-increasing volume of notifications about agreements with potential anti-competitive effects, from a growing number of Member States.³⁶² The procedural regime was eventually modernised through the enactment of Regulation 1/2003³⁶³, which replaced Regulation 17 and introduced several radical changes to the enforcement procedures. By the mid-2000s a subtler modernisation had also occurred with regards to the objective of competition policy that now regarded the term *consumer welfare* as its ultimate substantive premise.³⁶⁴

The procedural modernisation altered *how* EU competition law was enforced. The notification obligation, that was clogging the system, was abolished. National competition authorities were enlisted as enforcement agents, and the role of the Commission changed from being the sole

³⁶⁰ See David J. Gerber, 'The Transformation of the European Community Competition Law?' (1994) 35 *Harvard International Law Journal* 97, 126-35; David Gerber, *Law and Competition in the Twentieth-Century Europe: Protecting Prometheus* (Oxford University Press 2001) 371-85.

³⁶¹ The neo-liberal turn in US and UK politics during the 1980s was ignited and fuelled by several economics oriented scholars, most prominently represented at the University of Chicago during the 1960s and 1970s. Bork (see Robert H. Bork, *The Antitrust Paradox* (Free Press 1978)) and Posner (see Richard A. Posner, *Antitrust Law* (University of Chicago Press 1976)) adopted these ideas and criticised the orthodox approach to US antitrust law for being overly interventionist and naïve in terms of economic rationale. See David Gerber, *Law and Competition in the Twentieth-Century Europe: Protecting Prometheus* (Oxford University Press 2001) 384; David J. Gerber, 'Two Forms of Modernization in European Competition Law' (2007) 31 *Fordham International Law Journal* 1235, 1248-50.

³⁶² See David J. Gerber, 'Two Forms of Modernization in European Competition Law' (2007) 31 *Fordham International Law Journal* 1235, 1236-39.

³⁶³ Council Regulation (EC) 1/2003 on the implementation of the rules on competition laid down in Articles 81 and 82 of the Treaty [2002] OJ L001.

³⁶⁴ David J. Gerber, 'Two Forms of Modernization in European Competition Law' (2007) 31 *Fordham International Law Journal* 1235, 1247-48.

enforcer into being the central orchestrator of a Europe wide network of competition authorities. The net effect of the changes was that the procedure was simplified by skipping the notification duty, and the enforcement resources increased by requiring national authorities in each Member State to apply the competition provisions in cases of a Union dimension. The total capacity for enforcement thus increased. It is not entirely clear whether these changes should be viewed as a decentralisation of enforcement power for the benefit of the national authorities, or whether the enforcement power of the Commission was solidified by consolidation of the influence over how the national authorities apply competition law within their respective jurisdictions.³⁶⁵

The normative modernisation altered *what* was being enforced through EU competition law. The integration imperative supported by ordoliberal ideas about economic freedom of action, that had provided intellectual rigour to the enforcement until the 1990s, gave away by the early 2000s for what has been labelled a *more economic approach*. The term *more economic* in this context should be taken as a reference to the increased importance of efficiency considerations in deciding which behaviour should be deemed compatible with the competition law regime.³⁶⁶

The history of competition law in Europe and antitrust law in the US reveals two perceptions of anti-cartel and anti-monopoly legislation: the competitive process seen as a fundamental objective, or as subjected to normative elements considered hierarchically superior. The competitive process can thus be viewed either as a mean or as an end. In case it is viewed as a means, the character of antitrust and competition law becomes procedural; the objective of the legislation becomes the facilitation of a specific objective that exists independently of the process of competition.

³⁶⁵ David J. Gerber, 'Two Forms of Modernization in European Competition Law' (2007) 31 *Fordham International Law Journal* 1235, 1242-44. See also David J. Gerber, *Global Competition Law: Law, Markets and Globalization* (Oxford University Press 2010) 187-92.

³⁶⁶ Monti suggest that from a political perspective the core values of EU competition law have historically been three: market integration, economic freedom, and efficiency. The modernisation of the early 2000s should in that context be seen as an elevation of the efficiency value, not as an elimination of the other two. See Giorgio Monti, *EC Competition Law* (Cambridge University Press 2007) 20-22.

The initial objective of US antitrust law was simple enough; the process of competition should be protected to avoid the trust problem. By the mid-20th century this practical solution to a specific problem had evolved into a sophisticated legal doctrine supported by insights from economics referred to as the Harvard school of antitrust thought. The Chicago critique spearheaded by Bork and Posner in the 1970s was founded on a different perception of antitrust law; the protection of competition was not an end anymore; it had become a means towards achieving economic efficiency. This perception was constructed on the contentious premise that a specific understanding of economic efficiency ought to be the ultimate objective of public policy.

The objective of EU competition law was from the start perceived as something larger than the process of competition. The competition policy was auxiliary to a larger political ideal of an economically unified Europe. During the first decades of the Treaty of Rome, the Court of Justice actively invoked the grand political objective of the Treaty as the ultimate norm that ought to be pursued through the competition provisions. By the mid-1980s, this initial ideological thrust was becoming less relevant. The main strands of competition law doctrine had by then already been established, which lessened the need for an activist interpretation agenda. Gradually the focus thus turned to the process of competition as the practical objective of competition law, although the integration objective still existed at a more abstract and less pragmatically relevant level.

A focus on the process of competition in Europe was exposed to the economic efficiency based Chicago critique in similar ways as the Harvard antitrust doctrine had been earlier in the US. The *more economic approach* of the early 2000s was a European response to critique based on economic efficiency rationale; the focus again moved beyond the process of competition and onto a type of economic efficiency the process ought to facilitate. The more economic approach within the EU has however not yet become a pure economic approach. EU's competition policy still draws on normative elements from moral philosophy, in addition to the more recent neo liberal economic rationale. This is evident by the rhetoric emphasis on the consumer's interests in EU's current competition policy, which highlights an element of

normativity that is hard to reconcile with an entirely efficiency based approach as understood by neo liberal economics.³⁶⁷

EU's competition policy, from the start has consciously served a normative objective that is external to the actual process of competition. Yet there seems to be something about competition as such that makes it, in the instrumental sense, inherently suitable for achieving certain public policy objectives. Before turning to the objective of EU's competition law as stated in the positive law, I will devote few paragraphs on explaining the intrinsic value of competition as a process.

2. The process of competition

It is not entirely obvious that the process of competition should preferably be viewed as a mean towards achieving public policy, rather than as an end in of itself. While each of the normative ends that could be suggested as the ultimate objective of the process of competition is at a closer look contentious in some way upon inherently controversial political choices, the process of competition in the strict instrumental sense seems immune to controversy.³⁶⁸

The leading explanation, about what it is about the process of competition that makes it desirable, is based on Adam Smith's intuition about the invisible hand in the marketplace. In its modern representation, pioneered by Arrow and Debreu, the invisible hand is known as the two theorems of welfare economics. The first theorem predicts that in ideal circumstances a Pareto efficient equilibrium will emerge between buyers and sellers. The second theorem predicts that any Pareto equilibrium can be sustained artificially.³⁶⁹

³⁶⁷ See chapter one for examples of this rhetoric, such as Mario Monti's speech in Stockholm in 2000. See Mario Monti 'Fighting Cartels Why and How? Why should we be concerned with cartels and collusive behaviour?' (2000) 3rd Nordic Competition Policy Conference, Stockholm <http://europa.eu/rapid/press-release_SPEECH-00-295_en.htm> accessed 9 September 2016.

³⁶⁸ For a critical review of few of the mainstream normative objectives of EU competition law as presented in the literature see Ioannis Lianos, 'Some Reflections on the Question of the Goals of the Competition Law' (2013) UCL CLES Working Paper Series 3/2013, 2-32 <<https://www.ucl.ac.uk/cles/research-paper-series>> accessed 22 October 2015.

³⁶⁹ See Kenneth J. Arrow, 'An Extension of the Basic Theorems of Classical Welfare Economics' in Jerzy Neyman (ed) *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press 1951) 507-532; and Gerard Debreu, *Theory of Value - An Axiomatic Analysis of Economic Equilibrium* (Wiley 1959). For a decent nontechnical explanation of the theorems see Joseph E. Stiglitz, 'The Invisible Hand and Modern Welfare Economics' (1991) NBER Working Paper Series, Working Paper No 3641 <<http://www.nber.org/papers/w3641.pdf>> accessed 30 December 2016.

If the policy focus is primarily on the first theorem, i.e. to create the ideal circumstances for the formation of a Pareto equilibrium, the process of competition becomes the primary objective. A simultaneous focus on the second theorem justifies policy intervention for engineering the policy result desired, i.e. which Pareto equilibrium should be pursued. In that case, the competitive process becomes a mean towards achieving a specific equilibrium result, and that specific equilibrium becomes the pursued end. If the two theorems are perceived as complementary, the issue of distribution is dealt with successively. The first theorem strives for efficiency conditioned on a Pareto distribution, and through the second theorem, a specific Pareto distribution can be imposed.

An exclusive policy focus on the second theorem of welfare economics, allows for a separation of efficiency considerations dealt with under the first theorem and the distributional considerations. Separation of the efficiency issue from the distributional issue enables a toleration of market failures that result in efficiency equilibriums that do not generate Pareto distribution (i.e. Kaldor-Hicks efficiency equilibriums). In such cases, the focus of competition policy is the maximisation of economic efficiency irrespective of distributional issues. Distributional justice is perceived as being the issue of specialist redistribution mechanisms that are separate from competition policy and properly dealt with through the general taxation scheme.

Complexity of real world situations is the main obstacle to policy efforts to sustain a market environment that allows for the spontaneous formation of Pareto equilibriums in accordance with the first theorem. It is almost impossible to eliminate conditions of market errors that unduly enhance the strategic position of some agents and detriment the position of others. An intervention seeking to establish a vacuum, in which a Pareto optimal price can be established, has high probabilities of failing due to these difficulties. The problem of time is very hard to address through efforts to protect specific Pareto enhancing circumstances. As soon as the regulator has addressed a problem in accordance with a specific market dynamic, the chances are that the passing off time has made that response obsolete due to the time sensitivity of

the variables involved. The interaction of market actors on a competitive market is inevitably based on a fluid strategy that constantly evolves as new incentives emerge.

A minimalist approach to the first theorem of welfare economics relaxes the Pareto requirement in favour of a Kaldor-Hicks standard using a reference to the possibilities for the implementation of a distributional policy in accordance with the second theorem as a rationale. Under this approach a temporary Kaldor-Hicks equilibrium price, in which one agent can tax other agents, is justified based on the incentive for innovation that it creates. As soon as one agent can tax others, an incentive is created for others to try the same, and for the incumbent to innovate further to stay in the lead. The race for the ability to tax others thus enhances total welfare over time.

Assessed in terms of adaptability to time sensitive variables, the minimalistic approach to the first theorem in theory should be superior to the interventionists approach. Instead of attempting to freeze a specific stage of market innovation in time, efficiency is perceived as cyclical in which each cycle is initiated by an innovation that allows a taxation by the innovator. The cycle subsequently gravitates towards a Pareto equilibrium as competitors try to catch up, until a new cycle is initiated by an innovation.

The problems with the minimalist view are primarily of two types: the first type of problems relates to entrenchment on the status quo and the second relates to the time length of the cycle in which Pareto optimality is disrupted. From a public policy point of view, it would be desirable to maintain indefinitely a market structure that yields Pareto optimal results. In the same way, it would be desirable for an agent or a group of agents to maintain a situation indefinitely where they can tax other agents through a Kaldor-Hicks equilibrium. The minimalist view on public intervention risks misidentifying artificial entrenchment on the status quo (i.e. cartels and abuse of dominance) as a normal cyclical Kaldor-Hicks efficient taxation.

The problem of the cyclical length is less relevant where technical innovations are frequent. In such instances, the ability of one agent to tax others is bracketed by the short frequency of innovation and is an important instrument to incentivise investment in research and

development. In other cases, the frequency of innovation can be long, or even very long. In such cases, the incumbent could be able to tax others for a long, or a very long time, without resorting to any entrenchment tactics that would trigger intervention on the minimalist view. As the length of the innovation cycle increases (e.g. expensive infrastructure programs) and the market gradually takes the form of a permanent or a semi-permanent monopoly, the public policy argument for allowing some agents to tax others weakens, which in turn weakens the compatibility of the minimalist doctrine with common sense.

The minimalist view on competition policy, that pursues economic efficiency rather than the process of competition, may thus be superior in terms of factoring in dynamic efficiency considerations, but the dynamic responsiveness comes at the cost of practicality in cases of long innovation cycles which for practical purposes may seem permanent in the present, and at the cost of tolerance towards entrenchment tactics that delay or forestall natural gravitation towards Pareto optimality. These two views represent the extremes; the process of competition as purely instrumental, or as an ultimate end. Actual systems of competition law tend to be mixed in this sense, but do perhaps contain stronger elements of one over the other.

The intuitive appeal of competition as an instrument of public policy can be explained based on Adam Smith's *invisible hand* metaphor. Through the act of competing, something desirable happens that is hard to explain. Somehow, we feel that the result of a fair competitive process is equitable. Arrow and Debreu proved mathematically that a competitive equilibrium yields Pareto efficient results, which could be viewed as a rational explanation to the invisible object of desire in Adam Smith's metaphor. The results of a competitive process are intuitively appealing because they are Pareto efficient; or in other words, the *invisible hand* on the free market, is the gravitational pull towards Pareto optimality. Viewed in these terms, the instrument of competition is suitable for achieving Pareto efficiency and the use of the instrument in public policy would thus indicate an objective in which Pareto efficiency is either the primary target, or ancillary in achieving the primary target.

The ideal state of constant Pareto optimality, across time with regards to market prices, is unattainable for practical purposes. The practical issue thus becomes to find the best realisable

approximation. In that regard, it is not obvious whether treating the process of competition as an end or as a mean yields superior results. It is important to have this inconclusiveness in mind when discussing the current consensus on the objective of EU competition law. A contemporary emphasis on a specific standard of efficiency is not to be taken as a definite normative objective of competition law. Empirical data, advances in economic theory, or improved regulatory techniques at any moment might generate policy arguments in favour of a more process-focused approach to competition law.

3. The current consensus on the objective of EU's competition policy

The popular belief is that the current normative objective of EU competition law has something to do with being *more economic* than it used to be, and that the interests of *consumers* are of principal concern. This perception is derived from the modernisation initiative of the Commission in the 2000s, which through several legal acts elevated the interest of consumers to prominence and had the self-proclaimed objective of approaching the competition law provisions of the Treaties more economically.³⁷⁰ Notably, around the same time an interesting constitutional development occurred.

Following the Laeken declaration of 2001,³⁷¹ a new constitutional Treaty was drafted which proposed a change to how the competition law regime is constitutionally perceived. Among the primary objectives of the EU in the proposed constitution of 2004, was that the Union should offer its citizens '*an internal market where competition is free and undistorted*'.³⁷² This was in stark contrast with the previous layout in Article 3f of the Treaty establishing the European Economic Community (EEC)³⁷³ and Article 3g of the Treaty establishing the European

³⁷⁰ Monti explains that the turn towards the more economic approach was gradual and was initiated by the Commission's Green Paper on Vertical Restraints in 1996. See Giorgio Monti, *EC Competition Law* (Cambridge University Press 2007) 82-83. See also Ioannis Lianos, 'Some Reflections on the Question of the Goals of EU Competition Law' (2013) UCL CLES Working Paper Series 3/2013, 32 <<https://www.ucl.ac.uk/cles/research-paper-series>> accessed 22 October 2015.

³⁷¹ The Laeken Declaration on the Future of the European Union was issued as an annex to the press release following the European Council meeting in Laeken, Belgium, 14 and 15 December 2001. <http://europa.eu/rapid/press-release_DOC-01-18_en.htm> accessed 27 October 2015.

³⁷² See Article 1-3 (2) of the Treaty Establishing a Constitution for Europe [2004] OJ 310/47; '*The Union shall offer its citizens an area of freedom, security and justice without internal frontiers, and an internal market where competition is free and undistorted.*'

³⁷³ i.e. Article 3f of the Rome Treaty of 1957 (Treaty establishing the European Economic Community); '*For the purposes set out in the preceding Article, the activities of the Community shall include*' (...) '*the establishment of a system ensuring that competition shall not be distorted in the Common Market.*'

Community (EC),³⁷⁴ which listed the competition policy as one of the primary instruments for achieving the EEC's and the EC's aims listed in Article 2 of the respective Treaties. The proposed constitution infamously never became law. Few years later when the negotiation of the Lisbon Treaty had been completed, the elevation of the constitutional status of the competition policy to an objective of the EU was nowhere to be found. In fact, the competition policy had been demoted to a protocol status³⁷⁵ from its previous status among the primary instruments for achieving the objectives of the EEC and the EC.³⁷⁶ After the Lisbon Treaty, the constitutional status of the competition policy is obviously that of a mean rather than an end. Constitutionally, the promotion of the competition policy to an independent objective of the EU was explicitly rejected in favour of an instrumental view for achieving broader Union ends, in particular the internal market objective.³⁷⁷

The Commission's modernisation initiative that aims at an economic approach to competition law for the benefit of the consumers, and the constitutional status of competition policy within the Lisbon Treaty as an instrument for the facilitation of the internal market provide the outlines of the current consensus on the objectives of EU competition law. At least for the purposes of the positive law of the Union. This representation is however vague and thus requires further elaboration.

The Commission's focus through the modernisation initiative, on being more economic and to prioritise consumer interests, hints at normative elements. Protection of consumer interests

³⁷⁴ i.e. Article 3g of the Maastricht Treaty of 1992 (Treaty on European Union); *'For the purposes set out in Article 2, the activities of the Community shall include' (...)* *'a system ensuring that competition in the internal market is not distorted.'*

³⁷⁵ i.e. Protocol 27 to the Treaty of Lisbon (Treaty on European Union and the Treaty on the Functioning of the European Union); *'THE HIGH CONTRACTING PARTIES, CONSIDERING that the internal market as set out in Article 3 of the Treaty on European Union includes a system ensuring that competition is not distorted, HAVE AGREED that: To this end, the Union shall, if necessary, take action under the provisions of the Treaties, including under Article 352 of the Treaty on the Functioning of the European Union. This protocol shall be annexed to the Treaty on European Union and to the Treaty on the Functioning of the European Union.'*

³⁷⁶ Lianos attributes the relegation of the competition policy to the protocol section to the lobbying of Nicolas Sarkozy, the French president at the time. See Ioannis Lianos, 'Some Reflections on the Question of the Goals of EU Competition Law' (2013) UCL CLES Working Paper Series 3/2013, 37-40 <<https://www.ucl.ac.uk/cles/research-paper-series>> accessed 22 October 2015 Although the aim of Protocol 27 was to neutralise the effects of removing the competition policy from the list of the Union's primary instruments, the instrumental prominence of the competition policy was clearly reduced by the Lisbon Treaty.

³⁷⁷ Protocol 27 clearly bracketed the competition policy within the broader category of the internal market.

on the internal market could be explained based on economic theory; in particular, welfare economics. Initially, this seems a plausible approach to giving consumer interest a specific meaning, especially when being more economic in the methodological sense forms a parallel objective. Perceived in that way, the focus on consumer interest could be taken to represent the objective of a specific welfare standard within welfare economics, namely consumer welfare.³⁷⁸

From the perspective of economic theory in the strict sense, a focus on consumer welfare in public policy is however not the optimal objective.³⁷⁹ The theoretical problem with a rigid consumer welfare standard is its inability to respond to the temporal dimension of the market. If social welfare within the sphere of competition policy is composed of the aggregated welfare of producers and the relevant consumers in a specific market, an efficiency enhancement conditioned on consumer welfare is only allowed if it is neutral with regards to, or enhances, the welfare of consumers.³⁸⁰ This in effect means that producers are only allowed to implement Pareto improvements with regards to the consumers. The inability of producers to implement Kaldor-Hicks improvements in the short term with regards to consumers, disincentives innovation, which in the long-term harms aggregated total social welfare. Based on this type of rationale, it could be claimed that the optimal normative standard of competition policy should be total social welfare, rather than consumer welfare.³⁸¹ The

³⁷⁸ Ezrachi claims that the concept of consumer welfare in competition policy may be a universal benchmark of economic analysis in competition policy, but a universal agreement of the benchmark's properties is lacking. An agreement on the importance of the term thus mask a disagreement about its underlying rationale. See Ariel Ezrachi, 'Sponge' (2015) Working Paper CCLP (L) 42. 18-19 <<https://www.law.ox.ac.uk/sites/files/oxlaw/cclpl42-.pdf>> accessed 30 December 2016.

³⁷⁹ Nazzini explains comprehensively the basics of the welfare economics approach to competition law and concludes that a consumer welfare standard is theoretically inferior to a long term total welfare standard. See Renato Nazzini, *The Foundations of European Union Competition Law: The Objective and Principles of Article 102* (Oxford University Press 2011) 32-50. See also Giorgio Monti, *EC Competition Law* (Cambridge University Press 2007) 83-86.

³⁸⁰ Efficiency is measured on different scales. Allocative efficiency and productive efficiency measure statically how efficiently resources are allocated and goods produced at a specific point in time, while dynamic efficiency incorporates allocative and productive efficiency on a temporal scale that also measures the potential for innovation. These different methods for measuring the level of efficiency are methodological tools, while the standards of Pareto and Kaldor-Hicks are normative in character. See further Giorgio Monti, *EC Competition Law* (Cambridge University Press 2007) 45.

³⁸¹ Nazzini claims that long term social welfare with regards to competition policy is 'theoretically superior to any other objective', and in his mind this in particular applies with regards to the objective of consumer welfare. See Renato Nazzini, *The Foundations of European Union Competition Law: The Objective and Principles of Article 102* (Oxford University Press 2011) 50.

persuasiveness of the total welfare argument in competition policy depends the plausibility of severing distributional issues from efficiency issues in accordance with the second theorem of welfare economics. So long as only economic efficiency matters in competition policy, the normative objective of total welfare reigns.

As a matter of theory, it may be desirable to separate efficiency issues from distributional issues in competition policy. As a matter of a legal fact, the current competition law regime of the EU is not founded on that presumption. On the contrary, Article 101 TFEU, the central competition provisions of the Treaty, suggest that distributional issues are internal to EU competition policy.³⁸² This explicit legal fact lessens the compatibility of EU competition policy with a purely economic approach, especially those minimalistic ones that focus entirely on economic welfare. The Commission's focus on consumer interests on the internal market will thus not be properly explained based on a pure welfare economic theory. Such an explanation would additionally need to assume that the Commission's approach is illogical; i.e. that the Commission is somehow mistaken in preferring the inferior consumer welfare standard to the normatively superior total welfare standard.

A proper understanding of the normative objective of EU's competition policy thus requires the inclusion of the distributive issue, along with the efficiency issue. Economic theory is inapt at dealing properly with questions of distribution. Utilitarian calculus is insensitive to moral sentiments that are liable to influence the proper understanding of equitable distribution policy. Distribution is dealt with at a more advanced level within the domains of political and moral theory. The standards of Pareto and Kaldor-Hicks do address the distribution question to an extent, but the answers are not in line with most people's intuition about distributional fairness. The distributional issue is normally the primary task of the political, especially under democratic systems of government. The focus on the consumer in competition policy, as far as it is a distributional issue, is thus best understood as being from the perspective of the political in the sense of moral and political theory, rather than in the sense of economic theory.

³⁸² Monti points out that Article 81(3) (now 101(3)) explicitly contains a distributional clause which fits poorly with neoclassical economics theory's understanding of competition policy. See Giorgio Monti, *EC Competition Law* (Cambridge University Press 2007) 25-26.

By viewing the objective of consumer welfare and the focus on consumer interest in EU competition policy as a political distribution value, instead of as an efficiency standard, the objective starts to make more sense. It is of course preferable from a political point of view to strive for a competition policy that enhances total welfare in the economic sense, but regardless of utilitarian welfare calculus, it may nonetheless be even more preferable from a political perspective to accept a suboptimal welfare standard that is superior in terms of distributional fairness. The more economic approach pursuing the normative objective of consumer welfare and consumer interest on the internal market, should thus be seen as a compromise between a purely welfare enhancing objective in the sense of economic theory, and the objective of distributional fairness in the sense of moral and political philosophy.³⁸³ The equilibrium point between these two normative objectives, as settled by the political process that drafted the treaty provisions on competition and by the political process that decides the enforcement strategies of the Commission, is in fact a political compromise between the competing preferences of the stakeholders in EU's competition policy. While it may be rational for some stakeholders to prefer the objective of total welfare, at the same time it may be equally rational for other stakeholders to oppose the objective of total welfare in the absence of a rule of distribution that guarantees their interests. The political process, if sufficiently robust and efficient, resolves these conflicting claims equitably so that each stakeholder gets what he deserves, based on what he and other stakeholders want.

If we try to extract the current consensus on the normative objective of EU's competition policy from the relevant Treaty provisions and the Commission's stated enforcement policy in Article 101 and 102 cases, and merger cases, the normative elements of economic efficiency and consumer interests appear. If the primary stakeholders in the competition policy hold conflicting preferences with regards to the realisation of these normative objectives, neither can be fully realised without compromising the other. The political process has negotiated an

³⁸³ Ezrachi and Monti point out that the design of the institutions responsible for the enforcement of EU's competition policy is a factor in understanding the system's function. For example, the body of 28 politically appointed Commissioner's is the ultimate decision body in individual cases of competition law enforcement. This composition of the decision body invites political rationale into the decision making on top of the economics rationale that the case handlers at the lower administrative level may be more inclined to follow. See Ariel Ezrachi, 'Sponge' (2015) Working Paper CCLP (L) 42, 8 <<https://www.law.ox.ac.uk/sites/files/oxlaw/cclpl42.pdf>> accessed 30 December 2016.; Giorgio Monti, *EC Competition Law* (Cambridge University Press 2007) 2-6.

equilibrium between these two competing elements, which in turn is the sole normative objective of EU's competition law. The current equilibrium emphasises the maximisation of social welfare, subject to a condition of distributional fairness with regards to consumers. The distributional condition is not strict; it simply emphasises that efficiency enhancements should be Pareto with regards to consumers. Such a condition still leaves a room for Kaldor-Hicks improvements with regards to competitors. The welfare of individual firms can thus be increased to the detriment of the welfare of other firms, but not to the detriment of consumers.

This chapter provided a brief overview over *what* EU's competition procedure seeks to facilitate. The answer to the question of *how* it intends to do so depends on the applicable procedural rules. I have previously explained that in general, procedures need to strike a balance between achieving its task accurately and efficiently. The object or the task of accuracy and efficiency for the EU competition procedure is the political equilibrium I have described in this chapter. I have also explained that law enforcement procedures in general are instrumentally organised based on intervals of information framing and decision points. Next, I will analyse the Commission's competition law procedure against the backdrop of this framework and this ultimate optimisation objective. By doing so, I will attempt to identify the relevant stakeholders, the decision points in the procedure, and the framing intervals preceding them.

Stakeholders, decision points, and EU's competition proceedings

Assuming the success of the political process in producing a fair equilibrium between the main stakeholders with regards to the substantive goals of EU competition law, the normative objective of seeking economic welfare and protecting consumer interest is a common substantive objective of all the stakeholders involved.³⁸⁴ The procedural regime seeks in turn to implement this general objective through its various provisions and practices, both holistically and on a case by case basis. Due to the unattainability of doing so with absolute accuracy, discrepancies are liable to form with regards to how accurately the mutually agreed objective should be pursued and at which cost. As before, differently situated individuals with regards to the procedure are liable to hold diverging preferences with regards this balance of accuracy and the efficient use of enforcement resources. Essentially, this is the question of how many errors should be conceded through the procedural regime, and of which type. To establish the optimal balance in this regard I am suggesting using the *model of fair rules*, the function, and details of which I explained in an earlier chapter.

The key assumption that must be assumed on behalf of the procedural architect is that the stakeholder's preferences towards the accuracy/error question is dictated by a rational inference from the stakeholder's interest and practical circumstances. Each stakeholder that shares interests and practical circumstances with another stakeholder should thus through practical reasoning hold the same, or very similar, preferences towards the outcome of a specific regulatory decision. In the same vein, it would be perfectly rational for a differently situated stakeholder to hold different preferences towards the outcome of a specific regulatory decision. Based on this assumption, the key stakeholders in a procedure can be defined and categorised based on their practical circumstances and the preferences they are likely to hold in respond to their practical position concerning the consequences of the regulatory decision at hand.

Preferences towards specific consequences can be empirically measured or observed in number of ways. They could for example be estimated based on past choices (i.e. revealed

³⁸⁴ This was discussed and established in chapter 6.

preferences), and they could be estimated based on a survey or a questionnaire (i.e. stated preferences). A diligent procedural architect would be well advised to spend some resources on establishing through an empirical research an estimation of the actual preferences of the population in question towards the accuracy/error question. For the purposes of this research, I will not attempt to compile a dataset on the preferences of different stakeholders in an EU competition law procedure. The research resources at my disposal do simply not allow for such a study for the time being. I can however discuss briefly how the preferences are likely to trend in cases of rational actors. Many of the key variables are obvious to any observer and their impact on a rational choice can be approximated. Along these lines, I will now attempt to map the key stakeholders in EU competition law procedure, and discuss the probable composition of their preferences towards the question of procedural accuracy/errors. This mapping serves as a prelude to the actual application of the *model of fair rules* in chapter eight, where this chapter's analysis of EU's competition procedure will be operationalised with the aim of answering how fairness should be approached within the context of EU's competition procedure.

1. The stakeholders in a competition proceeding

For the purposes of the *model of fair rules* methodology, we need to simplify the segmentation of the population that holds a stake in EU's competition policy. The total population of the EU (and the EEA) forms the whole, of which all stakeholders are a part. As discussed in chapter six, the population has agreed to use the instrument of competition in the marketplace to pursue a competition policy that emphasises economic welfare, subject to a Pareto distribution condition with regards to consumers. The population also needs to agree on how to best achieve this objective of competition policy through the procedural regime. The choices made in the procedural regime affect the accuracy with which the objective is achieved and the types of errors conceded for those purposes. Segments of the whole are liable to hold different preferences as to the choices made in the procedural regime, with regards to the issue of accuracy and error concession. The role of the state, as a procedural architect, is to make regulatory choices that satisfy a condition of fairness with regards to the preference

satisfaction of the different segments, but at the same time strive for optimal efficiency of preference satisfaction within the boundaries of the fair.³⁸⁵

The *model of fair rules* uses the simplifying assumption that if the preference profile of the population is not unified, the whole should be split into segments based on their interest and circumstances and that each member of the population must belong to one of the segments, and one segment only. In the case of preferences towards a procedural regime in competition law, such a simplification is not entirely consistent with reality; an individual can and usually does belong to more than one segment at a time. A normal consumer for example can own stocks in a firm that becomes the subject of a competition law procedure. In such cases, the same individual holds stake in the procedure both as a consumer and as an equity owner in a firm. By using the simplifying assumption, an attempt is made to identify the primary concern that is likely to affect the individual's preference composition. It is to an extent optional how precisely the whole should be segmented. In some cases, it is sufficient to split it into two subgroups, while other instances would require a further segmentation to extract the essence of the conflicting preferences at stake.

For the purposes of the current procedural regime in EU competition law, a distinction is primarily made between interested parties and non-interested parties. The group of interested parties can further be divided into undertakings under investigation for having breached the competition provisions, and legal and natural persons that can assume the position of a complainant. The category of complainants can further be divided into undertakings that are competitors of the accused undertakings, and natural or legal persons that have the position of a consumer with regards to the accused undertakings or the relevant market at stake. The category of non-interested parties still holds stakes in the procedure, but only in their capacity as general taxpaying members of the public that pay for the enforcement system and suffer for the more general macro-economic consequences of competition law breaches. Non-interested taxpayers can both be undertakings and natural persons.

³⁸⁵ For a discussion on the definition of fairness in the laws and the role of the state in facilitating this definition see chapter two and chapter four.

For our purposes, we can extract from the above that the primary stakeholders in a competition law procedure are the firms liable to become the subject of such procedures. Their competitors, whose competitiveness relies on the lawfulness of the market behaviour of other firms, are also important stakeholders. Individual firms or members of the public, in their capacity as consumers form a third group of relevant stakeholders, and in their capacity as taxpayers form a fourth group of relevant stakeholders.³⁸⁶ We can refer to the first type as *the accused*, the second type as *the competitor*, the third as *the consumer*, and the fourth as *the taxpayer*.³⁸⁷

The role of the state and the state's enforcement institutions in the design of a procedure should not be viewed as a stakeholder role.³⁸⁸ The role of the state in the design should not be confused with the role of the state during the actual procedural process, where the state's enforcement institutions may have the role of acting on behalf of specific interests or stakeholders. The general consumer and the non-interested taxpayers, for example, have a limited access to the enforcement process of specific competition law cases, but implicit in the prosecutor role of the enforcement agencies is the pursuance of the consumer's interests and more generally the interests of the general public. In general, however, the state's decisional and framing role during a procedure, is to be considered above the interests of specific stakeholders. It is mostly about ensuring an equitable balance between competing interests during the procedural course of specific cases, which when acting in the prosecutorial capacity,

³⁸⁶ The definition of a consumer in the Commission notice on the application of Article 101 TFEU could be used for the former group; *'The concept of 'consumers' encompasses all direct or indirect users of the products covered by the agreement, including producers that use the products as an input, wholesalers, retailers and final consumers, i.e. natural persons who are acting for purposes which can be regarded as outside their trade or profession. In other words, consumers within the meaning of Article 81(3) are the customers of the parties to the agreement and subsequent purchasers'*; See Guidelines on the application of Article 81(3) of the Treaty [2004] OJ C101, para 84. The latter group could then be defined as any firm or individual that is not a member of any of the other groups of stakeholders, but who still is concerned with how public resources are spent on the enforcement of competition law and has a general macro-economic interest in minimising anti-competitive practices.

³⁸⁷ In praxis, individual stakeholders often have mixed stakes. They can for example be competitors and consumers at the same time. For simplifying purposes, an assumption needs to be made that a stakeholder belongs to one group and one group only. See further on the categorisation of stakeholders and the assumption of completeness of preference function in chapters three and four. Note also that in cartel proceedings, competitors are absent as a stakeholder group if all the relevant undertakings are part of the cartel.

³⁸⁸ This is consistent with the social contract view on society; the state as an institution acts on behalf of the people, not on behalf of itself. Enforcement institutions do thus have an auxiliary role and must act on a mandate from the stakeholders; an autonomous accord of institutions would be inconsistent with the democratic order.

may at times require acting on behalf of specific relevant but procedurally unprivileged interests.

2. The stakes in a competition proceeding

The dominant legitimate consideration for the accused during a competition proceeding, is to avoid becoming the subject of a false positive error. It could also be said, of course, that the dominant consideration is often not to be found guilty at all, regardless of true guilt. For our purposes, the assumption must however be made that all stakeholders are genuine in their commitment and support for the ultimate objective of competition law, and thus the accused is also in support of reprimanding true offenders against that objective, even if it turns out to be himself.³⁸⁹ The accused will thus ultimately prefer not to be found guilty unless he is. This will lead him to prefer high public investment in accuracy of the proceeding, which considering his proportion of potentially substantial error costs and his proportion of the public costs of the enforcement procedure, will remain marginally cost beneficial for him longer than for other types of stakeholders. This will also lead him to prefer distribution of errors that gives premium to the avoidance of false positive errors, over false negative errors.

The competitors of companies entangled in competition law proceedings will have an interest in seeing offenders against the objectives of the competition law regime brought to justice. The main motive for holding such interest would be their own position on the relevant market; if some actors are cheating on the competition rules, it is liable to deteriorate the competitiveness of those who stay within the boundaries of the permissible. In their capacity as operators on the market in question, an undetected breach against the rules can have serious consequences for individual competitors, while their share of the enforcement cost is not proportionated to their risk of becoming subjected to an error cost. The threshold for additional investment in accuracy to remain cost beneficial will thus be high, although not as high as the threshold for the accused. The competitor will have a dominating interest in avoiding false negative errors, while being at least neutral towards false positive errors. The competitor wants to avoid seeing offenders escaping justice through an enforcement error,

³⁸⁹ A preference for escaping justice cannot be considered as a legitimate preference in terms of designing public enforcement policy, a preference for not becoming the subject of injustice is however always a legitimate public policy consideration.

since that would deteriorate his competitiveness. A false negative error on the contrary would not inflict any cost on the competitor. The competitor would thus be neutral in his preferences towards such errors.³⁹⁰

Consumers have the most to gain if the objectives of the competition law regime are realised in an accurate manner. Accurate compliance implies competitive prices and frequent innovations that maintain a high level of social welfare in which the consumers receive an equitable share. The primary interest of the consumers concerns the price and the quality of the output of the relevant market. In terms of the procedural regime, the consumer should be willing to spend on additional accuracy, so long as his share of the enforcement expenditure is cost beneficial when compared with the consumer's gain in quality or price of the market output. The consumer is thus willing to pay for more accuracy, if he will gain at least as much in terms of price or quality of the relevant goods or services. The consumer is not conditioned on preferring a specific type of error allocation; he will simply prefer whichever allocation that is the most efficient in increasing consumer welfare.

The general taxpayer that does not qualify as an interested party in a competition proceeding does nonetheless carry a share of the enforcement costs and of the error costs that result from failures to accurately implement the objectives of the competition law regime. The individual taxpayer is detached from individual proceedings and is thus primarily concerned with larger macro-economic issues such as the efficient use of public resources and how the accurate enforcement of competition law influences social welfare in general. From the perspective of the taxpayer, it will remain rational to invest in procedural accuracy of competition proceedings as long as it remains cost beneficial in terms of total welfare, and as long as there are no other competing investment options that are likely to yield higher increase in total social welfare. This implies that the category of taxpayers would be quicker to reach the threshold where an additional investment in accuracy would not be cost beneficial. The taxpayer is not preconditioned on a preference for the type of errors to concede and will thus prefer whichever allocation that yields the greatest benefit in terms of social welfare.

³⁹⁰ It could be argued that the competitor would prefer that the accused would be found guilty irrespective of true guilt. Like before, such a bad faith preference cannot be accepted as a benchmark for designing public policy.

We can summarize the preference trends of the main stakeholders in a competition law procedure as follows:

Preference trends of stakeholders in a competition law procedure

	Total accuracy vs. level of investment	Distribution of errors
The Accused	Prefers high accuracy almost at any cost	Prefers type II errors
The Competitor	Prefers high accuracy and is insensitive to cost	Prefers type I errors
The Consumer	Cost beneficial accuracy in a narrow consumer sense	Neutral
The Taxpayer	Cost beneficial accuracy in a macro-economic sense	Neutral

By examining the provisions of Regulation EC 1/2003 and the Implementation Regulation EC 773/2004, we can see that firms subject to competition law proceedings are recognised as having entitlement to several procedural guarantees. Translated into the language of preferences towards accuracy and distribution of accuracy errors, these provisions on one hand ensure a certain level of investment on behalf of the EU in the accuracy of competition law proceedings, and on the other hand give guidance on the distribution of errors that favour the avoidance of false positive errors. Additionally, complainants (usually competitors, but sometimes consumers)³⁹¹ have standing that enables them to provide an input into the proceedings that facilitates accuracy selectively in accordance with their stakes.³⁹² The category of general consumers that do not file complaints have lesser access to individual proceedings, but the current normative objective of EU's competition law regime gives premium to their interest, which should be reflected in individual enforcement actions. The

³⁹¹ Consumers can in principle show legitimate interest in a case and become complainants in accordance with Article 7(2) of Regulation EC 1/2003, e.g. Joined Cases T-213/01 and 214/01 *Österreichische Postsparkasse v Commission* [2002] ECLI:EU:T:2006:151. In praxis, individual retail consumers do however rarely lodge effective complaints.

³⁹² The probabilities are that the competitors would mostly provide incriminating information, but would refrain from providing information that might undermine a prohibition decision.

stakeholder status of non-interested taxpayers is not as such recognised through the procedural rules. The public authorities governing budgetary decisions will however presumably be strongly influenced in their budgetary decisions by this category of stakeholders that, due to its numerical size, wields a considerable political influence.

Based on the analysis in chapter five on procedures, the instrumental function of procedures is an interval of framing phases and decision phases that seek to facilitate the substantive objective of a procedure. In doing so, the procedural instrument strives for an accurate output within the limits of the available material means. The optimal balance between the accurate and the material is an equilibrium between the preferences of the principle stakeholders involved, i.e. how much should be spent on accuracy and what kind of accuracy should be pursued. In the context of EU's competition law regime, the enforcement procedure is indeed instrumentally organised based on intervals of framing phases and decision phases in which accuracy and cost efficiency of the substantive result is the fundamental concern.

3. The procedural provisions in EU competition law

The enforcement of Articles 101 and 102 TFEU in typical infringement cases is regulated by Council Regulation EC 1/2003,³⁹³ that grants the Commission with enforcement powers in the field of competition law. Based on Article 33 of Regulation EC 1/2003, the Commission has further regulated its own enforcement procedures through Commission Regulation EC 773/2004.³⁹⁴ The competition procedure as laid down in the TFEU and Regulation EC 1/2003 evolves around the question of whether a specific behaviour constitutes a prohibited conduct with regards to Articles 101 and 102 TFEU. Article 7 of Regulation EC 1/2003 empowers the Commission to prohibit conduct that infringes the Treaty provisions. The Commission can also take commitment decisions, where the infringing firms offer behavioural or structural commitments that address the Commission's preliminary concerns.³⁹⁵ Additionally the Commission may decide that Articles 101 and 102 TFEU are inapplicable due to '*Community public interests*' and the Commission can by a decision order interim measures to bring urgent

³⁹³ Council Regulation (EC) No 1/2003 on the implementation of the rules on competition laid down in Articles 81 and 82 of the Treaty [2003] OJ L1/1.

³⁹⁴ Commission Regulation (EC) No 773/2004 relation to the conduct of proceedings by the Commission pursuant to Articles 81 and 82 of the EC Treaty [2004] OJ L123/18.

³⁹⁵ Article 9 of Regulation EC 1/2003.

competition law infringements to an end.³⁹⁶ For our purposes, I will limit the discussion to the procedure that leads up to the taking of a prohibition decision in accordance with Article 7 of Regulation EC 1/2003 and to the possibility of having such decisions appealed to the Court of Justice of the EU. This procedural avenue provides the most complete procedure for achieving accuracy in dealing with breaches to the competition law regime, while the others either cut the main procedure short or are exceptional from the norm and do thus not matter for the narrative of examining the accuracy and the efficiency of the normal competition procedure.

3.1. Decision points in EU's competition law procedure

Before an infringement to the competition law provisions of the TFEU has been indefinitely established, a series of intermediary decisions have been made that each is preceded by a sub procedure aimed at facilitating that specific decision. Each of the intermediary decisions form a part of a holistic procedure that cumulates in the final decision on the existence of an infringement or not. The significance of the intermediary decision points is that at each point leading up to an Article 7 prohibition decision, an intermediary decision can be made to end the procedure without pursuing any further enforcement actions. After the case reaches a court stage, the nature of the decision points alters a bit; the stake is still about ending further enforcement actions on behalf of the Commission, but the power to do so no longer rests within the Commission. The Commission has already reached its conclusion on the prohibition of a specific behaviour, but the subjects to that decision can attempt to have the prohibition annulled by seeking an overriding judgment from the CJEU.

An enforcement action in 101 and 102 prohibition cases starts with a complaint that notifies the Commission of an alleged breach, or through the Commission's own initiative, which may be based on market information received from interested or non-interested parties. After a potential breach has reached the attention of the Commission, an initial screening takes place that aims at facilitating a decision on whether to open a formal investigation. During this initial phase, the Commission may resort to inspections of company premises that aim at securing inculpatory evidences that are at risk of being spoiled by the subjects of the investigation. Less intrusive measures can also be used, such as information requests that require the firms

³⁹⁶ Articles 8 and 10 of Regulation EC 1/2003.

involved to provide the Commission with the requested information. The Commission can also interview people for the purposes of collecting information. Depending on the strength of the initial lead, a decision must be made within the Commission as to how thorough investigation is needed during the initial informal stage to reach a decision on whether to open a formal investigation. In some cases, the allegation made in a complaint might obviously be unfounded, or concern an infringement that is not sufficiently important for the Commission to take interest in pursuing an enforcement action. In such cases the Commission will simply close the case without entering an expensive investigation, or in some cases allocate it to a national competition authority (NCA). If, however the Commission resorts to inspections or information requests during the initial phase, it must be presumed that the Commission already has doubts that it intends follow up on through the opening of a formal investigation, unless the results of the inspections or the information requests rebuts the doubts the Commission already has about the alleged infringement.³⁹⁷

The opening of proceedings signals a commitment on behalf of the Commission to invest its resources in a particular case with the view of determining whether an infringement has been made to the competition provisions of the TFEU.³⁹⁸ Such a commitment is made on the implicit presumption that an infringement is probable and that the Commission has sufficient interest in the potential infringement to focus its resources there instead of somewhere else.³⁹⁹ The discretion of the Commission to open a formal investigation is not curbed by any official infringement probability standard. The Commission does however carry the burden of proof for infringement, and it must thus from the Commission's perspective, be probable from the start that it will be able to discharge that burden. The decision to focus on a specific case over others can be motivated by many practical reasons. It might be a case of a low hanging fruit that is easy to pick and conclude with a cost-efficient enforcement decision and thus show enforcement results. The incentive might also be a policy focus aimed at a specific market

³⁹⁷ See further on the initial assessment; Commission notice on best practices for the conduct of proceedings concerning Articles 101 and 102 TFEU [2011] OJ C308/03, paras 12-16.

³⁹⁸ See Commission notice on best practices for the conduct of proceedings concerning Articles 101 and 102 TFEU [2011] OJ C308/03, para 18.

³⁹⁹ The Commission of course cannot be explicit about what it believes with regards to a probable infringement and thus uses neutral language in its communication until a final decision on infringement is reached. See Commission notice on best practices for the conduct of proceedings concerning Articles 101 and 102 TFEU [2011] OJ C308/03, para 22.

behaviour on which the Commission wants to signal its standing. Sometimes the decision might even be motivated by a desire by the Commission to advance a new enforcement doctrine in which the circumstances of a specific case might form a good pretext.⁴⁰⁰ Irrespective of specific practical reasons for advancing a case, the overall objective of accurate and efficient enforcement always forms the superior reason towards which the case specific reasons function.

During the formal investigation, the Commission can resort to the same investigative tools, outlined in Chapter V of Regulation EC 1/2003, as during the informal stage. The Commission can thus take statements, make inspections, and issue information requests for the purposes of facilitating the decision on whether to issue a *statement of objections* (SO). The Commission may also hold *state of play meetings* with the subjects of the proceedings, the complainants, or third parties to discuss the case, either through its own initiative or at the request of the parties.⁴⁰¹ The SO forms the second major decision point in an EU competition law proceeding. Note though that due to the need for stealth during the investigation of cartel conspiracies, the opening of proceedings and the issuance of the SO usually coincides.⁴⁰² In cartel proceedings, the Commission has already at the point of a formal opening of investigation gathered the needed info to reach a preliminary conclusion and can thus simultaneously issue an SO.

At the SO decision point, the Commission must decide whether to reach a preliminary conclusion of an infringement, or whether to close the case for some or all the parties without issuing an SO. A third procedural option is also available in cases where the Commission does not intend to impose fines; the parties subject to the proceedings can offer commitments that address the concerns of the Commission. If deemed appropriate remedies, the Commission can adopt a decision where it accepts these commitments without concluding about the past

⁴⁰⁰ See further discussion and references on the selection of cases in the Notice on Best Practices; Commission notice on best practices for the conduct of proceedings concerning Articles 101 and 102 TFEU [2011] OJ C308/03, para 13.

⁴⁰¹ See Commission notice on best practices for the conduct of proceedings concerning Articles 101 and 102 TFEU [2011] OJ C308/03, paras 32-69.

⁴⁰² See Commission notice on best practices for the conduct of proceedings concerning Articles 101 and 102 TFEU [2011] OJ C308/03, paras 4 and 24.

or present existence of an infringement. The question of an infringement is thus left inconclusive, but the Commission nonetheless pledges to take no further action if the commitments are kept.⁴⁰³ An SO is only issued after an in-depth investigation and it signals a confident position of the Commission that an infringement has been made. The SO outlines the objections the Commission has towards the behaviour of the parties' subject to the proceedings and invites them to respond to these objections. A non-confidential version of the SO is made public and the complainants and interested third parties are invited to submit their comments on the SO to the Commission.⁴⁰⁴

The issuance of a SO functions as an official accusation that an infringement has been made, and the parties subject to the proceedings are challenged to reply to that accusation through a written reply and through an oral hearing orchestrated by the Commission's Hearing Officer. Additionally, the parties are invited to a state of play meeting to discuss the case. To assess properly the Commission's accusation, the addressees of the SO are granted with access to the Commission's case file prior to the submission of a reply to the SO and prior to the oral hearing. Depending on the content of the written replies, and the results of the oral hearings or any other communication with the parties after the issuance of the SO, the Commission moves towards the final decision point on whether to confirm in part or whole the preliminary conclusion reached in the SO, or whether to drop the case without finding an infringement.⁴⁰⁵

If the Commission finds an infringement and adopts prohibition decision in accordance with Article 7 of Regulation EC 1/2003, it signals the definite conclusion of the Commission. That conclusion can however be challenged before the CJEU. In reaching a conclusion about an infringement to Articles 101 or 102 TFEU the Commission has gone through three major intervals of framing and decision phases.⁴⁰⁶ At each of the three decision points, the

⁴⁰³ Commitment decisions are taken pursuant Article 9 of Regulation EC 1/2003. See further on Commitment decisions; Commission notice on best practices for the conduct of proceedings concerning Articles 101 and 102 TFEU [2011] OJ C308/03, paras 115-133.

⁴⁰⁴ See further Commission notice on best practices for the conduct of proceedings concerning Articles 101 and 102 TFEU [2011] OJ C308/03, paras 84-85 and 104-105.

⁴⁰⁵ See further Commission notice on best practices for the conduct of proceedings concerning Articles 101 and 102 TFEU [2011] OJ C308/03, para 77-114.

⁴⁰⁶ Note though cartel proceedings where the decisions to open investigation and to issue an SO are often taken simultaneously.

Commission can choose to drop a case or to move on towards a conclusion of an infringement. The three decision points are progressively demanding in terms of accuracy of the finding about the existence of an infringement, and each of the progressing framing intervals marks a procedural investment aimed at facilitating ever-increasing accuracy at the subsequent decision points. If already at the first decision point the Commission is confident that it will not be able to find an infringement, it can drop the case and thus save resources. The same applies at the SO decision point. The progression of increasingly demanding decision points thus facilitates the efficient use of enforcement resources.

At each of the three decision points, the major stakeholders in the procedure might have different preferences regarding the baseline accuracy to be pursued. They might also have different preferences about the types of errors conceded at each point. While the preference function of each stakeholder is likely to be persistent through each of the decision phases, the magnitude of the preferences towards a certain outcome increases at each interval, as the stakes get higher. For example, the accused on one hand might emphasize that a false positive error is not made during the decision to open an investigation, but he might feel much stronger about such mistakes at the final decision point where he might potentially receive a hefty fine without being guilty of an infringement. The competitor might on the other hand hold stable preferences about false negative errors as the proceeding progresses towards the final decision point. For him the stake is always the same at each point with the prospect of the case being falsely dropped. The consumer, being neutral about the types of errors conceded, at the same time might emphasize the cost efficiency ratio of the overall accuracy achieved and the corresponding enforcement investment. In that case, the consumer would require that the continuation of a case at each decision point should be beneficial towards reaching the accurate outcome, given the additional resources needed to proceed. The taxpayer would base his preferences on similar cost benefit analysis considerations, but would consider the investment in a more macro-economic manner, rather than just from the perspective of consumer welfare.

After the Commission has concluded the administrative procedure with the imposition of a penalty or a fine, parties to the case that can show *'direct and individual concern'*⁴⁰⁷ with the result can apply to the Court of Justice to review the decision based on Article 31 of Regulation 1/2003 EC and Article 263 TFEU. The provision stipulates that the Court of Justice shall have *'unlimited jurisdiction'* to review the Commission's decisions of that type, and that it may *'cancel, reduce or increase'* the imposed sanctions.

It follows from Article 256 TFEU, that the General Court has a first instance jurisdiction to hear applications to review competition decisions of the Commission, and that the General Court's findings in such cases can only be appealed on points of law to the Court of Justice. This effectively makes the General Court the last instance to argue facts in competition cases, but the Court of Justice is the ultimate instance for arguing on points of law. Article 281 TFEU stipulates that a special protocol should be made for the statute of the CJEU, which can be amended by the Council and the Parliament through the normal legislative procedure. Additionally, Article 253 stipulates that the Court shall establish its rules of procedure that should be approved by the Council. The Statute of the CJEU⁴⁰⁸ and the Rules of Procedure⁴⁰⁹, lay down in details the workings of the procedures before the Court of Justice and the General Court.

The nature of the decision mechanism in competition proceedings before the Court of Justice is different from its nature before the Commission. The role of the Commission is reduced to that of defending its own conclusion at the administrative stage and a separate judicial body overtakes the adjudicative function. The intervals of framing and deciding are however in place; before the General Court decides there is a framing stage that involves written submissions from the parties and an oral hearing before the panel of judges, and before the Court of Justice decides there is also the opportunity to submit written pleadings and to state

⁴⁰⁷ This criterion, which is now stipulated in Article 263 TFEU (ex Article 230 TEC), was further developed in Case 25/62 *Plaumann v Commission* [1963] ECLI:EU:C:1963:1 into what is now known as the *Plaumann test*. The test makes it very hard to judicially challenge a Commission decision unless one is a privileged party to the procedure leading up to it.

⁴⁰⁸ See Protocol (No 3) on the Statute of the Court of Justice of the European Union, annexed to the Treaties, as amended by Regulation (EU, Euratom) No 741/2012 of the European Parliament and of the Council [2012] OJ L228.

⁴⁰⁹ See Rules of Procedure of the Court of Justice [2012] OJ L265, as amended by [2013] OJ L173.

arguments in an oral hearing. Additionally, an *advocate general* normally submits an independent written opinion to the Court of Justice during the latter framing phase at the judicial stage. An important difference between the judicial and the administrative procedure is that an annulment decision by the General Court can be appealed to the Court of Justice on points of law; a decision finding no breach before the General Court does thus not end the procedure, unlike in the procedure before the Commission where the procedure could be terminated at any decision point through a no breach finding.

The preference function of the stakeholders is likely to correspond to the function during the administrative procedure before the Commission, but again, the intensity level is likely to increase as it draws closer to the final and the ultimate instance; before the Court of Justice on points of law, and before the General Court on points of fact.

3.2. Instrumental and normative function

EU's competition law procedures can be analysed based on their instrumental function. As previously explained in chapter five, the instrumental function of procedures consists of two functional elements: the framing function; and the decision function. By analysing the procedural regulations that control EU's competition law procedures it should be possible to identify and categorise provisions based on their function towards facilitating the framing phase of the procedure and in other cases with reference to the decisional phase of the procedure. Alternatively, procedures can be analysed based on their normative function. As previously explained, a procedure seeks to facilitate its substantive agenda with as much accuracy as materially possible; the normative agenda thus becomes accuracy and efficiency.

If this analysis holds, it should be possible to identify four types of procedural provisions: on one hand provisions that instrumentally focus on the framing phase and either seek to facilitate accuracy or efficiency of the framing, and on the other hand provisions that instrumentally focus on the decision phase and seek to facilitate accuracy or efficiency of that functional capacity. To make things a bit more complicated, some provisions might contain hybrid functions in either the instrumental capacity or the normative capacity, or both at the same

time. However, the hybrid provisions should nonetheless be composed of any of these two instrumental and two normative elements.

Some procedural provisions are on their face not obviously procedural in character, as opposed to substantive. This for example applies to provisions providing for reprimand powers and powers to take certain types of decisions, typically located among the procedural provisions. Such provisions can be viewed as being ends rather than means; i.e. a punishment or a specific decision type becomes the consequence the subjects of the procedure strive to avoid. Viewed more broadly however, the tool of reprimand is properly seen as an instrument for achieving a superior objective; a specific type of punishment is thus not an end but instead a means towards achieving something of substantive importance. The same applies to provisions describing the availability of types of decisions; in the intermediate term, they seem substantive in character, but viewed holistically, their procedural character dominates. What matters in this sense, is that in their capacity as intermediary substantive ends, these provisions depend on the rationale of somehow being facilitators of the superior objective. In what follows, I will treat this kind of provisions as procedural towards the grand objective of the competition law regime, not as substantive in the narrow intermediate sense of the procedural regime.⁴¹⁰

3.2.1. The procedural regulations

Through Regulation EC 1/2003, the Council grants the Commission with specific powers to enforce Articles 101 and 102 TFEU. Some of the provisions define the jurisdiction of the Commission both in and of itself and with regards to national competition authorities (NCA), national courts, and the cooperation between the different bodies, i.e. Articles 3, 4, 5, 6, 11, 12, 13, 14, 15, 16, 22, 29 and 32. The transitional, amending and final provisions of the Regulation, i.e. Articles 34-45, have little relevance for our analysis with the exception of Article 35, which commands the Member States to designate a national authority with the task of enforcing Articles 101 and 102. Such designation, signals an investment in enforcement

⁴¹⁰ For further discussion on the distinction between the substantive and the procedural a reference is made to chapter five.

resources that increases the overall enforcement capacity, which should translate into an increased total accuracy of the enforcement regime.

The provisions of Regulation EC 1/2003, that control EU's competition law enforcement procedure, are listed in the following table. Their instrumental and normative function is also defined:

Council Regulation (EC) 1/2003			
Art.	Description	Instrumental Function	Normative Function
2	Burden of proof: The party alleging infringement carries the burden, but reverses if the accused claims an exception	Decisional	Accuracy / Efficiency
7	Prohibition decision: Decision type available to DG COMP	Decisional	Efficiency
8	Interim measures: Decision type available to DG COMP	Decisional	Efficiency
9	Commitments: Decision type available to DG COMP	Decisional	Efficiency
10	Inapplicability: Decision type available to DG COMP	Decisional	Efficiency
17	Sector inquiry: Investigative tool available to DG COMP	Framing	Accuracy / Efficiency
18	Info request: Investigative tool available to DG COMP	Framing	Accuracy / Efficiency
19	Statement: Investigative tool available to DG COMP	Framing	Accuracy / Efficiency
20	Inspection: Investigative tool available to DG COMP	Framing	Accuracy / Efficiency
21	Further inspection: Investigative tool available to DG COMP	Framing	Accuracy / Efficiency
23	Fines: Reprimand available to DG COMP	Decisional	Efficiency
24	Penalties: Reprimand available to DG COMP	Decisional	Efficiency

25	Limitation of penalties: Curbs the reprimand powers of DG COMP	Decisional	Efficiency
26	Limitation of penalties enforcement: Curbs the reprimand powers of DG COMP	Decisional	Efficiency
27	Right to be heard: Investigative duty of DG COMP	Framing	Accuracy
28	Professional secrecy: Investigative duty of DG COMP	Framing	Efficiency
30	Publication of decisions: Decision duty of DG COMP	Decisional	Accuracy
31	Review by the CJEU: Limitation of DG COMP's decision powers	Decisional	Accuracy

The provisions of Regulation EC 773/2004, that control EU's competition law enforcement procedure, are listed in the following table. Their instrumental and normative function is also defined. Supplementary to Regulation EC 1/2003, the Regulation regulates in more details the initiation of Commission proceedings in Article 101 and 102 cases, and the handling of complaints and the conduction of hearings.

Commission Regulation (EC) 773/2004			
Art.	Description	Instrumental function	Normative Function
2	Initiation of proceedings: DG COMP can use investigative powers and reject complaints prior to opening of formal proceedings	Framing / Decisional	Efficiency / Accuracy
3	Power to take statements: Supplementary to powers to investigate in Reg. 1/2003	Framing	Accuracy / Efficiency
4	Questions during inspections: Supplementary to powers to investigate in Reg. 1/2003	Framing	Accuracy / Efficiency
5	Admissibility of complaints: Criteria for submitting complaints	Framing	Accuracy / Efficiency
6	Procedural participation of Complainants: Rights of participation	Framing	Accuracy

7	Rejection of complaints: Procedure for rejection of complaints	Decisional	Efficiency
8	Access to info: Access rights of the complainant	Framing	Accuracy
9	Overlapping complaints: Complaints can be rejected if being processed by another competition authority	Decisional	Efficiency
10	Replies to SO: Supplementary to rights in Reg. 1/2003	Framing	Accuracy
11	Right to be heard: Supplementary to rights in Reg. 1/2003	Framing	Accuracy
12	Right to an oral hearing: Supplementary to rights in Reg. 1/2003	Framing	Accuracy
13	Hearing of others: Supplementary to rights in Reg. 1/2003	Framing	Accuracy
14	Conduct of oral hearings: Supplementary to rights in Reg. 1/2003	Framing	Accuracy / Efficiency
15	Access to file: Supplementary to rights in Reg. 1/2003	Framing	Accuracy
16	Confidentiality: Supplementary to rights in Reg. 1/2003	Framing	Efficiency
17	Time limits: The articulation of various deadlines in a proceeding	Framing / Decisional	Efficiency

The tables above show that many of the provisions of the two main procedural regulations for EU competition law procedures, relate to the accuracy and the efficiency of the framing phases of the procedure by articulating various framing tools at the disposal of the Commission and by conditioning the access of various parties to different sections of the framing phases. Many of the provisions also relate to the efficiency of the decisional phases by defining the types of decisions that can be reached at various decision points, for the ease of the decision makers and the parties alike. Few of the provisions relate to the accuracy of the decisional phases, which still is of great importance for the overall success of the procedure. The provision on the

burden of proof implies a premium on a specific type of accuracy, but does not condition the level of overall accuracy. The exposure of the Commission's decision to the public through publication creates an incentive for decisions to be consistent with common sense, and the potential to appeal the decisions to the CJEU generates similar pressure of accuracy on the decision process.

The relative absence of provisions regulating the accuracy of the decision points in the main procedural regulations, can partly be explained with reference to two reasons: firstly the said procedural regulations mainly deal with what the Commission can decide in the context of competition law, but less with how it decides; and secondly, while individual provisions do perhaps not directly engage the problem of decisional accuracy, the aggregate effect of the system defined in the procedural regulations, organised in three intervals of framing and decisional phases, is towards decisional accuracy. But there is more to it. The provisions that seem absent from the formal procedural regime, have their substitutes in the informal practices of DG Competition.

3.2.2. The procedural practices of DG Competition

The Commission's internal decision-making mechanisms are of considerable importance for procedural accuracy in the context of competition law proceedings. The Commission acts in many different fields of EU law, but the application of the competition provisions is nonetheless in many ways procedurally special in terms of how the decisions are taken. The core Treaty provisions on how the Commission acts are found in Article 17 TEU and Articles 248-250 TFEU, and on their basis, the Commission has adopted the Rules of Procedure of the European Commission.⁴¹¹ Significantly, Article 17(6) TEU orders that the Commission shall act '*as a collegiate body*', and Article 250 TFEU dictates that the '*Commission shall act by a majority of its members*' conditioned on a quorum that in Article 7 of the Rules of Procedure is defined as the majority of the members of the Commission as defined in the Treaties (i.e. Articles 17 TEU and 244 TFEU). The basic structure of the decision mechanism of the Commission is a majority decision by the members of the Commission, i.e. the Commissioners, although in

⁴¹¹ See The consolidated version of the Rules of Procedure of the Commission [2000] OJ L308/26. Note though that the Rules were adopted before the reorganisation of the Treaties following the Lisbon Treaty in 2009 and were thus adopted with references the old Treaty numbers.

practice the Commission usually acts by consensus.⁴¹² The ultimate power to take decisions in the context of competition law proceedings, thus rest with the College of the Commissioners of the Commission.

The Commission is organised into different Directorates General (DG's) which are given different task for which the Commission is responsible. The Directorate General for Competition (DG Competition) is responsible for the enforcement of the competition provisions of the Treaties, including Articles 101 and 102 TFEU. The president of the Commission allocates one of the members of the Commission responsibility for the tasks carried out by DG Competition.⁴¹³ This member prepares proposals and makes suggestions to the president of the Commission, to put proposals relating to the enforcement of the competition provisions on the agenda of the Commission's meetings.⁴¹⁴ DG Competition is further organised into several directorates, which are composed of units, which are further divided into case teams that are responsible for individual cases. The head of unit manages and is responsible for the work of the case team. Any proposed action on behalf of DG Competition, that requires the approval of the Commission, goes through the Director General, the most senior official within DG competition, which makes proposals to the Commissioner responsible for the competition portfolio. The Commissioner, if in agreement with the proposal, requests to get it on the agenda of the Commission's regular meetings.

Before the Commission adopts a final decision in Article 101 and 102 TFEU cases it has gone through several layers of internal framing mechanism that are intended to increase the accuracy of the eventual decision making. The main object of accuracy is the work carried out by the case team and the proposed action the team makes. At the first instance, the case team must convince the Director General to make a proposal to the Commissioner. The Director General can seek advice from the Chief Competition Economist, who reports directly to him, on matters relating to the economics and the econometrics of the proposed action. He can also agree with the Commissioner to set up a peer review panel on selected cases either

⁴¹² Article 8(2) of the Commission's Rules of Procedure implies the practice of deciding by consensus. A '*vote shall be taken if any Member so requests*'. If no one asks for a vote, the Commission decides based on a consensus.

⁴¹³ See Article 17(6)(a) TEU, Article 248 TFEU, and Article 3 of the Rules of Procedure of the Commission.

⁴¹⁴ See Articles 6 and 19 of the Rules of Procedure of the Commission.

before, or after the issuance of an SO. If peer reviewing is decided, a special team is assembled to scrutinise the casefile and the proposed action by the case team, with the purpose of screening for errors and suggest improvements. The peer review team reports to the Director General and provides an input into his decision on whether, or what, to propose to the Commissioner. If the proposal of the case team survives the scrutiny of the Director General, it moves on to the second instance, where the Commissioner must decide on whether to propose it to the College of Commissioners. The Commissioner will normally take note of and base his decision on the input provided by the Chief Economist, the report of the peer reviewing team (if available), and the action proposed by the Director General.⁴¹⁵

The third and the final threshold for adopting a Commission decision on matters of competition law is the College of Commissioners. In addition to the inputs already mentioned, the Commission seeks the advice of the Commission's Legal Service in accordance with Article 23(3) of the Commission's Procedural Rules, and, in accordance with Article 14 of the Regulation EC 1/2003, an advice must be sought from an Advisory Committee composed of the representatives of the NCA of the Member States. The Hearing Officers of the Commission, institutionally organised to be independent of DG Competition, are responsible for monitoring the compliance of the Commission with procedural rights and guarantees throughout the proceedings in competition law cases.⁴¹⁶ Before a final decision is taken, the Hearing Officer reports to the competent Commissioner (and several others) on the compliance with procedural standards and on whether the parties to the case have been given opportunity to exercise their right to be heard with regards to all of the objections raised in the proposed decision.⁴¹⁷ In reaching a final decision in an infringement case of Article 101 or 102 TFEU, the College of Commissioners will thus have access to independent expert opinions on different

⁴¹⁵ The workings of these internal checks and balances mechanisms is explained in a short document titled *'Proceedings for the application of Articles 101 and 102 TFEU: Key actors and checks and balances'* available at the DG Competition's website; <http://ec.europa.eu/competition/antitrust/key_actors_en.pdf> accessed 10 December 2015.

⁴¹⁶ See further on the Hearing Officers; Decision C (2011) 5742 of the President of the European Commission on the function and the terms of reference of the hearing offices in certain competition proceedings [2011] OJ L275/29.

⁴¹⁷ See Article 14 and 16 of Decision C (2011) 5742 on the function and the terms of reference of the hearing offices in certain competition proceedings.

aspects of the proposed decision, which should assist in facilitating an accurate outcome of their deliberation.

The complex formal and informal institutional mechanism of DG Competition to an extent address and regulate the procedural aspect, concerning accuracy of the decision phase, that seems underrepresented in the formal procedural regulations. In the absence of complete transparency in this regard, these self-imposed restraints could become the subject of suspicion by parties' external to the European Commission.

3.2.3. The procedure before the Court of Justice of the EU

The powers of the Court of Justice to review competition decisions of the Commission are based on Article 263 TFEU and more specifically on Article 31 of Regulation 1/2003 EC. The former grants the Court with general powers to review the legality of the Commission's decisions, and the later grants the Court with a more specific unlimited jurisdiction to review competition decisions where a fine or a penalty has been imposed. The decisional function of the Court is conditioned by a requirement of independence of the Court's members, stipulated in Article 19 TEU and Articles 253 and 254 TFEU. Essentially, this means that the Court should be impartial with regards to the stakeholders and their claims. The Court is further required by Article 19 TEU, to ensure that *'that the law is observed'* in the application of the Treaties and the auxiliary secondary legislation. This requirement conditions the Court on the adjudicative mode of deciding; it should not decide based on its on preferences, the Court should observe the previously decided laws and use them to anchor its adjudication.

The Treaties articulate the decisional jurisdiction of the Court, dictate the decisional modality the Court should use to reach its conclusions, and stress the prerequisite of the Court's independence. The details of the decisional mechanism are further described in the Statue of the Court of Justice and in the Rules of Procedure of the Court of Justice. In many ways, these more detailed procedural regulations correspond to the procedural regulations of the Commission previously described with regards to normativity and functionality. The provisions that deal specifically with procedure can be categorised in the same way depending on

whether they deal with the framing or the decisional phase of the Court's procedure, and whether they deal with the accuracy or the efficiency of the two phases.

Before the Court, the quality of the framing process is dependent on the parties to the case. If they provide high quality written submissions and argue their case effectively at the oral hearings, the decisional task of the Court is facilitated. If, however the parties fail to plead their arguments effectively during the written or the oral part of the procedure, it is likely to be reflected in the quality of the Court's outcome, since the Court is to an extent bound by how the case is pleaded before it. The input of the Advocates General increases the quality of the framing phase; it provides the Court with impartial assessment of the case and the pleadings of the parties that it can choose to follow or, if in disagreement, seek an alternative approach.

A two-tiered court procedure is an expensive measure to enhance the decisional accuracy of the procedural process in competition law, especially considering the preceding multi-layered administrative procedure before the Commission, that should already guarantee a high level of accuracy. Judges of a high quality are expensive experts and the necessary institutional infrastructure is expensive as well. The Court of Justice is organised into chambers of judges (3, 5, 15, or the Grand Chamber of all judges) that precede over individual cases, and each judge has a cabinet of supporting staff that assist them in composing the draft judgements in cases where the judge acts as the *judge-rapporteur*, and assist them to take position in cases where other judges act as the *judge-rapporteur*. The organisation of the Court guarantees that each case is scrutinised by all the judges in the chamber and their respective assisting cabinet members. Many legal experts of the highest degree thus review important decisions that are assigned to the Grand Chamber. This organisation facilitates decisional accuracy, although at a high cost in terms of expert resources.

4. The Accuracy of EU's competition procedure

The political peculiarities of the College of Commissioners can be considered as compromising its integrity as a decision-making body in legal adjudication.⁴¹⁸ Although the Commissioners and

⁴¹⁸ Ian Forrester (now judge at the General Court) notes the risk of political bias in the Commission's decision mechanism where the Commissioners, usually politicians, decide the result of a public prosecution. See Ian S.

the Commission are supposed to work independently of any national interests, according to Article 17(3) TEU and 245 TFEU, they are usually chosen from the top layer of the political elites of the respective Member States. As such, the Commissioners are usually apt at political decision-making, where interests and preferences reign supreme as motives of action. In many fields of the Commission's work and decisions making, a political background is very useful since large parts of the Commission's competences concern policy and policymaking. However, as previously explained, the process of legal adjudication and political decision-making is distinct in important ways. While special interest and preferences are supposed to be inputs in political decisions with an *ex-ante* perspective, legal adjudication is on the contrary supposed to be based primarily on a rational inference from the relevant facts and the applicable laws from an *ex-post* perspective.⁴¹⁹

While the Commissioners of the Commission are often not specialist in the classic process of legal adjudication, and despite that all but one of them deal primarily with other issues than the competition portfolio, the institutional design of the competition law procedure still provides some guarantees that the outcome of individual competition law cases is decided based on the legal adjunction modality, rather than based on the political modality. The key guarantee is the one provided for in Article 263 TFEU and Article 31 of Regulation EC 1/2003 regarding the possibility to have the competition law decisions of the Commission reviewed by the Court of Justice of the EU. The requirement in Article 31 of Regulation EC 1/2003 that the Commission's decisions in competition law proceedings shall be published has a similar effect, although weaker. Unlike the Commission, the Court of Justice is required by Article 19 TEU to use the decision modality of legal adjudication in its interpretation of the competition provisions of the Treaties; *'[the CJEU] shall ensure that in the interpretation and application of the Treaties the law is observed.'* Naturally this requirement on the decision body at the final instance of a competition law proceeding has effects on all decision points leading up to the

Forrester, 'Due process in EC competition cases: a distinguished institution with flawed procedures' (2009) 34 European Law Review 817, 822 and 831-33.

⁴¹⁹ In politics, we think about what we should do in the future based on the current desires of people, while in the court room we think about what happened in the past and what were the applicable laws as previously decided. The distinction is in practice more nuanced and not always clear-cut, but as a rule of thumb this is a useful distinction, especially concerning what kind of thought processes ought to guide the different types of decision making.

final instance. The effect is however not uniform since only some decision types are caught by the effect during the administrative procedure before the Commission.

As previously explained, the Commission proceedings in Article 101 and 102 TFEU cases are organised in three intervals of framing and decision phases. At each interval, there is a decision point where the Commission can decide to drop the case, or to continue towards an infringement decision. At each of these decision points, the potential of an eventual review by the CJEU forces a legal adjudicative modality of deciding, if the Commission opts to progress towards an infringement decision. A decision to drop a case is however not subject to the same pressure; the case is simply closed without any further possibilities to have that decision reviewed.⁴²⁰ This has implications for the accuracy of the decision making; while the framing phase might be perfectly executed in terms of accuracy by the Commission's administrative services, the actual act of deciding might fail to produce the accurate result due to the final decision maker's call to drop a case on political grounds, instead of inferring a decision from the relevant facts and the applicable laws. There are, in other words, checks and balances that prevent the Commission from intentionally making false positive errors in competition cases, but less so if the Commission elects to make a false negative error.

Viewed differently but still within the accuracy/efficiency paradigm there might be, and usually are, valid efficiency reasons behind the Commission's decision to drop cases. Even if the facts and the law recommend the pursuance of an infringement decision, the limits of the Commission's enforcement budget might provide a valid reason for closing a case and concentrate the resources elsewhere. Such a decision can be taken validly based on the political modality of decision making, rather than based on the legal adjudicative modality. A system can be envisaged, where the Commission would be required to take a formal reasoned decision irrespective of the finding of an infringement or not, which in both cases could be

⁴²⁰ By citing 'community interests' the Commission has discretion to prioritise its investigative efforts in competition cases and close cases that do not qualify as of interest for the EU community. Note though the criteria for stating reasons articulated in the *Automec* case law. See Case T-24/90 *Automec v Commission* [1992] ECLI:EU:T:1992:97, paras 85-86; and more recently Case T-342/11 *CEES v Commission* [2014] ECLI:EU:T:2014:60, paras 58-60: '*As regards the factors determining the exercise of the discretion enjoyed by the Commission under Article 9(2) of Regulation No 1/2003, it should be remembered that the latter has only limited resources, which it must use in taking action against a potentially wide range of behaviour which is contrary to competition law.*'

appealed to the CJEU.⁴²¹ Such a system would force an adjudicative decision modality for both negative and positive infringement decisions, and thus lessen the probabilities of both false negatives and false positives. The choice of institutional design of not having such system in place can simply be viewed as a prioritisation of the available resources. The current institutional system thus emphasises the avoidance of false positives by requiring the legal adjudicative modality of taking decisions for reaching positive infringement decisions, while false negatives are *prima facie* permissible with reference to the political modality of deciding that permits political arguments to influence the conclusion.

By examining the framing phases, prior to each of the decision points, an evolving perspective on the information that are the object of accuracy quickly reveals. The framing phase prior to the opening of a formal investigation emphasises the gathering of raw information data, uncontaminated by the agenda of the stakeholders involved. The investigative powers articulated in Regulation EC 1/2003 are meant to facilitate this initial fact-finding. The first decision point is thus primarily based on the investigators perspective on the relevant information. The intermediary decision point of the SO relies on the same investigative tools, but allows for more input by the stakeholders and is thus not as concerned with the element of surprise as often is important during the very early stage of cases involving sophisticated conspiracies. The SO thus reflects the investigators perspective, slightly informed by the stakeholders. Before the final decision point, the Commission grants access to its preliminary findings and the sources underscoring that finding, and invites the parties to the case and other interested stakeholders to exercise their right to be heard before the case is concluded with a final decision. The final decision at the administrative level should thus reflect both the investigator's perspective and the stakeholder's perspective on the relevant information. The court procedure subsequently takes a fresh neutral view on the law and the facts that incorporates the arguments of those who participate in the court procedure at the first instance, and then again on the law at the last instance.

⁴²¹ Once the Commission has opened a formal investigation into an alleged infringement to the state aid provisions of the Treaties, i.e. Articles 107-108 TFEU, it must close the case with a formal reasoned decision on the existence of an incompatible aid or not. Such decisions can be contested and annulled by the CJEU.

It is important to note that while the right to be heard is potentially an important source of additional information for the facilitation of an accurate result, it can also be a venue for the parties to mislead and confuse the case. A guilty defendant facing a hefty antitrust fine will have an enormous incentive to spend his resources on attempts to confuse and divert the decision maker towards making a false positive error. Countering such attempts drains enforcement resources and leaves less for other enforcement uses. The appropriate timing and quantity of the execution of the right to be heard are thus important considerations, both in terms of ultimate accuracy of the result and the efficiency with which it is reached.

I have now analysed the broad structure of EU's competition law procedure and identified the main instrumental elements that seek to facilitate the framing of relevant information for the taking of decisions that aim at producing efficiently as accurate results as materially possible. Some of the regulatory provisions are concerned with efficiency, while others are concerned with accuracy. Some are decisional instruments, while others are framing instruments. When combined in a holistic procedural regime, the hope of the regulator is that the aggregated average result from individual cases is at an equilibrium in terms of the result accuracy achieved and the cost with which that level is achieved. In complex procedures, such as EU competition procedures, an equilibrium is to an extent necessary at each major decision point. The equilibrium will however not be the same at each point; as the procedure progresses towards a final finding of an infringement, the stakes get proportionally higher for some stakeholders, that justifies an increased weight of their preferences in the fair equilibrium solution.

Applied fairness in EU's competition procedure

Publicly accessible documents from competition law litigation before the Luxembourg courts provide an insight, on not only what the law is on specific issues, but also on what the parties to the cases want the law to be on specific issues. If the issue at hand is undetermined, or vaguely determined in the *ex-ante* sense, the parties argue essentially how the respective rule *ought to be* based on their respective preferences. In such cases the courts also assume the *ex-ante* legislative perspective and determine what the rule *ought to be* for the current and future instances, rather than identifying what the rule is *per se*. The advocates general (AG's) of the Court of Justice, assume a similar role in their opinions on legally undetermined questions, and thus reveal their preferences on how a specific issue should be resolved in the *ex-ante* legislative sense.

In this chapter, I shall begin by quickly reviewing the case law of the CJEU on Article 47 of the Charter in the context of competition law proceedings. Article 47 of the Charter is after the entering into force of the Lisbon Treaty, the primary source of reference for those claiming rights of the procedural fairness type, and thus the appropriate context to look for arguments relating to preferences over the fair design of competition procedures. Following this quick review, I will identify two previously undecided fairness problems in separate lines of cases, analyse the arguments in the court documents in more depth, and try to reveal how these arguments fit with the dichotomy of accuracy and efficiency of procedures. If the analysis outlined in the previous chapters holds, we should be able to predict how the parties argue in terms of emphasis on either accuracy or efficiency. More interestingly, we should also be able to reveal the methodology of how the AG's and the Court approach this problem of balancing.

After having studied two problems of procedural fairness in the case law on Article 47 of the Charter and examined how they have been resolved by the Court, the later part of the chapter will display how the *model of fair rules* can be utilised to resolve these dilemmas. This requires that the arguments of the parties to the cases be translated into the lingo of preferences and utility, and that the dilemma to be defined in terms of stakes, stakeholders, and the balancing of specific interests. In conclusion, an assessment will be made on similarities or differences of

the outcome of the actual process before the Court, and the potential outcome that the *model of fair rules* would produce.

1. The case law on Article 47 of the Charter of Fundamental Rights

1.1. Review of the cases⁴²²

The first mention of Article 47 of the Charter in the competition case law of the CJEU, after the entering into force of the Lisbon Treaty, was in the judgment in Case C-407/08 P *Knauf Gips v Commission*, which was handed down on 1 July 2010. In that case, the applicant noted, in relation with his third ground of appeal, that the Court of First Instance (the CFI) had in paragraph 359 of the judgement under appeal⁴²³ found that the applicant was estopped from arguing that the Commission's decision was erroneously addressed, since the applicant had not raised the issue during the administrative proceeding. The CFI had argued that the applicant should have been aware of this problem since the issuance of the statement of objections and could, and should, thus have raised this issue earlier. The applicant considered this in breach of the '*in dubio pro reo*' principle which dictates that doubt should be interpreted in favour of the defendant.⁴²⁴

Advocate General Mazák opined that the Court of First Instance erred in law by finding the applicant estopped from making the argument during the court proceeding and suggested the judgment be set aside with regards to the applicant.⁴²⁵ The CJEU agreed with the AG Mazák and referred to Article 47 of the Charter of Fundamental Rights in support of that view: '*In the absence of a specific legal basis, such a restriction is contrary to the fundamental principles of the rule of law and of respect for the rights of the defence. Moreover, the rights to an effective remedy and of access to an impartial tribunal are guaranteed by Article 47 of the Charter of Fundamental Rights of the European Union which, under the first subparagraph of Article 6(1) TEU, has the same legal value as the Treaties. Under Article 52(1) of that charter, any limitation on the exercise of the rights and freedoms recognised by the charter must be provided for by law.*'⁴²⁶ The CJEU subsequently set the judgment of the CFI aside with regards to the applicant.

⁴²² The review is based on a search on the Court's website - <http://curia.europa.eu/> - in November 2016.

⁴²³ See Case T-52/03 *Knauf Gips v Commission* [2008] ECLI:EU:T:2008:253.

⁴²⁴ See Case C-407/08 P *Knauf Gips v Commission* [2010] ECLI:EU:C:2010:389, para 59.

⁴²⁵ See Case C-407/08 P *Knauf Gips v Commission* [2010] ECLI:EU:C:2010:70, opinion of AG Mazák, paras 96-97.

⁴²⁶ See Case C-407/08 P *Knauf Gips v Commission* [2010] ECLI:EU:C:2010:389, para 96.

On 8 December 2011, the CJEU handed down highly anticipated judgements in three appeal cases, concerning a cartel in the market for copper plumbing tubes and in the market for copper industrial tubes.⁴²⁷ These cases are usually denoted as the *KME-Chalkor* cases⁴²⁸ with reference to the applicants in the three appeal cases.⁴²⁹ The applicants each argued that by referring to the Commission's margin of discretion with regards to certain factual findings, the General Court had failed to instigate a full judicial review of the Commission's cartel decision. This, they argued, amounted to a breach of their procedural rights as defined by Article 6 of the ECHR and Article 47 of the Charter. One of the main points of debate during the litigation concerned the categorisation of the Commission's competition enforcement proceeding. The applicants argued that the proceedings should be categorised as a criminal proceeding for the purposes of Article 6 of the ECHR, which accordingly would move the procedure within the scope of the procedural guarantees provided therein. The Commission however relied on a recent case law of the European Court of Human Rights (ECtHR), which seemed to categorise competition proceedings as being not fully criminal in character for the purposes of the procedural guarantees of Article 6 ECHR. Advocate General Sharpston agreed with the Commission's reading of the ECtHR's case law in her opinion. She further opined that while an administrative procedure, such as the Commission's competition enforcement procedure, might not satisfy the strict conditions of Article 6 ECHR when imposing substantial antitrust fines, it might nonetheless be sufficient for the purposes of the ECHR, that all aspects of the result of such procedure could be reviewed by a body that complies with the strict adequacy conditions for criminal proceedings.⁴³⁰

The approach to the case taken by the Court of Justice in its judgements resembles the one suggested by AG Sharpston. The Court implicitly recognised that to comply with the standards

⁴²⁷ The Commission split its initial proceeding against the copper tube cartel into three cases: Case COMP/E-1/38.069 (Copper plumbing tubes), Case COMP/E-1/38.121 (Fittings) and Case COMP/E-1/38.240 (Industrial tubes).

⁴²⁸ See also the discussion on this litigation, its prelude, and a review of the relevant literature in chapter one.

⁴²⁹ The applicants were respectively KME Germany AG, KME France SAS, KME Italy SpA and Chalkor AE Expexergasias Metallon. The KME group appealed in the context of the Commission's inquiry into the plumbing tubes cartel (Case C-389/10 P) and the industrial tubes cartel (Case C-272/09 P), but Chalkor appealed in the context of the plumbing tubes cartel (Case C-386/10 P).

⁴³⁰ See Case C-272/09 P *KME Germany and Others v Commission* [2011] ECLI:EU:C:2011:63, opinion of AG Sharpston, para 67.

of procedural fairness as articulated in Article 6 ECHR and Article 47 of the Charter, the Commission's competition procedure for imposing substantial antitrust fines had to be subject to the potential of a full judicial review at the request of the parties. The problem was however that ever since the landmark judgment of *Consten and Grundig* in 1966, the Commission was widely considered to enjoy certain discretion in establishing economic facts in competition proceedings, although the *Tetra Laval* judgement of 2005 had signalled a more conditional approach to that discretion.⁴³¹ The deference doctrine was grounded on the following phrase in *Consten and Grundig*, that was routinely referred to in the subsequent case law of the Court up until the *KME-Chalkor* litigation; '*[f]urthermore, the exercise of the Commission's powers necessarily implies complex evaluations on economic matters. A judicial review of these evaluations must take account of their nature by confining itself to an examination of the relevance of the facts and of the legal consequences which the Commission deduces therefrom.*'⁴³²

The Court's case law thus created an obstacle to upholding the legitimacy of the current institutional architecture in cartel fining procedures with regards to considerations of procedural fairness as laid down in the ECHR and the Charter of Fundamental Rights. In a somewhat unexpected move, the Court solved the dilemma by simply acting as if the Commission never enjoyed the discretion implied by the case law building on *Consten and Grundig*; '*[a]s regards the review of legality, the Court of Justice has held that whilst, in areas giving rise to complex economic assessments, the Commission has a margin of discretion with regard to economic matters, that does not mean that the Courts of the European Union must refrain from reviewing the Commission's interpretation of information of an economic nature. Not only must those Courts establish, among other things, whether the evidence relied on is factually accurate, reliable and consistent but also whether that evidence contains all the information which must be taken into account in order to assess a complex situation and whether it is capable of substantiating the conclusions drawn from it.*'⁴³³

⁴³¹ See para 39 of Case C-12/03 P *Commission v Tetra Laval* [2005] ECLI:EU:C:2005:87.

⁴³² See Joined Cases 56 and 58/64 *Consten and Grundig v Commission* [1966] ECLI:EU:C:1966:41, p 347.

⁴³³ Each of the three judgments contains a paragraph containing this text. See Case C-272/09 P, para 94; Case C-386/10 P, para 54; and Case C-389/10 P, para 121.

With the doctrine of deference out of the way, the Court could quickly fend off the challenge with a reference to the Treaties and Regulation 1/2003: *'The review provided for by the Treaties thus involves review by the Courts of the European Union of both the law and the facts, and means that they have the power to assess the evidence, to annul the contested decision and to alter the amount of a fine. The review of legality provided for under Article 263 TFEU, supplemented by the unlimited jurisdiction in respect of the amount of the fine, provided for under Article 31 of Regulation No 1/2003, is not therefore contrary to the requirements of the principle of effective judicial protection in Article 47 of the Charter.'*⁴³⁴

Sensing that the General Court had in the judgment under appeal assumed the validity of the deference doctrine, established in *Consten and Grundig*, the Court made a brief detour in its reasoning to explain what its deputy court had really meant: *'It must be noted in that regard that although the General Court repeatedly referred to the "discretion", the "substantial margin of discretion" or the "wide discretion" of the Commission, including in paragraphs 35 to 37, 92, 103, 115, 118, 129 and 141 of the judgment under appeal, such references did not prevent the General Court from carrying out the full and unrestricted review, in law and in fact, required of it.'*⁴³⁵

Given that the deference doctrine of *Consten and Grundig* still existed in some capacity prior to the *KME-Chalkor* litigation, the judgment signalled a definite departure.⁴³⁶ Viewed for what it was in the analytical terms of legal realism, this effectively meant that the Court amended the prevailing procedural system by reducing the deference doctrine to ensure compatibility with Article 47 of the Charter. The Court thus chose to slightly modify the procedural regime

⁴³⁴ See Case C-272/09 P *KME v Commission* [2011] ECLI:EU:C:2011:810, para 106.

⁴³⁵ See Case C-272/09 P *KME v Commission* [2011] ECLI:EU:C:2011:810, para 109. Nazzini is critical of this approach of the Court; *'Three arguments have been put forward to "save" the current system. The first is that, in certain cases, the Union courts do not in fact apply the "manifest error" test – although they say that they do – but carry out a comprehensive review of the evidence. This seems to be the approach adopted by the Court of Justice itself in KME, where the Court said that, although the General Court had set out a test of deferential review, it did in fact carry out a full review of the Commission decision. This approach is far from satisfactory.'* See Renato Nazzini, 'Administrative enforcement, judicial review and fundamental rights in EU competition law: a comparative contextual-functional perspective' (2012) 49 *Common Market Law Review* 971, 997.

⁴³⁶ Arguably, the Courts retreat from the deference doctrine with regards to complex economic facts started earlier with a narrowing of its scope following *Tetra Laval* of 2005.

under the cover of its interpretive powers, instead of deeming the regime unconstitutional which would have referred the issue to the legislator.

On 6 November 2012, the Grand Chamber of the Court of Justice issued a preliminary ruling in case C-199/11 *Otis and Others*.⁴³⁷ The case concerned an action for damages taken by the Commission against the members of a cartel that the Commission had previously found in its decision to be in breach of Article 81 EC (now Article 101 TFEU). Coincidentally a member of the cartel had provided the Commission with goods and services that were the object of the cartel (installation of elevators) and thus the Commission took a private action against the cartel member before a Belgian court. The Belgian court was unsure whether it was compatible with the procedural guarantees of Article 47 of the Charter that it was by Article 16 of Regulation 1/2003 bound to respect the result of the Commission's cartel decision in establishing the facts in the damages case initiated by the very same European Commission. The Belgian court thus referred questions to the Court of Justice for a preliminary ruling. The main issues raised in this context were whether a national court, obliged to respect a cartel decision by the Commission, would be able to reach an independent decision in a damages case for the purposes of Article 6 ECHR and Article 47 of the Charter.

In its ruling, the Court of Justice explained how the rule contained in Article 16 of Regulation 1/2003 was a jurisdictional rule, which did not interfere with the right of access to an independent tribunal under Article 47 of the Charter.⁴³⁸ The Court went on to explain that the correct forum for reviewing the finding of a breach of Article 101 TFEU was through the procedure established in the Treaties, which in any case has been considered compatible with the standards of Article 47 of the Charter in terms of independence and impartiality. The Court also explained that although the national court had to consider the breach of Article 101 TFEU as a fact, it still enjoyed complete independence in establishing the existence of a financial loss and in establishing a causal link between the loss and the breach of the competition rules.⁴³⁹

⁴³⁷ See Case C-199/11 *Otis and Others* [2012] ECLI:EU:C:2012:684.

⁴³⁸ See Case C-199/11 *Otis and Others* [2012] ECLI:EU:C:2012:684, paras 54-55.

⁴³⁹ See Case C-199/11 *Otis and Others* [2012] ECLI:EU:C:2012:684, paras 64-67.

On 11 July 2013, the Court of Justice handed down a judgment in case C-439/11 P *Zeigler v Commission*.⁴⁴⁰ The applicant contended that the Commission had failed to state adequate reasons for its cartel decision in contravention of Article 47 of the Charter. The Court found that the applicant had failed to establish the argument at the first instance before the General Court and thus found the plea inadmissible.⁴⁴¹ However, the Court also noted, with reference to its finding in *Solvay v Commission*, which the Commission was not a tribunal for the purposes of Article 47 of the Charter. Any faults in the administrative procedure leading up to the finding of restrictive practices should thus be dealt with in the context of Article 41 of the Charter, which enshrines the right to a good administration.⁴⁴²

Article 47 of the Charter was also mentioned in the context of the Court's judgment of 8 May 2013 in case C-508/11 P *Eni v Commission* and of 18 July 2013 in case C-499/11 P *Dow Chemical and Others v Commission*, but the Court rejected the arguments made by the applicants in both cases as unfounded.⁴⁴³

On 18 July 2013, the Court of Justice reinforced its conclusion from the *KME-Chalkor* cases in a decision in case C-501/11 P *Schindler Holding and Others v Commission*.⁴⁴⁴ The applicant relied on a similar argument as the applicants had done in the *KME-Chalkor* cases. In its conclusion, the Court pointed to a development in the case law of the ECoHR, which reinforced its earlier approach:

'In paragraph 59 of its judgment in A. Menarini Diagnostics v. Italy, the European Court of Human Rights explained that, in administrative proceedings, the obligation to comply with Article 6 of the ECHR does not preclude a 'penalty' from being imposed by an administrative authority in the first instance. For this to be possible, however, decisions taken by administrative authorities which do not themselves satisfy the requirements

⁴⁴⁰ See Case C-439/11 P *Zeigler v Commission* [2013] ECLI:EU:C:2013:513.

⁴⁴¹ See Case C-439/11 P *Zeigler v Commission* [2013] ECLI:EU:C:2013:513, para 128.

⁴⁴² See Case C-439/11 P *Zeigler v Commission* [2013] ECLI:EU:C:2013:513, para 154. The Court had already established the applicability of Article 41 of the Charter to competition proceedings before the Commission in two Grand Chamber judgments handed down on 25 October 2011 in cases C-109/10 P *Solvay v Commission* [2011] ECLI:EU:C:2011:686, para 53; and C-110/10 P *Solvay v Commission* [2011] ECLI:EU:C:2011:687, para 48.

⁴⁴³ See Case C-508/11 P *Eni v Commission* [2013] ECLI:EU:C:2013:289, paras 50-52 and 69; and Case C-499/11 P *Dow Chemical and Others v Commission* [2013] ECLI:EU:C:2013:482, paras 59-60.

⁴⁴⁴ See Case C-501/11 P *Schindler Holding and Others v Commission* [2013] ECLI:EU:C:2013:522.

laid down in Article 6(1) of the ECHR must be subject to subsequent review by a judicial body that has full jurisdiction. The characteristics of such a body include the power to quash in all respects, on questions of fact and law, the decision of the body below. The judicial body must in particular have jurisdiction to examine all questions of fact and law relevant to the dispute before it. [new para] Ruling on the principle of effective judicial protection, a general principle of European Union law to which expression is now given by Article 47 of the Charter and which corresponds, in European Union law, to Article 6(1) of the ECHR, the Court of Justice has held that, in addition to the review of legality provided for by the FEU Treaty, the European Union judiciary has the unlimited jurisdiction which it is afforded by Article 31 of Regulation No 1/2003, in accordance with Article 261 TFEU, and which empowers it to substitute its own appraisal for the Commission's and, consequently, to cancel, reduce or increase the fine or periodic penalty payment imposed (Chalkor v Commission, paragraph 63).⁴⁴⁵

Following the *Menarini*⁴⁴⁶ case of the ECoHR it was clear that the Commission could, in its capacity as an administrative body, impose substantial cartel fines of a criminal nature, as long as an independent and impartial tribunal, in terms of Article 6 of the ECHR, had the power to review all aspects of the fining decision. Following *KME-Chalkor*, the General Court was considered to have such unlimited jurisdiction to review the Commission's cartel fining decisions based on Article 263 TFEU and Article 261 TFEU in conjunction with Article 31 of Regulation 1/2003. The Court did not concede explicitly that the Commission's cartel fines were of a criminal nature; just short of such concession, it explained that even if they were criminal penalties, it would not in itself be a breach of Article 6 ECHR.⁴⁴⁷

On 26 November 2013, the Grand Chamber of the Court of Justice handed down three judgments in appeal cases relating to an earlier cartel decision of the Commission⁴⁴⁸ concerning the market for industrial bags. In each of the three cases: C-58/12 P *Groupe*

⁴⁴⁵ See Case C-501/11 P *Schindler Holding and Others v Commission* [2013] ECLI:EU:C:2013:522, paras 35-36. See also paras 33-34.

⁴⁴⁶ See Case No 43509/08 A. *Menarini Diagnostics S.r.l. v. Italy* [2011] ECLI:CE:ECHR:2011:0927JUD004350908.

⁴⁴⁷ See Case C-501/11 P *Schindler Holding and Others v Commission* [2013] ECLI:EU:C:2013:522, paras 33-34.

⁴⁴⁸ See the Commission's Decision in Case COMP/38354 – *Industrial Bags* [2005]

Gascogne v Commission;⁴⁴⁹ C-50/12 P *Kendrion v Commission*;⁴⁵⁰ and C-40/12 P *Gascogne Sack Deutschland v Commission*,⁴⁵¹ the applicants pleaded that an excessive delay in the General Court's procedure had breached their procedural rights under Article 47 of the Charter of Fundamental Rights. The procedure in each case had lasted for more than five years, and there was more than a three-year delay between the end of the written court procedure, and the opening of the oral court procedure. Without setting in stone the maximum permissible procedural time, the Court did not hesitate to conclude that '[t]he length of that period [i.e. between the written and the oral procedure] cannot be explained by the circumstances of the case, whether it be the complexity of the dispute, the conduct of the parties or the supervening procedural matters.'⁴⁵²

Due to the obvious breach to the rights of the parties to have a hearing within a reasonable time, as dictated by Article 47 of the Charter, the main issue of the case became to resolve how such breaches should be best dealt with in the context of competition procedures. The case was allocated to the Grand Chamber of the Court due to two incompatible alternatives of a solution in the existing case law. As a starting point the Court noted, with reference to its earlier case law, that since the delay in the proceeding would not have changed the outcome of the cases, the judgments under appeal would not be set aside.⁴⁵³ The question was thus, how the applicants should be compensated for the delay, without altering the substantive result of the cartel case. In case C-185/95 P *Baustahlgewebe v Commission*, the Court 'for reasons of economy of procedure and in order to ensure an immediate and effective remedy' decided in a similar situation to reduce the imposed fine.⁴⁵⁴ In case C-385/07 P *Der Grüne Punkt - Duales System Deutschland v Commission*, which did not involve a fine, the Court had taken a different approach by suggesting that the applicants needed to file a claim for damages against the Community in a separate new case.⁴⁵⁵

⁴⁴⁹ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:770.

⁴⁵⁰ See Case C-50/12 P *Kendrion v Commission* [2013] ECLI:EU:C:2013:771.

⁴⁵¹ See Case C-40/12 P *Gascogne Sack Deutschland v Commission* [2013] ECLI:EU:C:2013:768.

⁴⁵² See: Case C-40/12 P *Gascogne Sack Deutschland v Commission* [2013] ECLI:EU:C:2013:768, para 98.

⁴⁵³ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:770, paras 74-75.

⁴⁵⁴ See Case C-185/95 P *Baustahlgewebe v Commission* [1998] ECLI:EU:C:1998:608, para 48.

⁴⁵⁵ See Case C-385/07 P *Der Grüne Punkt - Duales System Deutschland v Commission* [2009] ECLI:EU:C:2009:456, para 195.

The opinion of AG Sharpston in *Groupe Gascogne v Commission* contains information about an unusual juridical methodology mustered by the Court in reaching its conclusion: ‘*The Court invited the 27 Member States, the European Parliament and the Council to indicate in writing their views on the approach taken in, respectively, Baustahlgewebe and Der Grüne Punkt. Seven Member States indicated a preference for the former, three favoured the latter and six Member States expressed no preference. The Council endorsed Baustahlgewebe whilst acknowledging that the two remedies coexist and neither is perfect. The European Parliament considered the Der Grüne Punkt approach to be better.*’⁴⁵⁶ The existence and the execution of this survey indicates how uncertain the Court was about which approach to take in the cases under appeal.

In the end, the Court followed the suggestion of AG Sharpston⁴⁵⁷ and adopted the *Der Grüne Punkt* approach, thus rejecting the pleas of the applicants in the three cases and instructing them to seek damages in a new action before the General Court.⁴⁵⁸ The decisive argument seems to have been the universal character of *Der Grüne Punkt* approach; ‘*a claim for damages brought against the European Union pursuant to Article 268 TFEU and the second paragraph of Article 340 TFEU constitutes an effective remedy of general application for asserting and penalising such a breach, since such a claim can cover all the situations where a reasonable period of time has been exceeded in proceedings.*’⁴⁵⁹ In comparison, the approach in *Baustahlgewebe* is limited to instances where a fine has been imposed, and relies on the Court’s own statute as a legal basis, instead of articles of the Treaties. Notably the argument of procedural economy of the *Baustahlgewebe* approach failed.

In the months following the Grand Chamber judgements in *Groupe Gascogne* and *Gascogne Sack Deutschland*, several other cases concerning lengthy procedure before the General Court

⁴⁵⁶ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:360, opinion of AG Sharpston, para 119.

⁴⁵⁷ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:360, opinion of AG Sharpston, paras 127, 129, 131-32 and 150.

⁴⁵⁸ The General Court has reached a conclusion in the first of the damages cases and the others are pending at the time of writing (Jan 2017). See Case T-577/14 *Gascogne Sack Deutschland and Gascogne v European Union* [2017] ECLI:EU:T:2017:1.

⁴⁵⁹ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:770, para 82.

were resolved using the same rationale.⁴⁶⁰ In the last of these cases, C-580/12 P *Guardian Industries and Guardian Europe v Commission* decided on 12 November 2014, Advocate General Wathelet made a last effort to oppose the approach taken in *Groupe Gascogne* by opining that *'the appropriate mechanism for remedying a breach by the General Court of the reasonable time principle in a case such as the present, would, for reasons of economy of procedure and also to ensure an immediate and effective remedy, be to reduce the fine rather than to leave it to the parties to bring an action for damages before the General Court which, necessarily, will have been found to have failed to observe that principle by being unable to deliver its judgment within a reasonable time.'* In its judgement, the Court disregarded the AG's opinion and maintained its approach in the *Gascogne Sack Deutschland* case.⁴⁶¹

In case C-295/12 P *Telefónica and Telefónica de España v Commission*, decided by the Court on 10 July 2014, the applicants argued that the proceedings before the General Court had taken unreasonably long time and that the General Court had failed to exercise its unlimited jurisdiction to review the Commission's decision, both in breach of Article 47 of the Charter. The Court dismissed both pleas, the former as unfounded and the later on grounds of lacking information to substantiate the claim of undue delays in the procedure.⁴⁶² An indication as to why the appeal was unsuccessful can be found in paragraph 23 of the judgment, where the Court summarises one of the Commission's arguments; *'the Commission contends that the appeal is extremely long and repetitive and frequently sets out a number of pleas on every page, so that it appears to contain several hundred pleas, which amounts to a record in the history of proceedings before the EU courts.'*⁴⁶³

In a judgment in case C-67/13 P *CB v Commission*,⁴⁶⁴ decided on 11 September 2014, the Court of Justice set the General Court's judgment under appeal aside. The Court found that the

⁴⁶⁰ See judgements in cases: C-238/12 P *FLSmidth v Commission* [2014] ECLI:EU:C:2014:284; C-578/11 P *Deltafina v Commission* [2014] ECLI:EU:C:2014:1742; C-243/12 P *FLS Plast v Commission* [2014] ECLI:EU:C:2014:2006; and C-580/12 P *Guardian Industries and Guardian Europe v Commission* [2014] ECLI:EU:C:2014:2363.

⁴⁶¹ See Case C-580/12 P *Guardian Industries and Guardian Europe v Commission* [2014] ECLI:EU:C:2014:2363, paras 17-21.

⁴⁶² See Case C-295/12 P *Telefónica and Telefónica de España v Commission* [2014] ECLI:EU:C:2014:2062, paras 58-60, 63, and 68-69.

⁴⁶³ See Case C-295/12 P *Telefónica and Telefónica de España v Commission* [2014] ECLI:EU:C:2014:2062, para 23.

⁴⁶⁴ See Case C-67/13 P *CB v Commission* [2014] ECLI:EU:C:2014:2204.

General Court had failed to observe the standard of judicial review established in the *KME-Chalkor* case law by giving the Commission too much discretion with regards to the assessment of certain economic facts. In a harsh language, the Court found that the General Court had shown indications of ‘*a general failure of analysis*’ that ‘*reveal[d] the lack of a full and detailed examination of the arguments of the appellant and of the parties*’.⁴⁶⁵

In case C-434/13 P *Commission v Parker Hannifin Manufacturing and Parker-Hannifin*, decided on 18 December 2014, the General Court in the judgment under appeal had exercised its unlimited jurisdiction to review a fine imposed by a decision of the Commission. In doing so, the Court of Justice however found that the General Court had ‘*failed to provide the information necessary to enable the parties concerned to understand why it had set at that level the amount of the fine attributable to Parker-Hannifin, and, moreover, to enable the Court to review the lawfulness of that reduction, in the light, inter alia, of the principle of equal treatment as invoked by the Commission.*’ The Commission’s appeal was subsequently upheld by the Court of Justice due to a breach by the General Court of its obligation to state reasons.⁴⁶⁶

In case C-583/13 P *Deutsche Bahn and Others v Commission*, decided on 18 June 2015, the applicants pleaded that their right to an effective judicial protection under Article 47 of the Charter had been breached by the General Court. They contended that they should have been able to have the Commission’s inspection decision, reviewed by the General Court prior to its execution, instead of post-inspection as the General Court had found to be sufficient. Based on the possibility to have a post-inspection judicial review of the Commission’s decision to undertake an inspection, and based on the possibility to prohibit the Commission to use documents and evidences obtained in an inspection that later is found to have been irregularly conducted, the Court of Justice rejected the plea.

1.2. The main fairness issues of the case law

By reading the cases concerning Article 47 of the Charter of Fundamental rights in competition proceedings found in case registry of the Court of Justice, two clusters of cases appear: first,

⁴⁶⁵ See Case C-67/13 P *CB v Commission* [2014] ECLI:EU:C:2014:2204, para 89. See also paras 42-46, and 90-92.

⁴⁶⁶ See Case C-434/13 P *Commission v Parker Hannifin Manufacturing and Parker-Hannifin* [2014] ECLI:EU:C:2014:2456, paras 74-86.

the *KME-Chalkor* cases and several subsequent cases reaffirming and extending that ruling; and second, the *Gascogne* cases and the subsequent cases confirming the line of argument established there. Both case clusters provide an interesting example for the purposes of balancing accuracy and efficiency in competition proceedings. The cases both demonstrate how the Court responds to situations where the law, as previously established, does not provide an obvious solution to the question at hand. The Court uses different methodologies to solve these issues, but in both instances, it effectively amends the law from its previous existing form by dictating a new balance between the primary competing elements in any law procedure; the element of procedural accuracy, and the element of procedural efficiency. Due to the law-making function, which the Court employs in these cases, implicitly in *KME-Chalkor* and explicitly in *Gascogne*, they provide an interesting example of how a *de facto* lawmaker approaches the preferences balancing in the composition of an actual accuracy and efficiency equilibrium.

In the *KME-Chalkor* cases the applicants argued that the limits on judicial review in competition proceedings, exemplified by the Commission's discretion to establish and evaluate certain types of facts, was not compatible with the strict accuracy standards for procedures of such type articulated in Article 6 of the ECHR and Article 47 of the Charter of Fundamental Rights. Although not argued by any of the parties, the Commission's discretion could have been rationalised with reference to the cost and accuracy implications entailed in allowing complex economic facts to be challenged before courts that do not possess expert personnel in economics. The Court's eventual finding effectively reduced the Commission's discretion, allowing parties subject to a cartel fine to challenge the Commission's findings on more factual grounds than before. The Court was silent as to the logic, or the theory, behind its decision to limit the Commission's discretion. Reading between the lines, we can assume that the Court considered the current limitations to judicial review unacceptable with reference to the obligations under Article 47 of the Charter. By taking that view, the Court was also stating that the standard of procedural accuracy would be inadequate if the Commission's discretion with regards to the assessment of complex economic facts were upheld.

By opening more grounds to appeal competition decisions of the Commission, and in particular complex economic grounds that the courts could previously defer on the basis of the discretion doctrine, the Court of Justice was in fact deciding that the jurisdiction of the courts were to be increased in comparison with the previous practice. A greater jurisdiction entails probabilities of increasing the volume and the complexity of litigation, which in turn requires the channelling of additional resources into taking and defending court actions, and to adjudicate an accurate result. The Court's decision in *KME-Chalkor* is thus not the least a suggestion that further resources should be used in competition enforcement, not just by those willing to pay more like those with disproportionately high stakes in the result (e.g. the addressees of a prohibition decision of the Commission), but also the EU society at large through the funding of the Commission and the Luxembourg courts. It seems evident that in the absence of increased public resources, a greater jurisdiction of the Courts will enable private parties to increase litigation resources and thus exploit the high standard of proof in their favour, to the effect that the risk of false negative errors, i.e. false annulments of prohibition decisions, will increase. The question that this raises is whether the Court was justified in taking such a decision, and based on which analysis was it reached?

The issue in the *Gascogne* cases was focused on the applicant's entitlement to a certain level of procedural efficiency, rather than on their entitlement to a certain level of procedural accuracy. While it was easy to establish that the procedure before the General Court had not complied with the standard of swiftness required by Article 47 of the Charter of Fundamental Rights, the Court struggled with establishing how to best remedy such problems. Again, the applicants wanted an efficient remedy, preferably an immediate reduction of the fine imposed, and the General Court had agreed in the judgments under appeal.

Advocate General Sharpston opposed arguments based entirely on considerations of procedural economy: *'The right to a fair trial within a reasonable time is a right that encompasses two key components, not one. Cutting every last possible corner in a search for swifter case-handling would not be compatible with maintaining the overall fairness of the*

proceedings.⁴⁶⁷ In this, she acknowledges the problem of balancing a need for economic efficiency, and a need for accurate results from law enforcement procedures, in an equilibrium of overall fairness of the procedure. Expanding on this point, she continued: *‘To the extent that the available resources of the General Court are inadequate to deal appropriately with the present and reasonably to be expected future case load, the responsibility must lie with the Member States.’*⁴⁶⁸ By this, she seems to be taking the view that courts should not be overly concerned about the economic implications of their decisions; they should primarily focus on producing accurate results, while deferring the issue of enforcement resources to the Member States.

By dismissing the economic rationale, AG Sharpston could conclude that the *Baustahlgewebe* approach is inferior to the *Der Grüne Punkt* approach, due to reasons of potential inaccuracy: *‘I understand that it is tempting to opt for the ‘procedural economy’ of reducing the fine (by some unspecified percentage, calculated on some unknown basis) in the context of an appeal. I am not satisfied by the intellectual foundations of such an approach (relationship between fine and conduct; jurisdiction; transparency) and it has, at worst, the potential to become almost entirely arbitrary.’*⁴⁶⁹

In its judgment, the Court of Justice substantively reached the same conclusion as suggested by AG Sharpston, but based on a different argument. Instead of dismissing economic efficiency considerations, the Court embraced them by refereeing to the efficiency advantage of having a uniform system for compensating for procedural delays: *‘Admittedly, the present case concerns a similar situation to that giving rise to the judgment in Baustahlgewebe v Commission. However, a claim for damages brought against the European Union pursuant to Article 268 TFEU and the second paragraph of Article 340 TFEU constitutes an effective remedy of general application for asserting and penalising such a breach, since such a claim can cover*

⁴⁶⁷ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:360, opinion of AG Sharpston, para 87.

⁴⁶⁸ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:360, opinion of AG Sharpston, para 88.

⁴⁶⁹ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:360, opinion of AG Sharpston, para 132. On this point see also paras 127, 129 and 131.

*all the situations where a reasonable period of time has been exceeded in proceedings.*⁴⁷⁰ Not convinced by this economic reasoning, AG Wathelet persisted to no avail in his subsequent opinion in the *Guardian Industries* case by restating the economic argument that had lost in *Gascogne*: *'Indeed, it would be paradoxical if the only way to obtain redress for excessively lengthy legal proceedings were to bring another legal action, which would necessarily entail additional costs (both for the parties and for the company) and further delay.'*⁴⁷¹

The opinions of the advocates general, and the judgment of the Court of Justice display how methodologically contentious the equilibrium between procedural accuracy and procedural efficiency can be. AG Sharpston wants to emphasise accuracy, but AG Wathelet economic efficiency. The Court gave priority to efficiency, but in a macro procedural economic sense that lead him to the same conclusion as AG Sharpston who had argued for accuracy, and to the opposite conclusion of AG Wathelet who subsequently argued for procedural economy in a more specific micro economic sense. In this, we are presented with three categorically different methodologies for solving an apparently simple problem of picking an optimal procedure for compensation for procedural delays.

For academic purpose, it is interesting to note the differing lines of arguments that are revealed in the opinions of the advocates general and in the judgment. They suggest a complete lack of consensus about the proper grounds for arguing about how a procedural rule should be designed. It is also noteworthy that the Court of Justice asked for, and presumably analysed the preferences of the Member States and several other stakeholders towards the two competing procedural alternatives. That suggests that the Court's ultimate judgment might be sensitive towards preferences of stakeholders, as is normal with predominantly *ex-ante* political decisions, but which is usually not a permissible argument in support of *ex-post* adjudicative decisions. Another significant revelation with regards to the judgments in the *KME-Chalkor* case line and in the *Gascogne* case line, is how underdeveloped the Court's theory of procedural rights is. In the former case, the equilibrium of accuracy and economic efficiency was nudged towards accuracy without proper motivation. In the latter case, the

⁴⁷⁰ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:770, para 82.

⁴⁷¹ See Case C-580/12 P *Guardian Industries and Guardian Europe v Commission* [2014] ECLI:EU:C:2014:272, opinion of AG Wathelet, para 111. See also para 110.

more accurate route of two choices was also chosen, but now on the grounds of an underdeveloped economic argument that contradicts what seems more economic and efficient in the immediate sense.

These two case clusters provide a good example of how an actual legal decision maker has approached two different issues of procedural fairness in which the law as it stood did not provide a definite *ex-ante* type of an answer to the dilemma. In the absence of an *ex-ante* answer, the possibility of providing an *ex-post* answer based on a previous *ex-ante* legislative decision is removed. The decision maker is thus forced into the *ex-ante* mode of deciding, which abides to a different rationale than the *ex-post* mode of deciding. Facing an *ex-ante* type of a decision, the Court showed signs of inconsistency in its methodology for approaching such questions, which resulted in a *de facto* rulemaking that was poorly motivated in terms of rights theory and in terms of the theory of economics. The opinions of the advocates general showed similar symptoms of confusion about the methodology that ought to be used for determining the content of a procedural rule in the context of competition proceedings, where the existing law does not provide a definite *ex-ante* type of an answer. In the next section, I will display how the *model of fair rules* could be applied to orderly solve the two dilemmas, in a consistent and coherent way based on an *ex-ante* decision methodology.

2. Solving fairness issues by applying the model of fair rules

So far, we have established that the mutually agreed purpose of EU's procedural regime in competition law is the accurate and efficient facilitation EU's competition law regime. We have also established that the normative aim of EU's competition law regime is consumer welfare on the internal market, understood as total economic welfare, restricted by a Pareto distributional requirement with regards to consumers. The work we would thus want the *model of fair rules* to undertake is to identify procedural alternatives that achieve the optimally fair balance of accuracy and efficiency in facilitating the normative aim of EU's competition law regime.⁴⁷²

⁴⁷² The model of fair rules in theory could also be used to calibrate the optimally fair aim of the competition law regime, but for our purposes, we assume that issue to be already settled at a constant that the procedural regime subsequently needs to facilitate.

The problem of fair balancing arises where decisions have the potential to inflict asymmetrical consequences on the stakeholders involved. Ideally, a procedure should not give rise to differences, once the stakeholders involved have agreed on the objective pursued through the procedure. However, the practical reality of limited resources imposes the acceptance of procedural errors. The concession of errors in a procedure, gives rise to two types of balancing problems: firstly, the total level of error concession; and secondly, the composition of the error types conceded. Thus, even if the stakeholders have an agreement on the substance pursued through the procedure, a disagreement can remain on how to deal with errors. At the identified decision points in EU's competition law procedure, a decision is taken that affects the error equilibrium between the stakeholders involved.

What I am thus suggesting, is that the stakeholders in an EU competition law procedure can disagree over two balancing questions at each of the major decision points in the procedure. They can disagree over how much should be spent on achieving the correct result, and they can disagree over the probabilistic distribution of the error cost that must be conceded.

Differences with regards to the first question can be explained based on the different effects a public investment in a procedure will have on the different types of stakeholders. If we simply assume that each stakeholder is equally sensitive towards a general increase in the tax burden due to an additional public investment in enforcement of the competition law regime, it might still be rational for the different stakeholders to prefer different levels of total investment due to their differing circumstances and thus differing stakes. The non-interested taxpayer will prefer an investment that is cost beneficial in the broader macro-economic sense, the consumer will prefer an investment that is cost beneficial measured by the benchmark of consumer welfare, the competitor will prefer an investment that is cost beneficial in terms of maintaining a level playing field in the marketplace, and the accused will prefer an investment that is cost beneficial towards avoiding unwarranted enforcement actions. Due to the differing stakes, the marginal utility of an additional investment in procedural accuracy is liable to be different between the stakeholders. In the same way, the breakeven cost point is likely to be located at different investment levels, depending on how important the consideration of accuracy is for the different stakeholders involved. All other things being equal, it would thus

be rational for the different stakeholders to prefer different levels of public investment in the enforcement of competition law.

Differences regarding the second question can be explained based on the same rationale. Different stakeholders have potentially varying preferences towards the concession of different types of errors. The concession of an error has general welfare implications, but in some cases also special implications that compound the general implications. Enforcement errors in competition law thus have general welfare implications that effect all non-interested taxpayers and interested consumers, but additionally the accused and the competitors are liable to suffer special consequences due to their special stakes. The special consequences will however vary between the accused and the competitor depending on the type of the error. The accused will prefer a false negative error, while the competitor will prefer a false positive error, assuming both are rational.

The role of the regulator, in a situation where different stakeholders have different preference trends towards the solution of a balancing dilemma, is to find an equitable solution in which the preference of each stakeholder is fairly reflected. To find the point of a fair equilibrium solution we can employ the *model of fair rules*. For the purposes of locating the fair equilibrium points, with regards to the two questions of procedural accuracy in the context of EU's competition law procedure, the model could be applied at the final decision point, where all the procedural steps and all the procedural possibilities aggregate in a grand balance of procedural accuracy and efficiency. The model could also be applied at each of the intermediary decision points, by aggregating the procedural steps and possibilities up to that point to identify the intermediate equilibrium of procedural fairness. In high profile competition law procedures, the balance at the intermediary decision points can matter a lot. It is for example of great importance for publicly listed companies, that a public accusation of an antitrust or cartel infringement is not made, unless a certain level of accuracy is guaranteed before it becomes permissible for public organs to intervene through such decisions. The consequences of an error at the intermediate decision points are of course not as severe, and thus the fair equilibrium point erects a lower accuracy threshold for making an accusation, than it would for finding an eventual infringement.

To apply the *model of fair rules*, an empirical input regarding the preferences of the relevant stakeholders is required. In the absence of an empirical research of such type, I will let do with explaining in few steps how the model could be applied and make few calculated guesses about the preferences of the stakeholders based on rational responds to the incentives that are obvious to any observer. To test the model, I will apply it on the two main problems that were at stake in the *KME-Chalkor* cases, and the *Gascogne* cases. By reading the cases and the accompanying opinions of the advocates general, some of the competing arguments are revealed and with them, the competing preferences for the ultimate equilibrium balance between the normative considerations of procedural accuracy and procedural efficiency.

2.1. Case Study I: The issue of fairness in KME-Chalkor

The issue of balancing in *KME-Chalkor* reflected an argument on behalf of the accused in the cartel proceedings of the Commission, to increase the procedural safeguards against a false finding of the Commission, especially in cases where a substantial fine had been imposed that amounted to a criminal sanction. The Commission did not engage with the argument in the abstract, but instead relied on the categorisation of the Commission's competition proceedings as administrative in Article 23(5) of Regulation 1/2003 EC rather than as criminal, and as a fall-back argument the ECoHR's distinction between hard-core and soft-core criminal proceedings, the latter of which did not require the same procedural safeguards at the first instance, given that at the later instance a full judicial review was available. Through the secondary argument, the Commission principally conceded the essential point made by the applicants, which had claimed that Commission decisions imposing substantial fines needed to be fully exposed to judicial review, and that the Commission could not in that regard be granted with any deference for appreciating complex economic facts.

Given that a concurrent view could be reached on the legal aspect of the problem, the disagreement moved onto factual issues with the Commission's alleged decisional discretion. The applicants maintained that the General Court had in its judgement explicitly showed deference to the Commission's finding. The Commission however maintained on the contrary that the General Court had granted no such discretion and that the applicants pleadings had

simply been rejected through the General Court's execution of its unlimited jurisdiction to review the Commission's penalising decision. The Court of Justice eventually sided with the Commission, and thus further reduced the *Consten and Grundig* deference doctrine. By this, the Court tipped the balance of the equilibrium between procedural accuracy and procedural efficiency, towards more accuracy. Subsequently, the new equilibrium constitutes the judicially approved benchmark for procedural fairness with regards to this specific issue.

If we assume that the Commission was granted with deference in assessing certain facts prior to the finding in the *KME-Chalkor* cases, the elimination or the reduction of the deference and the subsequent increased exposure of the Commission's competition decisions to challenges before the Luxemburg courts, implies additional enforcement costs to succeed in increasing procedural accuracy. Firstly, the Commission now needs to argue its decisions more thoroughly to withstand challenges on points that were previously non-challengeable. Secondly, the range of issues that can be challenged has been increased, which probabilistically translates into more frequent challenges, and more extensive pleadings; both of which drain resources from the parties to the case and from the decisional mechanism as such. Thirdly, the discretion, which the Commission supposedly had related to a set of expertise that the courts did not possess in the past, namely proficiency in fact finding grounded in the theory of economics. The elimination of the discretion implies that the courts must acquire this kind of expertise to be able to review confidently all aspects of the Commission's findings with an output that is at least as accurate as that of the Commission.

It seems apparent that in the absence of an increase of resources towards each of these additional costs, the overall accuracy of the procedure has a potential to decrease, or at the minimum fail to increase. Failure on behalf of the Commission to properly state reasons for its findings can result in false annulments by the Courts. Increased caseload on an underfunded enforcement mechanism is liable to increase errors. Absence of expertise at the Courts is also likely to result in enforcement errors. The Court's rulings in the *KME-Chalkor* cases were thus only probable to increase the accuracy of the enforcement process, if an increase in enforcement resources were to follow. If we take the liberty of interpreting the intention of the Court, it seems as if it wanted to increase the accuracy of the enforcement process with

reference to the due process standards of the Charter of Fundamental Rights, which unavoidably imposed various costs on the system as a whole. If this was the Court's intention, it essentially decided to increase accuracy by requiring additional investments in the enforcement process. It can be debated whether the Court was the best placed decision maker to take such an investment decision, on behalf of the relevant taxpaying population.

The apparent investment decision in *KME-Chalkor* rested on an ambiguous distinction advocated by the Commission based on ECoHR's finding in the *Jussila v Finland* case.⁴⁷³ The Grand Chamber of the ECoHR had concluded that the difference between criminal and administrative proceedings was gradual, rather than absolute, and that in-between the two there existed a category of proceedings that contained elements of criminal proceedings but not in the traditional hard-core sense. A distinction should thus be made between the two when assessing the applicability of the procedural guarantees of Article 6 of the ECHR and the level of protection granted for each class of criminal procedures.⁴⁷⁴ The Court's ruling in *KME-Chalkor* thus moved the competition proceedings of the Commission from the pure administrative sphere, to the in-between sphere where some of the criminal procedural guarantees apply, but not with full stringency in terms of Article 6 of the ECHR and Article 47 of the Charter.

⁴⁷³ See Case no. 73053/01 *Jussila v Finland* [2006] ECLI:CE:ECHR:2006:1123JUD007305301, para 43.

⁴⁷⁴ The Grand Chamber of the ECoHR was not unanimous on this distinction in *Jussila v Finland*. Judge Loucaides wrote a dissenting opinion which was joined by judges Zupancic and Spielmann: '*I find it difficult, in the context of a fair trial, to distinguish, as the majority do in this case, between criminal offences belonging to the "hard core of criminal law" and others which fall outside that category. Where does one draw the line? In which category does one place those offences which on their face value do not appear severe, but if committed by a recidivist may lead to serious sanctions? I believe that the guarantees for a fair trial envisaged by Article 6 of the Convention apply to all criminal offences. Their application does not and cannot depend on whether the relevant offence is considered as being in "the hard core of the criminal law" or whether "it carries any significant stigma". For the persons concerned, whom this provision of the Convention seeks to protect, all cases have their importance. No person accused of any criminal offence should be deprived of the possibility of examining witnesses against him or of any other of the safeguards attached to an oral hearing. Moreover, to accept such distinctions would open the way to abuse and arbitrariness. I firmly believe that judicial proceedings for the application of criminal law, in respect of any offence, by the omnipotent State against individuals require, more than any other judicial proceedings, strict compliance with the requirements of Article 6 of the Convention so as to protect the accused "against the administration of justice in secret with no public scrutiny". As rightly pointed out by Trechsel "... the principle of public trial in criminal cases has an importance which goes beyond personal interests". Therefore, once it was found (correctly) that the relevant proceedings in this case were criminal, the requirement of a public hearing in respect of them became a sine qua non. The failure to fulfil that requirement amounts, in my opinion, to a breach of Article 6 of the Convention.*'

Interestingly the argument made by the ECoHR rejects a strictly deontological interpretation of the conditions of Article 6, the Court hints at a more consequentialist approach by conceding that certain procedural guarantees are not absolute conditions, but rather indicators that should be interpreted in context with the protective interests at stake: *‘Notwithstanding the consideration that a certain gravity attaches to criminal proceedings, which are concerned with the allocation of criminal responsibility and the imposition of a punitive and deterrent sanction, it is self-evident that there are criminal cases which do not carry any significant degree of stigma. There are clearly “criminal charges” of differing weight.’*⁴⁷⁵ By taking this step the Court risks the cohesion of the system of protection under Article 6 ECHR, as was pointed out by judge Loucaides in his dissenting opinion: *‘[T]o accept such distinctions would open the way to abuse and arbitrariness.’* A consequentialist finding, in a system that relies on a deontological mode of arguing, invites confusion about the permitted types of arguments. Should the rights under the ECHR be argued for based on their deontological character, or can rights arguments be won or lost through a cost benefit analysis? *Jussila v Finland* seems to suggest the later and the *KME-Chalkor* by analogy relies on such rationale.

Despite the consequentialist character of the arguments for the main finding in *Jussila v Finland*, the typical tool for solving such balancing dilemmas in rights jurisprudence, the proportionality test was not employed, at least not explicitly. When the Court of Justice in *KME-Chalkor* subsequently adopted the argument, the Court not only withheld from making an explicit reference to the underlying rationale of the *Jussila* finding, but also abstained from commenting on the nature of the distinction between a criminal and an administrative proceeding, which was the crucial substantive finding in *Jussila*. Consequently, the sanctioning of the enforcement system in competition proceedings concerning compatibility with Article 47 of the Charter was insufficiently motivated, especially when considering that the Court had to alter the system in place slightly to be able to fit it within the parameters of the criteria of procedural fairness as decreed by the Charter and the ECHR. To be able to motivate its finding, the Court would have had to resort to a balancing exercise based on a consequentialist rationale. Without such an argument, the result lacks a rational foundation.

⁴⁷⁵ See Case no. 73053/01 *Jussila v Finland* [2006] ECLI:CE:ECHR:2006:1123JUD007305301, para 43.

In this context, it is also important to note that the Court of Justice refrained from referring to Article 6 of the ECHR in its finding in *KME-Chalkor*, but instead referred exclusively to Article 47 of the Charter.⁴⁷⁶ This created an additional obstacle; it barred the Court from using the interpretation of the ECoHR as a deontological anchor for its own finding. If the Court does not recognise the deontological force of the provisions of the ECHR and the interpretations of the ECoHR, it becomes problematic for the Court of Justice to use these same legal sources as anchors for its own findings without repeating in full the rationale on which the sources rest. The Court of Justice did not support its finding in the *KME-Chalkor* judgements with elaborate arguments explaining the deontological character of the procedural guarantees of Article 47 of the Charter, nor did it refer to the jurisprudence under the ECHR.⁴⁷⁷

In the absence of both a deontological argument and a consequential argument, it is challenging to identify the nature of the Court's reasoning in *KME-Chalkor*. It is hard to know whether the judgment was the result of a careful quantitative balancing between the competing considerations of procedural accuracy and procedural efficiency, or whether it was the Court's intuitive interpretation of the rights imperatives, integral to Article 47 of the Charter. Probably, it was a bit of both. The Court was probably under the influence of the *Jussila* ruling, which recognised the gradual nature of the right to a fair proceeding, but the Court did not explicitly employ any quantitative tools to establish where the balance between the competing interests should be located. Instead, the Court seems to have used its intuition to conclude that a fair equilibrium is reached, if the traditional criminal procedural guarantees partially apply to competition proceedings.

Given a consequential rationale along the lines of *Jussila*, the essential balancing issue in the *KME-Chalkor* litigation can be extracted through a simple question; how strict procedural

⁴⁷⁶ Sibony discusses the reasons for this in her note on the case, and comments that the Court avoided the question of categorisation as criminal or not by refereeing exclusively to Article 47 of the Charter. See Anne-Lise Sibony, 'Annotation of the judgment of the Court in Case C-272/09 P *KME Germany and others v. Commission*' (2012) 49 *Common Market Law Review* 1977, 1989-95.

⁴⁷⁷ The CJEU later altered its approach in this regard and made an explicit reference to Article 6 of the ECHR and the ECoHR's Menarini judgment. See Case C-501/11 P *Schindler Holding and Others v Commission* [2013] ECLI:EU:C:2013:522, paras 35-36.

guarantees should be employed to competition proceedings, given the protective interest at stake?

As noted in chapter seven, the essential balancing question for the competition proceedings of the EU can be assessed at various balancing points in the procedure. The point in question in the *KME-Chalkor* cases related to the judicial review of the General Court of the Commission's decision to fine certain participants in a cartel breach to the competition provisions of the TFEU. Referring to the previous discussion on the various balancing points in EU competition proceedings, the point at issue here follows the three decision points in the Commission's administrative procedure, but precedes the ultimate decision of the Court of Justice of the EU. The focus is thus at the fourth decision point and the question is how rigid it should be in terms of decisional accuracy.

2.2. Case Study II: The issue of fairness in Groupe Gascogne

The main fairness issue in the *Groupe Gascogne* cases was more explicitly argued in terms of the accuracy/efficiency dichotomy by the parties to the cases, than it was in the *KME-Chalkor* cases. There was little doubt about the existence of undue delays in the appeal procedure before the General Court, but the issue of debate was how to deal with such claims and how to compensate for such procedural faults. On one hand, it was argued that the Court of Justice should deal immediately with such claims and compensate for procedural faults by lowering the cartel fines that had been imposed in the cases. On the other hand, it was argued that the right approach was to initiate separate procedures for damages that would be dealt with by the General Court, which normally deals with questions of fact in competition cases, unlike the Court of Justice that primarily deals with questions of law.

The applicants, in line with their interests, argued for the simpler and immediate solution to this problem; they wanted the cartel fines to be reduced without having to launch a separate case before the General Court. The AGs were far from unanimous in their opinions on this issue; while AG Sharpston argued for a separate action for damages with reference to accuracy considerations, AG Wathelet argued for emphasis on procedural economy by lowering the fines immediately. The Court of Justice approached the problem more holistically and thus

found that a separate action needed to be taken by the applicants to conform with the general procedural route to claim compensation for procedural irregularities, implying that efficient exceptions from the norm might nonetheless create inefficiencies for the broader EU procedural system. The underlying rationale of the key arguments at odds in the *Groupe Gascogne* case line is partly reviled in the opinions of the AGs.

AG Sharpston at one point suggested that the *'right to a fair trial within a reasonable time is a right that encompasses two key components, not one'*. She then continued by contrasting these key components: *'Cutting every last possible corner in a search for swifter case-handling would not be compatible with maintaining the overall fairness of the proceeding.'*⁴⁷⁸ Sharpston identifies swiftness as being at odds with the concept of a fair proceeding, but it is not entirely clear which meaning she attaches to the normative objective of fairness. Her argument could be interpreted to mean that swiftness risks procedural errors that in turn undermine the fairness of the procedure; i.e. that fairness represents the accuracy of the procedure's result. This interpretation about the severability of procedural swiftness and procedural fairness is reinforced by Sharpston's remarks about the responsibility of the Member States to provide the courts with sufficient resources to enable swift case handling, and that *'difficulties caused by case overload, however real they may be, should be disregarded when assessing whether there has been excessive delay in handling a particular case.'*⁴⁷⁹ By this she seems to suggest that swift case handling is an entirely independent consideration that will not be excused with reference to limited availability of enforcement resources.

By treating procedural swiftness and procedural fairness as separate imperatives of the deontological type, both can be realised independently at the same time if, and only if, the issue of resources is ignored. When forced to choose between an economic procedural route and a more accuracy oriented procedural route, Sharpston puts the priority on accuracy: *'I understand that it is tempting to opt for the 'procedural economy' of reducing the fine (...). I am not satisfied by the intellectual foundations of such approach (...) and it has, at worst, the*

⁴⁷⁸ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:360, opinion of AG Sharpston, para 87.

⁴⁷⁹ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:360, opinion of AG Sharpston, para 90.

*potential to become almost entirely arbitrary.*⁴⁸⁰ Swiftiness should thus be secondary when at odds with procedural accuracy, or as already noted by Sharpston in the cited passage above; reduction of accuracy for the purposes of swift case handling is not compatible with procedural fairness.

AG Wathelet's remarks in *Guardian Industries v Commission* indicate a different understanding of the relation between procedural fairness and the expediency of the enforcement process. Wathelet cites the '*economy of procedure*' as a decisive reason for solving the problem of how to provide redress for undue procedural delays at the General Court. He further explains that '*it would be paradoxical if the only way to obtain redress for excessively lengthy legal proceedings were to bring another legal action, which would necessarily entail additional costs (both for the parties and for the company) and further delay.*'⁴⁸¹ Although not explicitly argued, this kind of reasoning presupposes that the expediency of a procedure is a component of a fair procedure. Undue delays are thus liable to impair the fairness of the procedure. On this view, further delays to the procedure to obtain redress are unlikely to provide an efficient way to rectify a breach to the principle of procedural fairness, caused by a procedural delay. Wathelet's approach, views case handling swiftiness as an efficiency function integral to procedural fairness. Procedural delays produce costs for the parties, which in turn are unfair in the procedural sense, if not rationalised based on other elements or components of procedural fairness, such as accuracy considerations. In contrast to AG Sharpston's deontological approach to procedural fairness, AG Wathelet's approach is more consequential in character by approaching the question of procedural fairness as a cost-benefit question that is sensitive to the issue of efficient use of public resources.

The Court of Justice, in its reasoning for the result, did not assess the costs or benefits of lowering cartel fines immediately in respond to undue delays in the procedure. The Court simply stated that the procedure provided for in Articles 340 and 268 TFEU provided a sufficiently effective remedy of general application to address the problem of procedural

⁴⁸⁰ See Case C-58/12 P *Groupe Gascogne v Commission* [2013] ECLI:EU:C:2013:360, opinion of AG Sharpston, para 132.

⁴⁸¹ See Case C-580/12 P *Guardian Industries and Guardian Europe v Commission* [2014] ECLI:EU:C:2014:272, opinion of AG Wathelet, para 111. See also para 110.

delays in cartel cases before the General Court. The Court nonetheless proceeded with establishing an undue procedural delay and thus a breach of Article 47 of the Charter, but abstained from quantifying the harm suffered by the applicants. The argument of the Court seems to weigh the benefit of having a single uniform procedure for claiming damages in the EU system of laws, against the cost of having specific procedures for specific types of damages claims. The argument is consequential rather than deontological, but instead of making the assessment of the costs and benefits of a single instance decisive, the Court subtly concludes that the general procedure '*constitutes an effective remedy*', implying that it is effective enough for this instance, but without prejudice to how effective it could be using the competing approach. By shifting the weight of the consequential reasoning away from the specific, over to the general, the Court can reach a conclusion that implies that the inconvenience of starting a new damages procedure in this case, should be outweighed by the convenience of having a single uniform universally applicable procedure for claiming damages caused by the actions of the EU institutions.

Despite the apparent consequential character of the reasoning of the Court, there are not indications in the text of the judgement that any actual quantitative assessment was made of the advantage of having a uniform system for claiming damages, or whether there actually is any advantage to emphasise uniformity over a specifically applicable efficient procedural solution, such as the one argued for by the applicants and AG Wathelet in the context of the *Groupe Gascogne* cases and the *Guardian Industries* case.

Different approaches to the procedural problem in the *Groupe Gascogne* cases reveal different methodologies to solving a problem that was predominantly legislative in character. AG Sharpston advocates a deontological methodology where the accuracy consideration is above other normative considerations, but she fails to consider the inescapable reality of definite enforcement resources. AG Wathelet acknowledges the importance of procedural efficiency for the concept of procedural fairness. The Court also acknowledges the role of procedural efficiency, but loses sight of the context of its assigned decisional task by adopting an external perspective that is irrelevant to the problem of competition law procedure that was at stake in the cases.

Essentially, the problem in *Groupe Gascogne* concerned the fairness of the cartel procedure at hand, and the fairness of the consequences for breaching the standard of procedural fairness in the cartel case. As we have previously established, the task of the competition procedural regime is to realise the normative objective of the competition law regime, and do so based on a fair balance between an accurate resolution of the task and an efficient use of the available enforcement resources. The fairness problem in question here is a derivative problem from the competition law procedure, and must as such be guided by the same essential agenda of realising the normative objective of the competition law regime. The question that was thus at stake in *Groupe Gascogne* was how the competition law procedure could be optimised in terms of fairness; would it be optimised by the immediate reduction of fines in cases of undue procedural delays, or would it be optimised by opening a new procedure to deal specifically with the problem of delay? Logically speaking, it would thus not be sufficient to show that the general procedure for claiming damages for procedural delays was adequate as the Court did, if it were not also adequate to deal with the problem of delays in a competition law procedure; i.e., the issue of fairness related to the competition law procedure at hand, and the obligation of providing fairness will only be discharged with regards to that same procedure. By its formulation of the problem, the Court thus only found fairness in the general sense, which does not necessarily imply fairness in the more specific sense of the competition procedure at stake.

By taking this view on the problem we can see that it evolves around the effects of each of the two procedural routes on the fairness of the competition procedure, and as we have seen the fairness of the procedural regime is determined by the balance of the accurate and efficient realisation of the normative objective of the competition law regime.

The balancing point affected by the problem in *Groupe Gascogne* relates to a delay in the court procedure before the General Court and how that delay should be alleviated to minimise the effects on the fairness of the whole procedure. The issue could only be brought up at the framing stage preceding the fifth and the final decision point, since the delay occurred after the conclusion of the framing stage before the fourth decision point before the General Court.

2.3. Applying the model of fair rules to KME-Chalkor and Groupe Gascogne

The central claim in *KME-Chalkor* was that accuracy of the competition procedure should be increased by abolishing or reducing the discretionary power of the Commission to establish certain complex economic facts at the end of the administrative procedure. This claim entails a presumption that the system in place was not accurate enough, and that the proposed change would increase the accuracy. The counter argument to this claim, although not raised during the proceeding, could focus on these presumptions; i.e. that the current system was accurate enough, and that the change did not appreciably increase accuracy. Additionally, it could be argued that the proposed accuracy enhancement was only cost beneficial for a narrow group of stakeholders, while for most of the others, the additional investment was not desirable given the level of the accuracy advantage gained.

The central claim in *Groupe Gascogne* was that the most efficient way to enhance and rectify the fairness of a competition procedure, that had been excessively delayed before the General Court, was to discount the cartel fine imposed immediately at the final instance before the Court of Justice. The counter argument, that prevailed, was that the accuracy of the competition procedure and the overall efficiency of the procedural system would be better served by using the universally applicable procedure for claiming damages caused by the EU institutions, instead of permitting the use of a special procedure in cases where a cartel fine had been imposed.

As previously noted, the problems in these two cases were undecided from a legislative point of view, which barred the Court from deciding based on an *ex-ante* answer previously provided by a legislator and thus required the Court to provide its own *ex-ante* type of an answer for the case at hand and for future cases involving the same or similar problems. Based on this understanding we can proceed with applying the *model of fair rules*, which is designed to deal with legal decision-making of the *ex-ante* type.

2.3.1. Step 1: Identifying the legislative plan

The grand legislative plan, in the context of EU competition law, is currently settled at an equilibrium point that strives for the protection of consumer interest on the internal market through the promotion of welfare in the economic sense. Translated into traditional concepts of normative political and economic theory, the current political equilibrium on the normative objective of EU competition law can be said to aim at total welfare in the economic sense, restricted by a Pareto distribution rule with regards to consumers. The primary instrument, with which the EU's competition law regime seeks to pursue this normative objective, is through the instrument of competition in the marketplace, the details of which are further articulated in the substantive provisions of the competition law regime.

Supplementary to the substantive provisions of EU's competition law regime are the procedural provisions that are described in various EU legal instruments and in the practice of DG Competition, the Commission, and the Court of Justice of the EU. These provisions and practices serve the same auxiliary agenda as any other procedural rules; to facilitate with accuracy and efficiency the grand agenda of the substantive regime, to which they provide their auxiliary service. The legislative plan of the procedural regime is thus different from that of the substantive regime. The plan of the procedural regime refers to an equilibrium point between the normative considerations of accuracy and efficiency, while the plan of the substantive regime is currently fixed at an equilibrium point between a utilitarian efficiency principle and a political distribution principle.

The object of preference for the stakeholders in the context of EU's competition proceedings is thus the balance between accuracy and efficiency. As previously explained, this balance can be influenced externally by the increase or decrease of the total available resources, or it can be influenced internally through the probabilistic distribution of the types of errors conceded. In both cases, the exact equilibrium point between the two variables becomes the object of preference for the stakeholders involved. If the regulator manages to find an equitable solution to the balance between efficiency and accuracy in a procedure, based on the preferences of the stakeholders involved, the equilibrium point represents procedural fairness. A claim about procedural fairness thus relates to the balance between accuracy and efficiency of a

procedure. In practice, the claim is often that more should be spent on accuracy, or that the risk of committing a certain type of accuracy errors should be reduced. Sometimes the practical claim is also that the procedural process should be more efficient, either through higher investment of public resources, or through the relaxation of the accuracy demand.

If we view the central claim in *KME-Chalkor* as a legislative proposal that the Court in its capacity as a decision-maker had to decide on whether to respond to, the plan behind the proposal is easily identified as towards increased accuracy of the competition procedure at the fourth decision point before the General Court. By expanding the jurisdiction of the General Court to scrutinise the decisions of the Commission, the total accuracy of the procedure is liable to increase if additional resources are committed to match the increase in workload. If, however the jurisdiction is expanded without committing additional resources, the rules of proof will require that the existing resources will be prioritised towards reducing false conviction errors, which will leave fewer resources towards scrutinising acquittal judgments for errors. In summary, the legislative plan in this case relates to the balance between procedural accuracy and procedural efficiency at the fourth decision point in EU's competition procedure.

If we view the competing claims in *Groupe Gascogne* in the same way, as proposals to adopt a rule on how to compensate for undue procedural delays before the General Court, the plan behind the proposals can quickly be revealed as: on one hand being towards increased efficiency of the procedure, and on the other hand towards increasing accuracy. By adopting a simple procedure for claiming damages due to procedural delays, total efficiency of the procedural process is enhanced and the incentive for the institutions to handle cases timely and efficiently is reinforced. The incentive that is reinforced, is however probable to undermine the accuracy of the procedure, unless additional resources are committed towards enhancing the case handling time. The other claim is based on accuracy considerations; mixing the compensation procedure with the cartel procedure, risks the accuracy of both. In summary, the legislative plan, in this case, relates to the balance between procedural accuracy and procedural efficiency at the fifth decision point in EU's competition procedure.

2.3.2. Step 2: Preference function of the stakeholders

Presuming that a given equilibrium point of accuracy and efficiency in competition proceedings will have different effects on people depending on their prior circumstances, it is logical that the preferences of people will differ with regards to where the equilibrium point should be located on the axis between efficiency and accuracy depending on those prior circumstances. It should thus be possible to categorise different stakeholders in a competition proceeding based on their role in the proceeding, which is determined by their prior circumstances.

Regarding EU's competition law proceedings, we have already defined four types of stakeholders: the accused, the competitor, the consumer, and the taxpayer. Each of these stakeholders is situated differently, which gives them incentive to hold diverging preferences with regards to the question, where the optimal equilibrium point of procedural accuracy and procedural efficiency should be located. The accused and the competitor would prefer large public investment in accuracy, but they would hold diverging preferences regarding the types of errors that should be the object of emphasis within the confines of the currently available resources. The consumer would prefer public investment in more accuracy if his share of the cost would not exceed his expected increase in welfare and he would be neutral with regards to the allocation of the conceded error types. The general non-interested taxpayer would support investment in procedural accuracy if it remains cost beneficial in the macro economic sense, which means that the investment should yield welfare returns that exceed the returns from other investment alternatives. The taxpayer is like the consumer not sensitive towards the typological allocation of conceded errors.

The preference function of the stakeholders towards the two reform claims can be assumed, based on the effects the proposed changes are liable to have on the stakeholders and based on the effects on the current regime. The effects of changing the competition procedure along the lines of the main claim in *KME-Chalkor* can be roughly assumed to be as follows: The accused would prefer to reduce the discretion of the Commission to establish complex facts, since he would benefit from increased accuracy of the procedure, or in the case of no additional enforcement resources to increase the emphasis on reducing false positive errors. The competitor would however be neutral towards reducing the discretion of the Commission

because only finding of a breach would be scrutinised, not the Commission's finding of no breach, which the competitor has an active interest in having as accurate as possible. If no additional resources would be committed alongside the change, the competitor would oppose it due to increased risk of false negative errors. The consumer would favour increased accuracy of the procedure so long as it could be shown that his cost of doing so would yield a net increase in his own welfare. He would also only favour more emphasis on reducing false positives if it could be shown to increase his welfare beyond the status quo. The taxpayer would support additional investment if it were to increase welfare more than competing investment options, and he would also only support increased emphasis of reducing false positives if it were to increase the taxpayer's welfare beyond the status quo.

The effects of adopting the procedures argued for in *Groupe Gascogne* involves comparing two alternative options, rather than comparing a reform claim with the status quo. The main problem of these cases involved ambiguity about the actual content of the status quo. In such cases, we simply assess the preferableness of the alternative options to each of the relevant stakeholders. The accused in the main competition proceeding would, as is indicated by the pleadings of the applicants, prefer the simple procedure of immediately lowering the ultimate cartel fine to compensate for undue procedural delays on behalf of the General Court. The channelling of public resources to a new damages procedure would leave less to get things right in the original cartel procedure and would additionally impute significant additional direct and indirect costs on the accused. Even in the case of additional public investment, it is probable that the gain by the accused of having a separate damages procedure would be eclipsed by the additional costs. The competitor would on the contrary prefer that the damages claim would not interfere with the cartel proceedings and would thus prefer it to be dealt with separately through a procedure that would be relatively accurate, given that additional public resources would be committed towards the task. In the case of no such investment commitment, the competitor would have to weigh the harm to accuracy on the cartel proceeding of including the damages claim, with the harm to accuracy of channelling some resources towards initiating a separate procedure. The consumer would choose the cheaper of the two procedural alternatives, unless the accuracy gains from choosing the expensive one would increase his own welfare. The same logic would apply if there were no additional public

investment involved. The taxpayer would in the same way chose the cheaper option, unless an investment in the expensive option could be rationalised as a beneficial investment compared with alternative investment options.

2.3.3. Step 3: Preference index for antitrust procedure

Viewed objectively, the relative importance of a stakeholder individual preference profile can vary. By establishing a preference index, an attempt is made to compare the relative importance of what each stakeholder wants. The usual assumptions about rationality apply. Not attempting to be surgically accurate, the tool is meant to provide the legislator with a useful tool for comparing the relative importance of different lobbying arguments, which is methodologically more robust than instantly applied intuition or a reduction to welfare inputs digestible for a cost benefit analysis.⁴⁸²

The first step in building a *preference index* is to define magnitudes of the different preferences at stake. A simple way to do that is to define the minimum and maximum magnitudes of the different preference functions; i.e. if this stakeholder gets all of what he wants the consequence will be this, and if he gets nothing of what he wants the consequence will be that. This provides two numerical values that represent specific consequences at the opposite extremes. By doing this with two competing stakeholder preference profiles (or more) we have a pair (or pairs) of corresponding magnitudes, i.e. we know the consequence of a 0% and a 100% preference fulfilment for the different preference profiles. This information enables an objective comparison of the different profiles. We simply assess the relative importance of a specific level (e.g. 0% or 100%, or any other comparable magnitude) of preference fulfilment for two or more stakeholders with a diverging preference function. The perspective of assessment must be neutral with regards to the stakeholders involved; their subjective interest must not influence how their preferences are objectively viewed. We want to establish how important the claim of a stakeholder is compared with the claim of another, as observed by a neutral bystander.

⁴⁸² As previously mentioned in chapter four, the Preference Index is inspired by the utility inference method of Van Neumann and Morgenstern, Savage's extension of the Bayesian probability model, the common priori assumption of Harsanyi, and the reference point dependence insight articulated by Kahneman and Tversky.

Previously I have suggested two variants of the *preference index*: a simple one that modestly compares comparable magnitudes of different preferences and determines their relative importance through Van Neumann's and Morgenstern's utility inference method; and a more complex one that utilises the status quo reference point to increase descriptive accuracy.

Using the balancing problems in *KME-Chalkor* and *Groupe Gascogne* as examples, we simply start by defining a best and a worst-case scenario for each of the stakeholders involved with regards to the balancing problems. We have already established that the balancing problems relate to the fourth and the fifth decision points in an EU competition proceeding. We would thus define the best and the worst outcomes in relation to these decision points; the worst and the best outcome for each stakeholder with regards to decision point four in *KME-Chalkor*, and the fifth point in *Groupe Gascogne*.

Once the best and the worst case scenarios have been defined we could choose to assemble a simplified version of the *preference index* by simply inferring objective utility values in a numerical form on the corresponding magnitudes and thus reveal their relative importance.⁴⁸³ If we choose the more complex version, we adopt a more case specific point of view by defining the current status quo in terms of preferableness for the stakeholders involved and locate the status quo on the axis between the worst and the best-case scenarios.⁴⁸⁴ Irrespective of which variant is used, we need, for simplifying purposes, to assume that the best possible consequence is at a point where the preferences of all stakeholders are completely fulfilled, and the worst possible consequence is where the preferences of all the stakeholders are completely ignored. All possible combinations of preference fulfilment of the stakeholders involved should thus fall on an axis between these two extremes, including the current status quo. Assuming this, all that needs to be done is to compare the impact of a change to one stakeholder type on the axis of overall preferableness to the impact of a change to another stakeholder type. By comparing the impact of comparable magnitudes of different types of stakeholders on the overall preferableness of a specific state of the world, an insight is gained

⁴⁸³ The question could simply be; 'how important is it to fulfil all of this stakeholder's desires, compared with fulfilling all of another stakeholder's desires?'

⁴⁸⁴ The inclusion of the reference points adds a dimension into the comparison. Instead of comparing absolute levels of preference fulfilment at a single measuring point, the comparison considers the preferableness of changes to the current level of preference fulfilment.

on the objective importance of the preferences of different stakeholders in that world. The composition of the preference index can be summarised as follows:

Step-by-step guide to the composition of a Preference Index

Step #	Action	Result
Step 1	Define best and worst-case scenarios for each stakeholder with regards to the legislative agenda at stake (e.g. the agenda at the fourth and the fifth decision point in EU competition procedure)	100% and 0% magnitudes of preference fulfilment established for each stakeholder with regards to the legislative agenda at stake
Step 2	Identify the status quo for each stakeholder between the extremes of the best and the worst-case scenarios (e.g. how preferable is the current procedure at the fourth and the fifth decision point in EU competition procedure?)	Information about the status quo preference level for each stakeholder
Step 3	Compare the total loss and the total gain of each stakeholder counting from the status quo as observed by a neutral bystander (e.g. how would a 100% gain and a 100% loss at the fourth decision point in EU competition procedure, objectively affect the different stakeholders?)	Information about the relative importance of each's stakeholders claim

Again, using the problem from the two competition litigations as examples, we would simply aggregate the worst-case scenarios of each of the stakeholders and the best-case scenarios and put at the opposite ends of an axis. Then we would use Van Neumann's and Morgenstern's utility inference method to locate the status quo point between the extreme points, and following that we could record the impact of any conceivable change to the status quo as a change relative to the opposite extremes and relative to the status quo. Depending on how accurate we would want or need to be, we could compare how a certain percentage change (negative or positive) of the preference fulfilment of a stakeholder would influence the overall

scale of preferableness with the impact a corresponding change in the preference fulfilment of another stakeholder would have. The result from our examples could be, speaking hypothetically in the absence of actual numbers, that for each percentage point that the desire of the accused for increased accuracy of the proceeding is realised, the overall objective preferableness of the solution to the balancing problem increases at a 2 to 1 ratio as compared with a corresponding percentage point fulfilment of the desires of the consumer.⁴⁸⁵ It would thus be more efficient towards overall preferableness to fulfil the desires of the accused, with regards to this particular balancing problem, than it would be to fulfil the desires of the consumer. The *preference index* of the consumer thus weighs less than the *preference index* of the accused at a *preference ratio* of 2 to 1.

By establishing systematically, the relative importance of different lobbying arguments of stakeholders towards specific balancing problems, the legislator gains a tool that enables him to tune with more accuracy the optimal solution to regulatory problems, given that he seeks to optimise the preferences of his constituency through his legislative acts. The *preference index* for different stakeholder interests, gives an objective perspective on what matters in the balancing and which weight should be given to individual types of claims. The objective importance of individual types of claims, do however not necessarily reflect how the stakeholders subjectively feel about the actual claims being made. Although, objectively speaking, a specific level of importance is attached to certain types of preference claims, the intensity of the claims within the category can differ as seen from the subjective perspective of the stakeholders involved. An objectively important type of a claim at a low intensity can thus weigh less than an objectively unimportant type of a claim at a high intensity.

2.3.4. Step 4: Preference matrixes for the key stakeholders

A range of actions can be proposed to achieve a specific legislative plan. These actions can be received differently by the various stakeholders involved, depending on the exact consequences of the actions or, in the case of legislative proposals, their expected consequences. In the context of competition law procedure, the plan, as previously noted, is to identify the optimal balance between the accuracy and the economic efficiency with which

⁴⁸⁵ I use the ratio 2:1 just as an example; I do not have any data to extract the actual ration from.

the primary objective of the competition law regime is pursued. The stakeholders previously identified are liable to hold different preferences towards the optimal location of this balance due to their differing stakes. Numerous legislative actions can be envisioned as having effect on the balance between what the stakeholders want. The effects these actions have on the preference fulfilment of the individual stakeholder matters, if the legislator is seeking to find a palatable legislative solution with regards to the stakeholders involved. It is not sufficient to establish that an action would be preferred or disliked by a stakeholder; information is also needed on the level of intensity of that position.

By listing the proposed actions and the intensity with which individual types of stakeholders like or dislike the probable consequences of a proposed action in a matrix, it becomes easy to compare and process further the subjective value of an action, as seen from the perspective of each of the stakeholders. Using the two examples previously mentioned, a matrix would simply be assembled for each stakeholder in a competition proceeding and the intensity of the effects on the preferences function of each would be measured with regards to the proposed actions. In both instances, the actions concern the decision points at the end, or nearly at the end of a competition procedure. The status quo, from which the effects of the proposed changes would be measured, thus needs to reflect that perspective.

Before establishing a subjective value of a proposed action for each of the stakeholders, we would have to envision the expected consequences of the proposed action for the relevant interests at stake. The subjective value of an action for a stakeholder is determined by how preferable the consequences are for him, irrespective of the preferences of the other types of stakeholders. The balancing problem in competition procedure is to establish a fair equilibrium between the normative consideration of procedural accuracy and procedural efficiency. At an individual level, each of the four identified stakeholders are liable to hold different preferences towards where the equilibrium point should be located on an axis between accuracy and efficiency.

By surveying how preferable a proposed action is for each of the stakeholder types, we are surveying how close to the point of optimisation the action is, from the perspective of each

stakeholder type. If the probable consequences of the proposed action were close to the optimum for a stakeholder, the action would measure as preferable at a certain intensity in the preference matrix for that stakeholder. The probable consequences of the same action could then at the same time suggest an outcome far from the point of optimisation for a different stakeholder, and thus measure in his preference matrix as un-preferable with a certain intensity. As before, Van Neumann's and Morgenstern's utility inference method can be used to numerically record the intensity of the change of preferableness each proposed action is liable to have, measured from the status quo, for each stakeholder.

For the purposes of the balancing problems in *KME-Chalkor* and *Groupe Gascogne* we would simply need to envision the probable consequences of limiting the discretion of the Commission to establish complex economic facts, and the consequences of the two competing options for resolving damages claims for procedural delays in cartel proceedings, and survey among the relevant stakeholders how preferable the different actions are, measured from the current status quo. The result of the survey would then be recorded in a numerical form into a preference matrix for each of the stakeholders, with a special provision for each of the proposed action. The recorded data would represent the subjective value of each of the proposed action, with regards to each of the relevant stakeholder type.

Example of a preference matrix for the KME-Chalkor and the Groupe Gascogne problems

Proposed Action	Reduce COM Discretion		Lower Cartel Fines		Separate Damages Proceeding	
Level of Resources	+ Resources =		+ Resources =		+ Resources =	
The Accused						
The Competitor						
The Consumer						
The Taxpayer						

2.3.5. Step 5: Finding the fair antitrust procedure

The fifth and the final step in applying the *model of fair rules* evolves around interpreting and working the numbers gathered during step three and four into a comprehensive solution to the problem of procedural fairness. During step four, we established the subjective preferences of the stakeholders involved to the problems in question, and during step three, we established the objective importance of the differing preference functions of the different stakeholders. We have thus established what each prefers subjectively, and how important that preference is objectively. During the final step, these two variables are merged into a third variable, which seeks to represent an objective value of the preferences of the stakeholders involved through the following formulation:

The objective value (OV) of an individual preference claim is a derivative from the subjective value (SV) of an individual preference claim and the applicable individual preference index (PI);

$$OV = PI * SV$$

For public policy purposes, we would want to maximise the aggregated objective value (OV) of a policy choice, conditioned on an equitable distribution of the individual subjective preference value (SV). We would thus want to maximise the overall efficiency of public policy, but only to the extent that individual subjects of the policy would have their private preferences fulfilled fairly, in comparison to others. To establish the combined objective value of a specific choice the numbers of the stakeholders holding a specific preference would need to be considered by multiplying their number with the individual objective value ($OV * (\text{Stakeholder} * n)$) or by multiplying their proportion of the whole with the individual objective value ($OV * (\text{stakeholder} * 1/n)$). In practice this means that not only the importance of the preference that individual stakeholders hold matters (i.e. the *preference index*), but also how many hold that particular preference of the relevant population.

The level of importance of a specific preference, directly affects its combined objective value, and the relative number of stakeholders holding that specific preference can either inflate, or deflate its combined objective value. Often there would be an inverse relation between the relative number of stakeholders holding a particular preference, and the individual objective

value of the preference, i.e., the accused in a cartel proceeding might hold preferences towards the accuracy of the proceeding that would objectively be recognised as very important, but due to how few the accused are as a proportion of the relevant population of stakeholders, the combined objective value of the accused's preference might be much lower proportionally than the individual objective value of his preference. The opposite could often be the case with consumers and taxpayers; they may hold comparatively unimportant individual preferences, but due to their vast numbers, the combined value of their preferences becomes an important consideration for designing public policy.

Referring to the balancing problems in *KME-Chalkor* and *Groupe Gascogne*, the preference matrix created during the fourth step should simply be adapted by adjusting the subjective preference values with the preference index for each stakeholder and the relative numerical size of each stakeholder's type. The matrixes should be updated using the following formulation on each of the preference values:

The combined objective value (COV) of the preference of a particular type of stakeholders is a function of the individual subjective preference value (SV) multiplied with the stakeholder's preference index (PI) and the relative size of the stakeholder type (1/n Stakeholder); $COV = SV * PI * (1/n \text{ Stakeholder})$

Following this adaptation, we have the total utility of each of the proposed action for each of stakeholder types, and thus by adding the utility of each of the stakeholder types together we have the total utility of each of the proposed action. If we were interested in maximising the utility of public policy, this would provide the information to enable the identification of the most efficient policy action for the purposes of maximising total welfare. We are however interested in identifying the optimally fair policy option, which means that the optimum of maximising total welfare is restricted by a distribution condition; only those policy options that guarantee equity of preference utility among the stakeholder types are eligible. The task is thus to identify among the eligible policy options, the single option that provides the highest aggregated utility of preference fulfilment. In practice, it could prove difficult to compose a policy proposal that would guarantee absolute equity; it could thus be necessary to define a

margin within which some inequality of preference fulfilment would be permitted, to enable the creation of more efficient policy options.

The basic reference value for comparing equity of preference fulfilment is the subjective preference matrix created in the context of step four. To account for differences in the objective value of the preferences, the preference index is used to transform the subjective values, into objective values. However, contrary to where the objective value was found for the purposes of establishing the combined objective value (COV), the subjective preference values (SV) would need to be divided by the preference index (PI), instead of multiplied. This is to facilitate comparison of the numbers, so that equal numbers represent an equitable distribution of preference fulfilment between different stakeholders holding preferences that are unequally important. Low intensity of unimportant preference would thus be required to match high intensity of an important preference to achieve equitable distribution of preference fulfilment. Using the cartel proceeding example, an objectively important preference of the accused would thus need to be fulfilled at a higher intensity, than an objectively unimportant preference held by the taxpayer, to achieve equitable distribution of preference fulfilment.

At the end of the fifth step, two decision matrixes should thus have been established: one representing the combined objective value (COV) of the preference fulfilment of the various stakeholders for each of the proposed actions, and another representing the relative level of objective preference fulfilment for each of the stakeholders involved with regards to each of the proposed actions. Using these two types of matrices, an efficient option can be identified that also fulfils the criteria of distributional fairness. For the purposes of procedural regulation, the legislative option that satisfies the criteria of equitable distribution of preference fulfilment among the relevant stakeholders, and does so with the utmost efficiency, can be branded as optimal in terms of procedural fairness.

Considering the *KME-Chalkor* problem in these terms, it seems evident that the accused in the initial cartel proceeding have most to gain from limiting the Commission's discretion to establish complex economic facts. The other stakeholders would seem to have less incentive

to support such changes and would potentially stand to lose from such changes. The net effect of the change could thus easily reduce total welfare by reducing the combined objective preference value (COV) towards the procedure at stake. This however shifts the attention to the distribution side of the equation. If the current status quo is fixed at an equilibrium that is very differently preferred by the relevant stakeholders in the objective sense, the consideration of procedural fairness might very well bar the continued use of the status quo due to unequal distribution of preference fulfilment. In that case, the other stakeholders might have to accept a reduction in their own preference fulfilment and a reduction in the total combined output of preferableness to accommodate the fairness claim of the accused. In the case of *KME-Chalkor*, this might mean that even if all the stakeholders involved (except for the accused) would stand to lose from limiting the discretion of the Commission, the right thing to do considering the optimisation of procedural fairness could still be to adopt the change if the result of the preference survey would provide an empirical validation.

The problem in *Groupe Gascogne* is slightly different; the status quo is a bit unclear and thus the competing options are two new alternatives rather than the status quo versus a change. A potential outcome of the comparison of these two options could be that one would be more efficient, while the other would be more equitable in terms of the distribution of preferableness. A call would then have to be made about whether both or only one option is sufficiently equitable. If both satisfy the criteria, the efficient option should be chosen, but if only the inefficient option satisfies the criteria, it would nonetheless have to be chosen as the only available option that satisfies the condition of procedural fairness. In *Groupe Gascogne* the applicants preferred an immediate reduction of the cartel fines imposed, over the alternative of having to start a separate procedure for claiming damages. The balancing problem could thus be said to hinge on whether the less preferable option was, from the point of view of the accused, nonetheless preferable enough to satisfy the condition of equitable distribution of preference fulfilment among the relevant stakeholders. The Court of Justice decided that the later procedure was indeed fair enough, and that it would be more efficient to use a universal procedure for settling damages claims that arise in the context of competition procedures. Empirical data about the preferences of the stakeholders in the cases could show whether the intuition of the Court about the preferableness of the competing

procedural alternatives was right or wrong, and whether it managed to identify the optimally fair procedure.

3. Assessing the application of the model of fair rules

In the first part of this chapter, we reviewed the cases before the Court of Justice where the question of procedural fairness in terms of Article 47 of the Charter of Fundamental Rights has been raised in the context of competition law litigation. The Court has already resolved several problems, most prominently two clusters of cases that evolved around the *KME-Chalkor* and the *Groupe Gascogne* litigations. The publicly available court documents provide an example of the arguments that can be raised for the purposes of protecting relevant stakeholder interests in the interpretation of the adequate fairness standard of competition law procedures; they do however not provide a definite account, although some of the most prominent stakeholders could submit their reasons. The two case clusters provided interesting examples for our purposes due to the Court's forced quasi-legislative approach to the problems, in the absence of a clear prior legislative decision on how the balancing issue should be resolved.

In the second part of the chapter, we used the main balancing problems of procedural fairness from the *KME-Chalkor* and the *Groupe Gascogne* litigations to show how the *model of fair rules* could be applied to such dilemmas, if framed as purely legislative problems. The arguments used by the Court and the parties to the cases, were used to approximate the preferences of the relevant stakeholders in the absence of an empirical survey. What we found from this exercise was that the arguments used by the Court for its conclusion in *Groupe Gascogne* were consequential in nature and could have been backed up with empirical evidences about the preferences of the relevant stakeholders, instead of relying on the intuition of the members of the Court about what was sufficiently fair, or optimally efficient. The finding in *KME-Chalkor* was supported by a more traditional deontological type of an argument about procedural imperatives, but again could have benefited from a more consequential approach in which the preferences of the relevant stakeholders would have been analysed in more details.

In the absence of actual data about the preferences of the stakeholders, it is hard to conclude whether the quasi legislative decisions of the Court of Justice in *KME-Chalkor* and *Groupe Gascogne* were warranted, viewed in terms of the optimally fair balance between procedural accuracy and procedural cost efficiency. We do however know that the *KME-Chalkor* decision to reduce the discretion of the Commission to establish complex economic fact is liable to increase the cost of competition procedures in general, but it is highly uncertain whether doing so will appreciably increase the accuracy of such procedures. We do also know that by choosing the expensive procedural alternative in *Groupe Gascogne*, of starting a separate procedure for claiming damages for undue delays in competition procedures, there is a risk of an overinvestment in the accuracy of an easy legal and factual question. It can be doubted whether the additional investment is cost beneficial for the relevant stakeholders involved in terms of the balance of procedural accuracy and cost efficiency. By applying the *model of fair rules* to these kinds of problems, methodological uniformity and accuracy is achieved in dealing with hard questions of what the law ought to be; both for legislators in the traditional parliamentary capacity, and for the occasional legislators wearing the scarlet robes of the Court of Justice of the European Union.

Conclusions

The aim of the thesis was broadly to provide an answer to *why a procedure is fair*, and to provide a methodological platform for establishing *when a procedure is fair*, using the context of EU's competition procedure as an example.

By examining the philosophical foundations of fairness in the laws I established a conception built on social contract theory that provides normative conditions that the laws (including the procedural laws) of the democratic state ought to comply with to enable a claim to fairness. A procedure is thus fair when it complies with the maxims of the social contract. By using insights for decision theory and expected utility theory I suggested the methodological alternative of the *model of fair rules* to implement the defined conception of fairness into legislative practice, thus providing the means to determine when a procedural design succeeds in complying with the fairness maxims of the social contract. The case study of the EU competition procedure provided a practical example of how the methodology could work in practice.

1. Thesis summary

Chapter by chapter, the output of the research can be summarised as follows:

The objective of this research, as explained in **chapter one**, is to examine the concept of fairness in the context of the laws, using EU's competition procedure and its interaction with the concept of procedural fairness as articulated by Article 6 of the ECHR and Article 47 of the Charter of Fundamental Rights as a case study. In contrast with the abundant, but inconclusive, literature on these issues, which focuses on the traditional juridical method for providing an answer to the question of what constitutes a fair competition procedure by reading and interpreting the relevant black letter legal sources, this research set out to analyse the philosophical foundations of the concept of fairness in the laws. The prospective utility of undertaking such a task is to provide an alternative methodological approach to fairness dilemmas in the laws. EU's competition procedure provides a practical context for testing the hypothesis of this novel methodology.

Chapter two starts by using the astrophysical concept of the dark matter as a metaphor for what might be absent in what we can visibly observe through the black letter law. The two orthodox theories of law, the naturalist and the positivist views, are at odds on whether there is anything necessary to the laws that cannot be observed by consulting the black letter sources. The positivist view would explain fairness in the laws with reference to the actual letter of the law, while the naturalist view would additionally refer to moral norms that exist independently beyond the written laws. For the purposes of answering the research question, this distinction is of cardinal importance; if the positivist view holds, then fairness in the laws is determined by the laws; if, however the naturalist view holds, then fairness in the laws is additionally subject to moral criteria that exists independently of the letter of the law. These orthodox legal theories both assume how morality connects with laws (or does not) without analysing more broadly how the laws connect with the overarching branch of political philosophy. A suggested route past the Hart/Dworkin debate, is to assess the laws in the context of the political organ that creates them. Arguably, a political organ that is based on the social contract idea, as operationalised through a democratic system of government, is bound to create laws that abide to the foundational premises of the social contract. As shown by Binmore, social cooperation can be formed and maintained instinctively, if it is efficient and the output of the cooperation is distributed through an egalitarian fairness principle. An equilibrium of egalitarian distributional fairness and efficiency thus forms the nucleus condition on the democratic state and the laws of the democratic state. An instance of law in the democratic state cannot claim compliance with the condition of fairness, unless it abides to the foundational equilibrium premise of the social contract, which can be viewed as the unseen dark matter of the laws.

Chapter three analyses potential methodologies for translating optimisation requirements, such as the one established in chapter two, into laws. An analysis of two categorically different orthodox approaches (the proportionality test and the cost benefit analysis) reveal weaknesses with each that make them unsuitable for the complete approach to optimisation that is required for balancing the fairness requirement of the social contract. While the proportionality test is good at promoting moral considerations, it is less fitting for considering efficiency and uses a primitive quantification methodology. The cost benefit analysis is more

robust in quantifying the relevant economic variables, but lacks rigour when assessing moral factors that are resistant to monetisation. Insights and concepts from decision theory provide clues to how the orthodox approaches can be methodologically improved. Expected utility theory as advanced by Von Neumann and Morgenstern, and later refined by Savage and Harsanyi is instrumental for understanding how rational decisions based on conflicting preferences can be optimised. The alternative model of prospect theory suggested by Kahneman and Tversky explains how deviations from the rational choice model can be dealt with and incorporated in decision models.

Chapter four starts by explaining how the role of the legislator should be understood in the context of social contract theory and how his essential task is to facilitate the basic premise of the social contract through his legislative work. Next, the act of legislating is framed as a decision problem so that the concepts and tools of decision theory can be applied in the legislative context. On the premise that the democratic legislator ought to engage his legislative task as a decision problem, where the optimum solution is in harmony with the egalitarian fairness concept of the social contract, chapter four suggest the *model of fair rules* as a methodological alternative to the orthodox methodologies reviewed in chapter three. The *model of fair rules* offers a comprehensive quantitative method for comparing competing legislative alternatives, considering the full spectrum of relevant moral and economic arguments. A comprehensive comparison of stakeholder preferences in a legislation, enables the identification of the legislative option that best satisfies the normative fairness criteria imposed by the social contract. The *model of fair rules* is thus a quantitative instrument for translating philosophical fairness into provisions of black letter law. In principle, the instrument is universally applicable for finding fair legislative alternatives, but for applying in specific context the precise balancing problem needs to be identified and described in terms that synchronizes with the decision theory terminology employed by the model.

Chapter five analyses law enforcement procedures with the aim of extracting their primary normative components and their primary instrumental function. The normative components are a source of potential balancing dilemmas, if disagreeing interests emphasise incompatible balancing solutions. The instrumental function explains the different typologies of procedural

provisions and how they affect the balancing towards the fair procedure in different ways. By examining the historical pedigrees of three types of enforcement procedures, a common instrumental function appears. Each procedural type is instrumentally focused on a decisional task that is preceded by a preparation phase where the relevant facts are gathered and processed to facilitate the decisional task. Depending on the procedure, one or several intervals of this instrumental mechanism of framing and deciding is required to complete the relevant procedural process. There can thus be several decision points within a single procedure, where balancing dilemmas may occur. The primary normative component of a procedure is accuracy; a procedure primarily ought to facilitate its instrumental task with accuracy. Accuracy however requires usage of resources that are limited in nature. Due to these limitations, a second normative component appears; accuracy should be achieved through the efficient use of the available resources. The practical limits on achievable accuracy creates a balancing dilemma on how many errors should be conceded, and which type of errors (type one or type two) should be preferred. This balancing dilemma can occur at each of the decision points identified through the instrumental analysis, and the procedural legislator needs to identify a fair solution to these dilemmas.

Chapter six identifies the object of facilitation for EU's competition procedure. While the procedural designer is occupied with balancing the normative components of accuracy and the efficient use of enforcement resources, he must assume a constant which his procedural project is meant to facilitate. In the case of EU competition procedure this constant is the normative objective of EU's substantive competition law regime. Historically, EU's competition policy was subject to the market integration imperative, which gave it different intellectual underpinnings than its US counterpart. Following the completion of the internal market in the 1990s and after the modernisation of the 2000s, the integration objective became less important and ideas about economic efficiency rose in prominence. A focus on the process of competition is an alternative to the pursuance of an external economic efficiency objective in competition policy. A focus on the process presumes the intrinsic value of competition per se, irrespective of external objectives. Following Arrow and Debreu, the two theorems of welfare economics predict that an ideal competitive process will result in a Pareto equilibrium between buyers and sellers, and that a specific distribution can be imposed separately. On this view, the

process of competition is a sensible policy if we want Pareto optimality in the marketplace and a separate mechanism (e.g. the tax system) for solving distributional issues. The current constitutional consensus on the objective of EU's competition policy decisively rejected a focus on the process of competition. Constitutionally, the competition policy is subordinated to the internal market objective and in regulatory practice to an economic welfare objective that allows Kaldor-Hicks improvements between competitors, as long as the improvements remain Pareto with regards to consumers. This objective is usually branded as consumer welfare or consumer interests, and is a mix of an economic welfare objective and a political distributional objective that favours consumers. The political compromise that forms the object of EU's competition policy, is the constant which the procedural regime seeks to facilitate as accurately as possible, given the available enforcement resources.

Chapter seven analyses EU's competition procedure based on the rational choice concepts of decision theory by identifying various stakeholders and the stakes that they seek to promote. The identity of the stakeholders is determined by their position with regards to potential breaches to the substantive provisions of the competition law regime. Their position controls their preferences with regards to the error balancing that occurs through the procedural regulation, both with regards to the public investment in total accuracy, and with regards to the types of errors conceded. An examination of EU's competition procedure reveals five major decision points that each is preceded by a framing process. Three of these decisions points fall within the competence of DG Competition and the Commission, but the last two decision points are within the jurisdiction of the Luxembourg courts. Individual procedural provisions can be defined both in terms of their instrumental function relating to either the framing or the decisional task, or in terms of the normative function as either accuracy enhancing or efficiency enhancing. Some provisions are mixed in either the instrumental or the normative sense, or both. In addition to the formal procedural provisions, DG Competition's internal code of practice tries to further increase the quality of the decisional process by the imposition of various quality mechanism.

Chapter eight starts by reviewing the case law of the Court of Justice where the fairness of EU's competition procedure has been challenged based on Article 47 of the Charter of Fundamental

Rights. Two case clusters appear through this review that each deal with a procedural dilemma that had not been answered decisively *before the fact* by a legislator. The absence of a *before the fact* prescription, forces the Court into the shoes of the legislator. In the *KME-Chalkor* cases the Court accepts this responsibility implicitly, but in the *Groupe Gascogne* cases it does so explicitly. The case clusters provide an account of a *before the fact* procedural balancing, where the optimisation requirement is procedural fairness in terms of the Charter of Fundamental Rights. After having analysed the arguments of the Court and the parties to the cases, the balancing dilemmas are measured against the *model of fair rules* to assess if any clarity could be gained. Given the conflicting and confusing arguments presented by the Court and the parties to the cases, a quantitative and methodological clarity could be increased by employing the *model of fair rules*. Consequently, the task of identifying the optimally fair procedural solution could be solved categorically with methodological soundness if a standardised method, such as the *model of fair rules*, would be accepted as the norm for solving such dilemmas. An acceptance of such a methodology by a legislator, would require courts to use an equivalent methodology when deciding legislatively undecided fairness issues, and when deciding legislatively decided issues, the interpretive methodology would need to respect the balancing of the legislator.

2. Concluding reflections

An understanding of morality in the laws, as an efficient equilibrium between the desires peoples that are subject to the laws, has important methodological implications. A juridical method that treats acts of legislation and past judicial findings as definitive argumentative constants risks being at odds with the current perception of society's fairness equilibrium. A legislation may reflect what once was perceived fair, but it does not necessarily follow that the current understanding of fairness is coordinated with what once was. An argument about fairness in the laws must thus be based on the underlying social contract equilibrium, instead of on what once was considered a current interpretation of that equilibrium as expressed by a piece of legislation, or by a finding of a court.

A juridical method in the democratic state that takes the temporal problem seriously, treats the fairness equilibrium as an ultimate norm of the legal system. An argument about legality

based on the written laws is thus defeasible by reference to the fairness equilibrium. The legislator is the ultimate arbitrator of what the current equilibrium is, and an argument meant to defeat a current legislation must thus maintain that the current laws are no longer in tune with what currently is perceived as fair. For a juridical method of this type to work, a coordinated methodology for establishing a present status of the fairness equilibrium with regards to distinct problems of law must be available, and the courts will need to assess the validity of arguments about fairness with reference to such methodology by measuring whether the common temperament has changed since the legislator addressed the problem during the preparation of the legislative act. For undecided legal questions, the courts would be permitted to resolve issues based on the same methodology.

To an extent, we can see that the classical juridical method, based on deontological maxims and interpretive canons, is already under a methodological stress due to the general awareness that few arguments are absolute and most claims need balancing with competing claims. This explains the successful penetration of the cost benefit analysis and the proportionality test in recent decades into mainstream legal practice; both methods complement the classical model by providing means to achieve the balancing of competing normative considerations. The downside of this patching is that the focus of the normative rationale of the system is lost, and the diverse methodologies for establishing legality risk giving incoherent answers to corresponding questions (i.e. is this proportional or cost beneficial?). The suggestion that this research would provide is that the best way forward is to rethink the entire legal method by assuming in its foundational premises the need for coherent balancing towards a defined ultimate normative consideration. To these ends I have proposed both a normative concept and a methodology for transforming that concept into legislative practice.

The essential novelty of the research rests in the synthesis of existing knowledge from political philosophy and economics theory, with a narrative of how fairness could be practically approached through the laws within the field of EU's competition law procedure. I would consider the research successful if the narrative is found to provide a plausible concept of fairness, suggest a plausible methodology for implanting the concept, and have demonstrated

how both the philosophical concept and the quantitative methodology could work in practice towards identifying procedural fairness in the design of EU's competition law procedure.

Firenze – Luxemburg - Reykjavik

September 2012 - May 2017

Bibliography

--*Brennu-Njáls saga* (~1270-1290) <http://www.sagadb.org/brennu-njals_saga> accessed 3

January 2017

Adler M, 'A Socio-Legal Approach to Administrative Justice' (2003) 25 *Law and Policy* 323

-- 'Fairness in Context' (2006) 33 *Journal of Law and Society* 615

-- 'Understanding and Analysing Administrative Justice' in Michael Adler (ed), *Administrative Justice in Context* (Hart Publishing 2009)

-- *Well-Being and Fair Distribution: Beyond Cost Benefit Analysis* (Oxford University Press 2011)

-- and **Posner EA**, *New Foundations of Cost-Benefit Analysis* (Harvard University Press 2006)

Alexy R, *A Theory of Constitutional Rights* (first published 1985, Julian Rivers tr, Oxford University Press 2002)

-- 'On Balancing and Subsumption – A Structural Comparison' (2003) 16 *Ratio Juris* 433

-- 'The Weight Formula', in Jerzy Stelmach, Bartosz Brożek, Wojciech Załuski (eds), *Studies in the Philosophy of Law - Frontiers of the Economic Analysis of Law* (Jagiellonian University Press 2007)

-- 'Constitutional Rights and Proportionality' (2014) 22 *Revus* 51

Allais M, 'Le Comportement de l'Homme Rationnel devant le Risque, Critique des Postulats et Axiomes de l'Ecole Americaine' (1953) 21 *Econometrica* 503

Arrow KJ, 'An Extension of the Basic Theorems of Classical Welfare Economics' in Jerzy Neyman (ed) *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press 1951)

Aumann RJ, 'Agreeing to Disagree' (1976) 4 *Annals of Statistics* 1236

Banchich TM, Marenbon J and Reid CJ, 'The Revival of Roman Law and Canon Law' in Fred D. Miller and Carrie-Ann Biondi (eds) *A Treaties of Legal Philosophy and General Jurisprudence – Vol 6 – A History of the Philosophy of Law from the Ancient Greeks to the Scholastics* (Springer 2007)

Barberis NC, 'Thirty Years of Prospect Theory in Economics: A Review and Assessment' (2013) 27 *The Journal of Economic Perspectives* 173

Barbier de La Serre E, 'Procedural justice in the European Community case-law concerning the rights of the defence: essentialist and instrumental trends (2006) 12 *European Public Law* 225

-- 'Standard of review in competition law cases: Posten Norge and beyond' in Carl Baudenbacher, Philipp Speitler and Bryndís Pálmarsdóttir (eds), *The EEA and the EFTA Court: decentred integration: to mark the 20th anniversary of the EFTA Court* (Hart Publishing, 2014)

Binmore K, *Game Theory and the Social Contract Vol. I: Playing Fair* (The MIT Press 1994)

-- *Game Theory and the Social Contract Vol II: Just Playing* (The MIT Press 1998)

-- *Natural Justice* (Oxford University Press 2005)

-- *Rational Decisions* (Princeton University Press 2011)

Bork RH, *The Antitrust Paradox* (Free Press 1978)

Bronckers M and Vallery A, 'Fair and effective competition policy in the EU: which role for authorities and which role for the courts after Menarini?' (2012) 8 *European Competition Journal* 283

Burazin L, 'Can there be an artefact theory of law?' (2016) 29 *Ratio Juris* 385

Cappelletti M and Garth BG, 'Introduction – Policies, Trends and Ideas in Civil Procedure' in Mauro Cappelletti (ed), *International Encyclopaedia of Comparative Law – Volume XVI - Civil Procedure* (Martinus Nijhoff Publishers 1987)

Castillo de la Torre F, 'Evidence, proof and judicial review in cartel cases (2009) 32 World Competition 505

Coleman J, *The Practice of Principle: In Defence of a Pragmatist Approach to Legal Theory* (Oxford University Press 2003)

Cooter R and Ulen T, *Law and Economics* (6th edn, Pearson 2012)

Craig P, *Administrative Law* (7th edn, Sweet & Maxwell 2012)

Dabbah MM, *International and Comparative Competition Law* (Cambridge University Press 2010)

Dawkins R, *The Selfish Gene* (Oxford University Press 1976)

Debreu G, *Theory of Value - An Axiomatic Analysis of Economic Equilibrium* (Wiley 1959)

Dworkin R, 'The Model of Rules' (1967) 35 The University of Chicago Law Review 14

-- *Taking Rights Seriously* (Harvard University Press, 1977)

-- 'Is wealth a value?' (1980) 9 The Journal of Legal Studies 191

-- 'Rights as Trumps' in Jeremy Waldron (ed), *Theories of Rights* (Oxford University Press 1984)

-- 'Principle, Policy, Procedure' in Ronald Dworkin, *A Matter of Principle* (Clarendon Press 1986)

-- 'Thirty years on: Book Review of The Practice of Principle: In Defence of a Pragmatist Approach to Legal Theory by Jules Coleman' (2002) 115 Harvard Law Review 1655

-- 'Hart's Postscript and the Character of Political Philosophy' (2004) 24 Oxford Journal of Legal Studies 1

-- *Justice for Hedgehogs* (Harvard University Press 2011)

Einarsson ÓJ, 'EC competition law and the right to a fair trial' (2006) 25 Yearbook of European Law 555

Einstein A, 'On the Method of Theoretical Physics' (1934) 1 *Philosophy of Science* 163

Ellsberg D, 'Risk, Ambiguity and the Savage Axioms' (1961) 75 *The Quarterly Journal of Economics* 643

Esmein A, *A History of Continental Criminal Procedure with Special Reference to France* (John Simpson tr; Little, Brown, and Company 1913)

Ezrachi A, 'Sponge' (2015) Working Paper CCLP (L) 42. 18-19
<<https://www.law.ox.ac.uk/sites/files/oxlaw/cclpl42.pdf>> accessed 30 December 2016

Flattery J, 'Balancing efficiency and justice in EU competition law: elements of procedural fairness and their impact on the right to a fair hearing' (2010) 7 *The Competition Law Review* 53

Forrester IS, 'Due process in EC competition cases: a distinguished institution with flawed procedures' (2009) 34 *European Law Review* 817

Frank RH, 'Why is cost-benefit analysis so controversial?' (2000) 29 *The Journal of Legal Studies* 913

Fuller LL and Winston KI, 'Forms and Limits of Adjudication' (1978) 92 *Harvard Law Review* 353

Gardner J and Macklem T, 'Scott J Shapiro: Legality' (2011) *Notre Dame Philosophical Reviews*
<<http://ndpr.nd.edu/news/27609-legality/>> accessed 19 January 2017

Gauthier D, *Morals by Agreement* (Oxford University Press 1986)

Gerber D, 'The Transformation of the European Community Competition Law?' (1994) 35 *Harvard International Law Journal* 97

-- *Law and Competition in the Twentieth-Century Europe: Protecting Prometheus* (Oxford University Press 2001)

- 'Two Forms of Modernization in European Competition Law' (2007) 31 *Fordham International Law Journal* 1235
- *Global Competition: Law, Markets, and Globalization* (Oxford University Press 2010)
- Gintis H**, *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences* (Princeton University Press 2008)
- Greenberg M**, 'How Facts Make Law' (2004) 10 *Legal Theory* 157
- Hansson SO**, *Decision Theory: A Brief Introduction* (revised edn, KTH Stockholm 2005)
- <<http://home.abe.kth.se/~soh/decisiontheory.pdf>> accessed 29 December 2016
- *The Ethics of Risk: Ethical analysis in an uncertain world* (Palgrave Macmillan 2013)
- and **Grüne-Yanoff T**, 'Preferences' in Edward N. Zalta (ed), *The Stanford Encyclopaedia of Philosophy* (Winter 2012)
- <<http://plato.stanford.edu/archives/win2012/entries/preferences/>> accessed 29 December 2016
- Harlow C and Rawlings R**, *Process and Procedure in EU Administration* (Hart Publishing 2014)
- Harsanyi J**, 'Games with Incomplete Information Played by Bayesian Players I-III' (1967-1968) 14 *Management Science* 159 (I); 320 (II); 486 (III)
- Hart HLA**, 'Positivism and the Separation of Law and Morals' (1958) 71 *Harvard Law Review* 593
- 'The Presidential Address: Prolegomenon to the Principles of Punishment' (1959) 60 *Proceedings of the Aristotelian Society* 1
- *The concept of law* (first published 1961, 2nd edn, Clarendon Press 1994)
- 'The Postscript' in *The concept of law* (first published 1961, 2nd edn, Clarendon Press 1994)

Hauger NFW and Palzer C, 'Investigator, prosecutor, judge ... and now plaintiff? The Leviathanian role of the European Commission in the light of fundamental rights' (2013) 36

World Competition 565

Himma KE, 'Inclusive Legal Positivism' in Jules L. Coleman and Scott Shapiro (eds), *The Oxford Handbook of Jurisprudence and Philosophy of Law* (Oxford University Press 2004) 125.

Hobbes T, *On the Citizen* (first published 1642, Richard Tuck and Michael Silverthorne eds, Cambridge University Press 1998).

-- *Leviathan* (first published 1651, Richard Tuck ed, 2nd edn, Cambridge University Press 1996)

Hume D, *Treaties of Human Nature* (first published 1738-40, David Fate Norton and Mary J. Norton eds, Oxford University Press 2007)

Inwood B and Miller FD, 'Law in Roman Philosophy' in Fred D. Miller and Carrie-Ann Biondi (eds) *A Treaties of Legal Philosophy and General Jurisprudence – Vol 6 – A History of the Philosophy of Law from the Ancient Greeks to the Scholastics* (Springer 2007)

Jaeger M, 'The standard of review in competition cases involving complex economic assessments: towards the marginalisation of the marginal review?' (2011) 2 *Journal of European Competition law & Practice* 295

Jestaedt M, 'The Doctrine of Balancing – Strengths and Weaknesses' in Matthias Klatt (ed), *Institutionalized Reason: The Jurisprudence of Robert Alexy* (Oxford University Press 2012)

Kagan RA, 'The Organisation of Administrative Justice Systems: The Role of Political Mistrust' in Michael Adler (ed), *Administrative Justice in Context* (Hart Publishing 2009)

Kahneman D, *Thinking fast and slow* (Farrar, Straus and Giroux 2011)

-- and **Tversky A**, 'Prospect Theory: An Analysis of Decision under Risk' (1979) 47 *Econometrica* 263

-- and **Tversky A**, 'Advances in Prospect Theory: Cumulative Representation of Uncertainty' (1992) 5 *Journal of Risk and Uncertainty* 297

Kant I, *Groundwork for the Metaphysics of Morals* (first published 1785, Mary Gregor and Jens Timmermann eds, 2nd edn, Cambridge University Press 2012)

Kelsen H, *Pure theory of Law* (first published 1934, University of California Press 1967)

Kissinger H, *World Order* (Allen Lane, London 2014)

Köszegi B and Rabin M, 'A model of reference-dependent preferences' (2006) CXXI *The Quarterly Journal of Economics* 1133

-- 'Reference-dependent Risk Attitudes' (2007) 97 *American Economic Review* 1047

-- 'Reference-dependent Consumption Plans' (2009) 99 *American Economic Review* 909

Laudan L, 'The Rules of Trial, Political Morality, and the Cost of Error: Is Proof Beyond Reasonable Harm Doing More Harm than Good?' in Leslie Green and Brian Leiter (eds), *Oxford Studies in Philosophy of Law: Volume I* (Oxford University Press 2011)

Leiter B, 'American Legal Realism' (2002) *The University of Texas School of Law Public Law and Legal Theory Research Paper No. 042* <<https://ssrn.com/abstract=339562>> accessed 3 January 2017

-- 'The end of Empire: Dworkin and the Jurisprudence in the 21st Century' (2005) 36 *Rutgers Law Journal* 165

-- 'Legal Positivism About the Artifact Law: A Retrospective Assessment' in L. Burazin, K. E. Himma, C. Roversi (eds) *Law as an Artifact* (Oxford University Press, *forthcoming 2017*) <<https://ssrn.com/abstract=2870877>> accessed 16 January 2017

Lenaerts K and Vanhamme J, 'Procedural rights of private parties in the community administrative process (1997) 34 *Common Market Law Review* 531

Leucht B and Marquis M, 'American Influences on EEC Competition Law' in Kiran Klaus Patel and Heike Schweitzer, *The Historical Foundations of EU Competition Law* (Oxford University Press 2013)

Lianos I, 'Some Reflections on the Question of the Goals of EU Competition Law' (2013) UCL CLES Working Paper Series 3/2013 <https://www.ucl.ac.uk/cles/research-paper-series>
accessed 22 October 2015

Locke J, *Two Treatises of Government* (first published 1689, Peter Laslett ed, Cambridge University Press 1988)

Luce DR and Raiffa H, *Games and decisions: Introduction and critical survey* (Courier Corporation 1957)

Machiavelli N, *The Prince* (first published 1513, Quentin Skinner and Russel Price eds, Cambridge University Press 1988)

Marmor A, 'Exclusive Legal Positivism' in Jules L. Coleman and Scott Shapiro (eds), *The Oxford Handbook of Jurisprudence and Philosophy of Law* (Oxford University Press 2004)

Mashaw JL, *Bureaucratic Justice: Managing Social Security Disability Claims* (Yale University Press 1983)

McCarty N and Meirowitz A, *Political Game Theory: An Introduction* (Cambridge University Press 2014)

Metzger E, 'An Outline of Roman Civil Procedure' (2013) 9 Roman Legal Tradition 1

Montage F, 'The case for a radical reform of the infringement procedure under regulation 17' (1996) 17 European Competition Law Review 428

Monti G, *EC Competition Law* (Cambridge University Press 2007)

Monti M, 'Fighting Cartels Why and How? Why should we be concerned with cartels and collusive behaviour?' (2000) 3rd Nordic Competition Policy Conference, Stockholm
<[http://europa.eu/rapid/press-release SPEECH-00-295 en.htm](http://europa.eu/rapid/press-release_SPEECH-00-295_en.htm)> accessed 9 September 2016

Mousourakis G, *Roman Law and the Origins of the Civil Law Tradition* (Springer 2015)

Nazzini R, *The Foundations of European Union Competition Law: The Objective and Principles of Article 102* (Oxford University Press 2011)

-- 'Administrative enforcement, judicial review and fundamental rights in EU competition law: a comparative contextual-functional perspective' (2012) 49 *Common Market Law Review* 971

Oliver P, "'Diagnostics" – a judgment applying the Convention of Human Rights to the field of competition' (2012) 3 *Journal of European Competition Law & Practice* 163

Posner RA, *Antitrust Law* (University of Chicago Press 1976)

-- *Economic Analysis of Law* (7th edn, Aspen 2007)

Rawls J, *A Theory of Justice* (1971, 2nd edn, Belknap Press 1999)

Rousseau J-J, *A Discourse on the Basis and Origin of inequality Among Men* (first published 1754, Bedford/St. Martin's 2010)

-- *The Social Contract* (first published 1762, Christopher Betts tr, Oxford University Press 1994)

Sartor G, 'The Logic of Proportionality: Reasoning with Non-Numerical Magnitudes' (2013) 14 *German Law Journal* 1419

Savage LJ, 'The Foundations of Statistics Reconsidered' in Jerzy Neyman (ed), *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability - Volume 1: Contributions to the Theory of Statistics* (University of California Press 1961) 575

-- *The Foundation of Statistics* (first published 1954, 2nd edn, Dover Publications 1972)

Scherer FM and Ross DR, *Industrial Market Structure and Economic Performance* (3rd edn, Houghton Mifflin 1990)

Schweitzer H, 'Judicial review in EU competition law' in Damien Geradin and Ioannis Lianos (eds), *Handbook on European competition law* (Edward Elgar Publishing 2013)

Sen A, 'The Discipline of Cost-Benefit Analysis' (2000) 29 *The Journal of Legal Studies* 931

Shannon CE, 'Programming a Computer for Playing Chess' (1950) 41 *Philosophical Magazine* 314

Shapiro SJ, *Legality* (Belknap Press, 2011)

Sibony A-L, 'Annotation of the judgment of the Court in Case C-272/09 P *KME Germany and others v. Commission*' (2012) 49 *Common Market Law Review* 1977

Skyrms B, *Evolution of the Social Contract* (1996, 2nd edn, Cambridge University Press 2014)
-- *The Stag Hunt and the Evolution of the Social Structure* (Cambridge University Press 2004)

Slater D, Thomas S and Waelbroeck D, 'Competition law proceedings before the European Commission and the right to a fair trial: no need for reform?' (2008) *The Global Competition Law Centre Working Paper Series - Working Paper No. 04/08*
<<https://www.coleurope.eu/content/gclc/documents/GCLC%20WP%2004-08.pdf>> accessed 3 January 2017
-- 'Competition law proceedings before the European Commission and the right to a fair trial: no need for reform?' (2009) 5 *European Competition Journal* 97

Solum LB, 'Procedural Justice' (2004) 78 *Southern California Law Review* 181

Spencer JR, 'Introduction' in Mirielle Delmas-Marty and J.R. Spencer (eds), *European Criminal Procedures* (Cambridge University Press 2002)

Stiglitz JE, 'The Invisible Hand and Modern Welfare Economics' (1991) NBER Working Paper Series, Working Paper No 3641 <<http://www.nber.org/papers/w3641.pdf>> accessed 30 December 2016

Summers SJ, *Fair Trials: The European Criminal Procedural Tradition and the European Court of Human Rights* (Hart Publishing, Oxford 2007)

Sunstein CR, 'The Storrs Lectures: Behavioural Economics and Paternalism' (2013) 122 Yale Law Journal 1826

-- 'The Limits of Quantification' (2014) 102 California Law Review 1369

Sweet AS and Mathews J, 'Proportionality Balancing and Global Constitutionalism' (2008) 47 Columbia Journal of Transnational Law 72

Tollenaar A and De Ridder K, 'Administrative justice from a Continental European Perspective' in Michael Adler (ed), *Administrative Justice in Context* (Hart Publishing 2009)

Tullock G, *Trials on Trial: Pure Theory of Legal Procedure* (Columbia University Press 1980)

van Caenegem RC, 'History of European Civil Procedure' (1973) in Mauro Cappelletti (ed), *International Encyclopaedia of Comparative Law – Volume XVI - Civil Procedure* (Martinus Nijhoff Publishers 1987)

van Rhee CH, 'Civil Procedure: a European *Ius Commune*?' (2000) 8 European Review of Private Law 589

-- 'Introduction' in C.H. van Rhee (ed) *European Traditions in Civil Procedure* (Intersentia 2005)

-- 'The Influence of the French Code de Procédure Civile (1806) in 19th Century Europe' in: L. Cadiet and G. Canivet (eds) *De la Commémoration d'un code à l'autre: 200 ans de procédure civile en France* (LexisNexis Litec 2006)

Vandenborre I and Goetz T, 'EU competition law procedures' (2012) 3 *Journal of European Competition Law & Practice* 578

von Neumann J and Morgenstern O, *The Theory of Games and Economic Behavior* (first published 1944, 3rd edn, Oxford University Press 1954)

Wade HWR and Forsyth CF, *Administrative Law* (9thedn, Oxford University Press 2004)

Waelbroeck D and Fosselard D, 'Should the decision-making power in EC antitrust procedures be left to an independent judge? – The impact of the European Convention of Human Rights on EC antitrust procedures' (1994) 14 *Yearbook of European Law* 111

Weiler JHH, 'The Transformation of Europe' (1991) 100 *The Yale Law Journal* 2403

Weitbrecht A, 'From Freiburg to Chicago and beyond – the first 50 years of European Competition law' (2008) 29(2) *European Competition Law Review* 81

Whelan P, 'Cartel criminalization and the challenge of "moral wrongfulness"' (2013) 33 *Oxford Journal of Legal Studies* 535

Wils WPJ, 'The combination of the Investigative and prosecutorial function and the adjudicative function in EC antitrust enforcement: a legal and economic analysis' (2004) 27 *World Competition* 201

-- 'The increased level of EU antitrust fines, judicial review and the ECHR' (2010) 33 *World Competition* 5

-- 'EU anti-trust enforcement powers and procedural rights and guarantees: the interplay between EU law, national law, the Charter of Fundamental Rights of the EU and the European Convention on Human Rights' (2011) 34 *World Competition* 189

-- 'The compatibility with fundamental rights of the EU antitrust enforcement system in which the European Commission acts both as investigator and as a first-instance decision maker' (2014) 37 *World Competition* 5

Witt AC, 'From Airtours to Ryanair: is the more economic approach to EU merger law really about more economics?' (2012) 49 *Common Market Law Review* 217

Zweigert K and Kötz H, *Introduction to Comparative Law* (Tony Weir tr, 3rdedn, Oxford University Press 1998)

