
The Transaction Costs of Semi-Public Institutions

Railway History as a Challenge for Coase

Martin Holterman



THE TRANSACTION COSTS OF SEMI-PUBLIC INSTITUTIONS

Railway History as a Challenge for Coase

Ph.D. dissertation committee

Supervisors:

prof. dr. N.S. Groenendijk

prof. dr. R.A. Wessel

dr. E. Woerdman

Members:

prof. dr. ir. A.G. Doree (University of Twente)

prof. dr. L.L. Roberts (University of Twente)

prof. dr. O. Couwenberg (University of Groningen)

prof. dr. dr. M.P. Finger (École Polytechnique Fédérale de Lausanne)

ISBN 978-90-365-3254-9

Printed by PrintPartners Ipskamp, Enschede

Copyright © 2011 by Martin Holterman

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the written permission of the author.

**THE TRANSACTION COSTS OF SEMI-PUBLIC INSTITUTIONS:
RAILWAY HISTORY AS A CHALLENGE FOR COASE**

DISSERTATION

TO OBTAIN

THE DOCTOR'S DEGREE AT THE UNIVERSITY OF TWENTE,

ON THE AUTHORITY OF THE RECTOR MAGNIFICUS,

PROF.DR. H. BRINKSMA,

ON ACCOUNT OF THE DECISION OF THE GRADUATION COMMITTEE,

TO BE PUBLICLY DEFENDED

ON THURSDAY SEPTEMBER 8, 2011 AT 16.45

by

Martin Willem Holterman

born on 16 November 1981

in Emmen.

This dissertation has been approved by the promotor:

Prof. dr. N.S. Groenendijk

Prof. dr. R.A. Wessel

Dr. E. Woerdman

Table of Contents

Table of Contents	6
Preface	10
1. Introduction	14
1.1 Introduction of the Problem	14
1.2 Preliminary Remarks	19
1.2.1 Law and Economics	19
1.2.2 The Nature of Economics	21
1.3 Research Questions and Outline	26
2. Theory	30
2.1 The Law and its Incompleteness	32
2.1.1 The Nature of Law	32
2.1.2 Contractual Incompleteness	36
2.1.3 Obligational Incompleteness	37
2.1.4 Insufficient State Contingency	42
2.1.5 Observability and verifiability	47
2.2 The Coase Theorem	51
2.2.1 Introduction	51
2.2.2 Torts and the Coase Theorem	52
2.2.3 Exceptions to the Coase Theorem	55
2.3 The Nature of Transaction Costs	63
2.4 Williamson	72
2.4.1 Characteristics of Transactions	72
2.4.2 Differences between Governance Structures	79
2.4.3 Cable television in Oakland, California	82
2.5 Incomplete Contract Theory	87
2.5.1 Grossman & Hart (1986): The Costs and Benefits of Ownership	87
2.5.2 Hart & Moore (1988): Incomplete Contracts and Renegotiation	89
2.5.3 Hart & Moore (1990): Property Rights and the Nature of the Firm	90
2.5.4 Hart & Moore (1999): Foundations of Incomplete Contracts	92
2.5.5 Hart, Shleifer & Vishny (1997)	93

2.5.6 Evaluation.....	95
2.6 Alignment Mechanisms.....	100
2.6.1 Introduction.....	100
2.6.2 Evolution.....	101
2.6.3 Competition.....	106
2.6.4 Design.....	109
2.6.5 Conclusion.....	115
2.7 Conclusion.....	117
3. History of the Railways.....	122
3.1 Beginnings.....	123
3.1.1 Company Law.....	126
3.1.2 Stockton & Darlington.....	134
3.1.3 The Liverpool & Manchester.....	139
3.1.4 The Netherlands.....	144
3.1.5 France.....	150
3.1.6 Belgium.....	157
3.1.7 Conclusion.....	161
3.2 Consolidation.....	167
3.2.1 Great-Britain.....	167
3.2.2 The Netherlands.....	176
3.2.3 France.....	180
3.2.4 Belgium.....	187
3.2.5 Germany.....	188
3.2.6 Conclusion.....	192
3.3 Nationalisation.....	200
3.3.2 Belgium.....	206
3.3.3 The Netherlands.....	210
3.3.4 France.....	214
3.3.5 Great-Britain.....	217
3.3.6 Conclusion.....	219
3.4 Privatisation: Thatcher and Beyond.....	224
3.4.1 Great-Britain.....	225
3.4.2 The EU Railway Packages.....	234

3.4.3 Conclusion.....	247
3.5 Conclusion.....	254
4. Railways: Alternative Approaches.....	260
4.1 Public Choice Theory.....	262
4.1.1 Introduction.....	262
4.1.2 Rent Seeking.....	264
4.1.3 Politicians.....	266
4.1.4 Bureaucracy.....	267
4.1.5 Dictators and Monarchs.....	269
4.1.6 Conclusion.....	270
4.2 Public Choice Theory and Railways.....	272
4.2.1 Beginnings.....	272
4.2.2 Consolidation.....	276
4.2.3 Nationalisation.....	281
4.2.4 Privatisation.....	285
4.2.5 Conclusion.....	290
4.3 Path Dependence.....	294
4.4 Path Dependence and Railways.....	300
4.4.1 Beginnings.....	300
4.4.2 Consolidation.....	302
4.4.3 Nationalisation.....	308
4.4.4 Privatisation.....	313
4.4.5 Conclusion.....	317
4.5 Conclusion: Alternative Approaches.....	320
5. Conclusion.....	324
5.1 Theoretical Findings.....	324
5.2 Historical Findings.....	327
5.3 Findings on Benchmarks.....	332
5.4 Positive and Normative Statements.....	335
5.5 NIE & the Public Sector: The Future.....	337
Literature.....	340
Nederlandstalige Samenvatting.....	368

Preface

At this point, it is customary to express glowing gratitude to those who have made this thesis possible, all the while ignoring to the best of one's ability the contributions of the author himself. Much as the “proefschrift” – *nomen est omen* – is the ultimate display of one's ability as an academic, for just a page or two this fact is cast aside in favour of feigned or genuine modesty. We can only speculate that our forebears have found it appropriate to celebrate the moment at which they receive their degree of *Philosophiae Doctor* by commemorating one of the most important lessons taught by the man to whom all other philosophy is but footnotes: “He, O men, is the wisest, who, like Socrates, knows that his wisdom is in truth worth nothing.”¹ In this light, I too will turn to those around me before displaying the fruit of my efforts.

To begin with, none of this would of course have been possible without the support of those who obtained the funding for this project, and who consented to hire me to carry it out. While I am to this day not entirely sure where the money came from all these years, it is clear that Bert Steenge, Dick Ruiter and Edwin Woerdman had a lot to do with it. In addition to those three, Michiel Heldeweg, Marieke van Genugten and Wim van de Griendt were also involved in the decision to hire me. To them I am grateful above all others, because ultimately they made it possible for me to spend five years sitting on my backside surfing the internet, reading everything of note that was even vaguely connected to law or economics, as well as a lot of things that were quite clearly not work-related at all. Thank you.

Of course, such work methods were only possible because I enjoyed the privilege of being supervised by two properly “old-school” professors, who did not care much for such innovations as “hands-on supervision” and “progress reports”. I was sincerely comforted by the discovery that such things are still possible in modern day academia. Long before I came to Enschede, I was already accustomed to working slowly and independently, and thanks to Bert and Dick I was able to continue to do so. Of course, the result was that I took a year longer than I was supposed to, but I think that is a price well worth paying. Moreover, whenever I did turn up to speak with one of them, I always ended up captivated by fascinating conversation, albeit fascinating in much the same way that my on-line reading was fascinating,

¹ Plato, *Apology*, par. 10.

that is to say interesting but often only coincidentally useful. Thank you Bert and Dick for being near-infinite sources of knowledge.

My gratitude also goes out to my colleagues in the LEGS department. Not because this thesis would not have been possible without them, since – with the exception of Ria Lusschen – that is clearly not the case, but because my time at the University of Twente would not have been the same without them. In this regard, thanks in particular to the various people with whom I have shared an office over the years: Alina Fazal, Marija Bockarjova, Tsjalle van der Burg, Carlie Geerdink, Chris Breuer, Sabine Hoogstad, Evisa Kica and Nupur Chowdhury. I am sorry for all the times I have kept you from your work, and grateful for the fact that you let me. Similarly, I would like to thank the various colleagues whose expertise I have borrowed over the years for reasons unconnected to this dissertation. Unfortunately, despite the best efforts of prof. Wessel, I still do not understand why the EU Treaties are not “international agreements” in the sense of art. 103 UN Charter. Perhaps AG Maduro will be able to explain this to me someday.

Prof. Wessel also warrants mention because he graciously agreed to step in as one of my new thesis supervisors following the retirement of prof. Ruiter. Despite my best efforts to write a thesis that was intelligible for economists and jurists alike, this must have been no small task. The same goes, of course, for prof. Groenendijk, who replaced prof. Steenge. The result is that this dissertation is the fruit of my collaboration with no fewer than five different supervisors, each with their own distinct style and academic background, and I can only say that I hope that I have done you all proud. This goes particularly for Edwin Woerdman, with whom I have spent many pleasant – and useful – hours talking about my thesis when I visited him in Groningen. His feedback has been invaluable.

That leaves only the people who have generously agreed to be my paranimphs, Chris Breuer and Saskia Holterman. In the case of the former, it was my honour and privilege to act in the same capacity for him on April 20, 2011, whereas my dear sister was so enthusiastic about being the second paranimph that I did not dare refuse. Thanks to both of you for your help on the day of my promotion, thank you Chris for your friendship, and thank you Sister for, well, everything. The same ill-defined expression of gratitude goes out to my father, Jan Holterman, who also made all of this possible in more ways than I can list.

Finally, I have to thank the person without whom this preface would not have existed. I would have undoubtedly recoiled from the task of putting to words the gratitude that I would

otherwise have preferred to leave unspoken, but for the fact that my mother Geesje Holterman insisted that this would have been the height of ingratitude. She was right, of course, as she always is. Socrates said that his wisdom might be worthless, but that he was still wiser than everyone else in society. Perhaps he should have spent a little more time talking to his mother.

1. Introduction

1.1 Introduction of the Problem

In the social sciences, there are two things one can do with a theory and the aspect of reality it is meant to explain: one can use the theory to improve reality², or one can use reality to improve the theory. The purpose of the present thesis is very much to do the latter. The empirical problem concerns the proper scope of government³, which we will use to examine the new institutional economics (hereafter: NIE) as it was developed starting with the work of Ronald Coase.⁴ For reasons that will be explained below, we will focus particularly on the history of the railway industry in Western Europe.

Starting with case studies by Coase⁵ and Oliver Williamson⁶, and more recently with theoretical work by Hart, Shleifer & Vishny (1997) and Williamson (1999), it has been suggested that the notion of transaction costs, which is the cornerstone of NIE, could be used equally well to analyse the proper scope of government as it could be to analyse the proper scope of firms. In the private sector, the theory describes the choice of a firm to outsource an activity or to do it in-house, decisions about mergers and split-offs. The argument was that the decision of the government to privatise or nationalise was functionally equivalent, meaning that it could be understood using the same theoretical tools.

The problem with this idea came to light when Williamson decided to tackle the problem head-on. He wrote an article about the possibility of privatising the US State Department.⁷ Since no one ever proposed privatising the foreign ministry, one would suspect that there are no immediate benefits to be expected from such a move. It follows, Williamson argued, that it should be possible to cast this problem in NIE terms and get the same result. He discovered, however, that he needed to invoke a new kind of transaction, which he called

² Eg. Winston (2006).

³ Cf. Hart, Shleifer & Vishny (1997).

⁴ Coase (1937, 1960). While it is a tedious distinction to make, we will have to insist on the difference between the new institutional economics and neo-institutional economics. The latter includes the former, but also related fields such as path dependency and – depending on who you ask – public choice theory. Since these last two will not be discussed until chapter 4, and are in any event not the main subject of this dissertation, we will use the term “new institutional economics” to refer only to Coase and his followers. Cf. Woerdman (2002), p. 45.

⁵ Coase (1959).

⁶ Williamson (1976).

⁷ Williamson (1999).

“sovereign transaction”, which was characterised by something called “probity” in addition to the ordinary transaction characteristics that his model normally relied on.⁸ Whether that conclusion really followed from the facts is a question we will have occasion to return to later on in this thesis. For now, the point of interest is his choice of problem. Is the foreign ministry really a fair and sensible aspect of reality with which to confront an economic theory of governance?

In 1966, Stanley Hoffmann analysed the process of European integration in terms of High Politics and what later became known as Low Politics.⁹ Put simply, the former concerns security and diplomacy, while the latter deals with money and business. Stating it in institutional economic terms, his argument was that functionalism, a theory from the field of international relations akin to various theories in economics regarding governance structures, only worked in low politics because only in those areas were governance structures designed in any manner that an economist would recognise. Treaties made in the area of high politics are sovereign transactions in a way that their low politics counterparts are not, causing the normal logic to break down. Similarly, impeaching an economic theory of public sector governance because it does not adequately explain the high politics problem of whether the foreign ministry could be privatised seems somewhat rash, at best.¹⁰

Other authors have generally followed this rationale, focusing instead on activities that are further away from the core political function of the government. For example, Hart, Shleifer & Vishny (1997) and Volokh (2007) wrote about US prisons, Ménard & Saussier (2002) chose the French drinking water industry, while Van Genugten (2008) used a case study taken from the Dutch garbage collection industry. Not only are such activities *a priori* more likely to be governed by economic principles, the advantage of their study is also that there is a wider variety of governance structures available for study. Unlike diplomacy, the running of a prison has in fact been privatised in many US states, allowing the economist to study whether the observations correlate with the variables identified by the model in the manner hypothesised. That said, one can take this method too far.

⁸ Cf. Ruiter (2005).

⁹ Hoffmann (1966).

¹⁰ We will have occasion to examine whether this diagnosis is correct when we examine Williamson’s approach more fully below. There, as throughout this thesis, we will argue that there is no need to introduce a special category of “sovereign transactions”, but that the problem is better understood as one of ordinary moral hazard. This diagnosis is supported by the recent rediscovery of the fact that war may be eminently sovereign, but that it is still quite possible to privatise it. Cf. Kaplan (2007), Francioni (2008) and Chesterman (2008).

There is a well-known joke about an economist who arrives home after a night's heavy drinking to discover that he has lost his house keys. To the surprise of his friend, he immediately proceeds to the nearest streetlight and starts searching for his keys under its illuminating light. Being asked why he would search there, he replies: "Because obviously in the darkness this is the place where I'd be most likely to find them!" This willingness to look for evidence in unexpected places is a very useful quality in economists, even if it may occasionally appear a little odd to outsiders, as long as the researcher is careful to generalise his results only insofar as this is appropriate. Unfortunately, there are a number of reasons why results gleaned from public governance problems that are easy to examine may be limited to those industries. Some of these will only become apparent after the theory has been fully discussed, but one obvious reason can be explained here.

In the next chapter, there is a section devoted to alignment mechanisms, by which I mean the mechanisms that tend to push the result that occurs in reality towards the theoretical optimum. Such mechanisms are extremely useful, because they allow for the application of the traditional method of economics, the combination of positive and normative science in a single model. The three alignment mechanisms that will be discussed are evolution, competition and rational design. In other words, we will consider in what sense governments and governance choices made by governments can be said to evolve or compete, and how likely they are to be designed optimally, if at all. Upon brief reflection, it should be obvious to the reader that all three of these mechanisms work much better in a setting with many different points of comparison over a wide range, such as the prison, drinking water and garbage collection studies cited above, than in a setting with less scope for benchmarking, such as the governance of the foreign service¹¹, of the internet¹² or the governance structure of the railway industry.

This is at the same time an issue of theory and of methodology. As soon as we add alignment mechanisms to the model, the number of instances where the model is to apply becomes an important variable when it comes to the likelihood that the model's predictions will in fact be observed. At the same time, there are great methodological problems involved in studying a problem that has only a small number of observations. For example, when Ménard & Saussier (2002) studied the supply of drinking water in France, they had one observa-

¹¹ Cf. Williamson (1999).

¹² Cf. Helfer & Dinwoodie (2001) and Wilske & Schiller (1997) for an introduction to some of the problems in this field.

tion for every French *département*, for a total of around 90 observations.¹³ Having so many observations was not only convenient for their analysis; it was also convenient for the politicians and civil servants in charge of deciding how the water supply should be organised in a given *département*. Every *département* could look to all the others in order to discover best practices. If a certain *département* should be particularly inefficient, this fact was sure to be discovered sooner or later, leading to pressure – through exit and voice¹⁴ – from its inhabitants. It follows that the average governance structure in the sector is highly unlikely to be less efficient than it would have been if the industry had been organised nationally, when there would have been much less scope for comparison. For this reason alone, it is impossible to generalise results obtained in settings with ample scope for comparison to all settings. In fact, to the extent that researchers are biased towards studying empirical cases with fewer methodological difficulties, it is likely that their work will overstate the explanatory power of the model. The present thesis will add to the literature by exploring this exact lacuna: a small-n public governance problem.

Within the set of such problems, the choice of case study is to some extent arbitrary. It is convenient if the case involves as few as possible technological difficulties, which would only tend to complicate the analysis for a researcher who will tend to be ill equipped to deal with them. Another factor to consider is history. If the activity that is the subject of the study has a long history, there is the opportunity of learning effects, meaning that the outcome is more likely to be efficient. Also, to the extent that there are path dependencies in play, it will be easier to filter them out if the study covers a longer period of time.¹⁵ For these reasons, this thesis will study an “aspect of reality” taken from the railway industry: 185 years of railway history in Western Europe to be exact.

The choice for railways as a case study is also useful for less mundane reasons. After three EC railway packages¹⁶, national law in this field is still very much in flux, with Member States uncertain as to the best approach to market ordering, within the range still permitted by European law. The same goes for similar markets, like many utility markets where a number of companies are aggressively consolidating across the Common Market, with the occasional

¹³ For a survey of similar studies, cf. Pérard (2009).

¹⁴ Hirschman (1970).

¹⁵ Cf. chapter 4.2, below.

¹⁶ The first railway package, agreed on 26 February 2001, consists of Directives 2001/12/EC, 2001/13/EC and 2001/14/EC, OJ L 75, dd. 15.3.2001, p. 1, 26 and 29, respectively. The second railway package consists of Directives 2004/49/EC, 2004/50/EC and 2004/51/EC, OJ L 164, dd. 30.4.2004, with corrigenda in OJ L 220, dd 21.6.2004. The third railway package, finally, consists of Regulations 1370/2007, 1371/2007 and 1372/2007, as well as the Directives 2007/58/EC and 2007/59/EC, OJ L 315, dd. 3.12.2007.

encouragement or discomfort of Member State governments. While the present thesis does not aim to give concrete solutions to these questions, its results will likely contribute to a careful understanding of the issues involved and the degree to which the new institutional economics can ultimately contribute to finding such solutions.

In the remainder of this chapter, we will have occasion to formulate and consider the research questions more carefully. First, however, it is necessary to take up some general questions regarding the nature of economic research generally, and in the context of the discipline of law & economics in particular.

1.2 Preliminary Remarks

1.2.1 Law and Economics

Traditionally, the study of law is divided in the study of *lex lata* and the study of *lex ferenda*, the study of the law as it is, and the study of the law as it should be.¹⁷ *Lex lata* is a system unto itself, with its own style of arguing and its own methodology, seemingly unconnected to the rest of academia¹⁸, although it is philosophically connected to descriptive questions posed by social scientists, such as: How did the law come to be as it is? *Lex ferenda*, on the other hand concerns normative questions posed by, for example, economists, but also philosophers, sociologists, and psychologists. Many scholars working in these latter fields do not even work in law faculties, nor do they necessarily consider themselves a part of the legal discourse. That is unfortunate, because their work will have to fit within a legal frame of reference in order to be applied.

To see the problem, consider Volokh (2007), a paper which we will return to in section 2.5, below. This author has a background in law as well as in economics, and at the time of writing he was affiliated with the Georgetown University Law Center. His work consists of an extension of the Hart, Shleifer & Vishny (1997) model, and like that earlier paper Volokh applies his work to the US prison industry. In fact, he begins with a detailed examination of US jurisprudence on the permissibility of privatising prisons, followed by an examination of the earlier literature which also discusses a wealth of practical examples. Having done that, the author proceeds to develop his own model using plenty of intimidating-looking algebra, before returning only very briefly to his practical problem. To be sure, he derives some interesting conclusions, but nevertheless the article appears disjointed, with too wide a gap still looming between the law and the economics. Apparently even for someone who is trained in both fields, it is difficult to make the law and the economics fit together.

To avoid this problem, this thesis will do things differently in a number of ways. It will reduce the degree of abstraction of the economic modelling as much as possible. The literature will be discussed and analysed verbally, and where results are still presented in the language of mathematics or formal logic, they are at least not derived that way. That is not to

¹⁷ Cf. Thirlway (2001).

¹⁸ A fact that leads to recurring debates about whether or not the study of the law should be considered a “science”. Cf. for example Stolker (2002) and De Geest (2004).

discount the usefulness of mathematics in economics¹⁹, far from it, but rather to note that the current level of abstraction and detail of economic modelling often far exceeds that which can usefully be applied to practical problems.²⁰ In order to assure practical applicability, theory will be illustrated with actual jurisprudence whenever possible, instead of using simplified hypotheticals as economists frequently do.²¹ Finally, economic models will be cast in terms of legal concepts as much as possible.

This last suggestion merits some further consideration. Ruiters has been suggesting for some time now that it should be possible and helpful to use concepts taken from legal theory in order to strengthen institutional economic analysis.²² Unfortunately, he has never been able to articulate exactly how that would work. Of course, to the extent that legal theory helps form a general understanding of the law, it is certainly useful. After all, for many institutional economists the law is what the fruit fly is for many a biologist: the object of all his efforts.²³ It is for that reason that we will briefly consider Wesley Newcomb Hohfeld's rights framework in section 2.1, for example. Notions such as "rights" and "entitlements"²⁴ are essential to any understanding of institutions, since they are the genetic material that defines these institutions. This is particularly true for those readers who do not have the benefit of a background in law, and who therefore miss a certain implicit understand of such essential legal concepts.

However in addition to such legal theory, we will find it useful to connect our economic variables with concepts taken from positive law, i.e. the law as it is used by courts and jurists every day. For instance, we will connect the economic notion of a transaction not only with the contract that might govern it²⁵, but also with the common law problem of what should or should not count as "consideration", i.e. that which is given in return for a legally binding promise.²⁶ We will emphasise the difference between an organisation that is incorporated, making it a separate person under the law, and an organisation that is nothing more than a partnership governed by a contract. In most cases, introducing legal concepts to clarify or indeed define economic variables not only makes the model more practicable in general, it

¹⁹ However, cf. Leontief (1971), who used his 1970 presidential address at the annual meeting of the American Economic Association to criticise the direction modern economics was taking in this respect.

²⁰ Cf. Lipsey (2001), particularly his observation 1, p. 172-174.

²¹ In chapter 2, the initial theoretical chapter, the illustrations will deliberately not be taken from the railway industry, which will only be discussed in chapter 3.

²² Originally in Ruiters (1994), most recently in Ruiters (2008a).

²³ There are also institutional economists who study institutions that are not law, eg. North (1990).

²⁴ Calabresi & Melamed (1972). Cf. Ruiters (2008b).

²⁵ Cf. Ruiters (2005), p. 290-291.

²⁶ Cf. Halsbury's Laws of England, 2nd ed., ed. Hailsham (1932), vol. 7, p. 136.

also has the specific benefit of making the variables more readily observable. While one may doubt whether a certain entity should be considered an “organisation”, there is rarely any doubt as to whether it is a “corporation”. As a result, even when defining an organisation as a corporation is too drastic, noting that all corporations are, in any event, organisations already makes the latter concept clearer.

1.2.2 *The Nature of Economics*

Regarding the nature of economic science, there are two things that need to be established if the reader is to follow the reasoning in the remainder of this thesis. On the one hand, there is a problem of economic methodology, concerning the nature of truth in economics, the purpose of models and their application, where I would like to note my adherence to the approach of Robbins (1932), as well as the Austrian School. On the other hand, there is the issue of rationality, which is forever a problem for economists generally and institutional economists in particular.

In 1932, Lionel Robbins wrote his *Essay on the Nature & Significance of Economic Science*²⁷, where he made two statements about the nature of economics that we will take to heart in this thesis. Regarding the subject matter of economics, he wrote that “Economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses.”²⁸ As for economic method, his key proposition was that “the propositions of Economics (...) are deductions from simple assumptions reflecting very elementary facts of general experience.”²⁹ It goes without saying, of course, that these propositions are only interesting to the extent that they are not obvious.³⁰

The most important consequence of this last proposition, which is echoed in the work of many economists of the Austrian school, including Knight³¹ and Von Mises³², is that there is a difference in economics between the question of whether a theory is true and the question of whether it has any great explanatory power, i.e. the extent to which variation in the inde-

²⁷ Cf. Masini (2009).

²⁸ Robbins (1932), p. 16. Cf. Backhouse & Medema (2009).

²⁹ Robbins (1932), p. 104. For a general critique, cf. Hands (2009) and Lipsey (2009). For a critique of these so-called “indisputable facts” in particular, cf. Sugden (2009).

³⁰ Lipsey (2001), p. 169.

³¹ Eg. Knight (1940).

³² Eg. Von Mises (1949), p. 862: “The ultimate yardstick of an economic theorem’s correctness or incorrectness is solely reason unaided by experience.”

pendent variables identified by the model “explains” the variation in the dependent variable.³³ Because of the complexity of human society, the social sciences always provide many different explanations for the relationship between any two variables. While many of these hypotheses might be true, even when they contradict each other, inevitably some will have greater explanatory power than others.

As an example, consider the idea of Ricardian Equivalence, the idea that “a tax cut financed by issuing government debt may fail to stimulate private consumption because consumers discount the future tax burden required to serve and repay the debt, in the belief that this will cancel out the current tax alleviation.”³⁴ This hypothesis was first proposed by David Ricardo in 1820, before immediately rejecting it, as most economists have done since³⁵, on the grounds that it runs so much counter to our everyday experience as to how consumers respond to a tax cut, not to mention the bulk of statistical evidence³⁶, that it must be rejected as a viable economic theory. At the same time, it must be admitted that this hypothesis follows unavoidably from the assumption of perfect rationality, when combined with certain facts about discounting and government budgeting. Instead of rejecting it altogether, it is probably more accurate to say that the Ricardian Equivalence Hypothesis is in some sense true, because it is derived by sound logic from sound assumptions³⁷, but that it has little explanatory power, because the considerations that it emphasises play only an infinitesimal role in consumers’ decision making process.³⁸

Another, perhaps more obvious way to tackle the problem raised by the Ricardian equivalence hypothesis is to point out that its key assumption may be incorrect. Is it really reasonable to impute economic actors with the kind of perfect rationality that is necessary in order to make this conclusion follow from the premises? And, more importantly, is it useful to do so?

³³ Cf. a discussion of R^2 in any statistics textbook. Eg. Kleinbaum et.al. (2008), p. 96-99.

³⁴ Briotti (2005), p. 5.

³⁵ However, cf. Barro (1974).

³⁶ Cf. Briotti (2005) for a survey.

³⁷ Which explains why many textbooks on macroeconomics do discuss it, eg. Heijdra & Van der Ploeg (2000), chapter 6, LeRoy Miller & VanHoose (2004), p. 186-191, and Abel & Bernanke (2005), p. 124-125.

³⁸ N.B. This is not the same as the approach advocated by Friedman (1952). In his view, positive science is judged only by its ability to yield “valid and meaningful predictions about phenomena not yet observed” (p. 7), making the underlying assumptions irrelevant as long as they work. The Friedman approach is flawed because it fails to appreciate that the capacity of science to illuminate social phenomena requires more than the production of correct testable predictions. The manner at which these predictions are arrived at is equally important.

The substantive merits of the rationality assumption depend very much on the context of the problem³⁹, for example on whether the theory is used to analyse specific individuals or average tendencies among larger groups. Such considerations aside, one could also argue that a degree of rationality follows from the very definition of economics. Lionel Robbins, for example, argued that his definition of economics, as quoted above, implied a minimum level of purposiveness.⁴⁰ That is not the same as saying that economic models should always be based on perfect rationality. They should, however, generally explain an actor's choice of means by reference to the ends that the actor is seeking to achieve. Abandoning this definition would make economics the study of all human behaviour, which is a task already claimed by psychologists.⁴¹

There is also a key pragmatic reason to favour rationality in some form: it allows the theory to be simultaneously descriptive and normative. People are assumed to behave in the manner the theory recommends. The benefit of this assumption goes well beyond the fact that recommendations are usually more closely determined – outside game theory there is usually only one optimal course of action, even if we are not always sure what it is – to the interrelationships between recommendation and description. Take the contracting problem, for instance, which is the focus of all NIE theory. Here, the theory makes a recommendation to one of the contracting parties by assuming the other party will (also) behave in accordance with the theory.⁴² No recommendation to the one is possible without a descriptive theory for the other. It follows that assuming a degree of rationality is essential in order to allow the economist to wrap it all up in a nice strong model that makes clear predictions with a minimum number of variables.

One could argue that this approach is the defining characteristic of economics, setting it apart from its sister social sciences. It is the secret to the science's success, allowing it to produce policy recommendations that make up for in practicality and accessibility what they lack in nuance. The discipline has been able to cast its shadow well beyond its “core constitu-

³⁹ Cf. at various instances in chapters 2 and 3, below, where we will discuss the merits of the assumption of rationality in NIE generally and in NIE's application to the railway sector in particular, respectively.

⁴⁰ Robbins (1932), p. 93.

⁴¹ Cf. <http://www.merriam-webster.com/dictionary/psychology>, definition 1: “[Psychology is] the science of mind and behavior.” (Last visited on March 17, 2009.)

⁴² Cf. Friedman (1952), p. 5-7.

ency” of trade and industry, making itself relevant to almost all aspects of human society.⁴³ The present thesis, of course, is also an example of such “economic imperialism”.⁴⁴

In order to preserve this competitive advantage, strict adherence to perfect rationality is not necessary. As long as actors are assumed to be optimising their utility, however defined and determined, the only requirement is that the constraints under which they operate are sufficiently clearly defined to fully determine the outcome. Stigler (1961), for example, models the search for information in terms of the expected marginal benefit of additional information relative to the marginal cost of obtaining it. The result is somewhat awkward, and in need of further development, but it was a useful first step in the direction of endogenising the information actors possess without abandoning the economic mode of reasoning.⁴⁵

With this in mind, we can consider some of the versions of bounded rationality that have been proposed over the years. The idea of satisficing⁴⁶, for example, is enormously underdetermined by economics standards, at least in its basic formulation. Even if an adequate account is given as to the origins of the actor’s “aspiration level” – a task that is in any event more descriptive than normative – the fact remains that the model provides no way of predicting which of the many solutions that are “good enough” will be chosen. When Simon notes that “the *first* satisfactory alternative has precise meaning”⁴⁷, he is right, of course, but only to the extent that the sequence in which the alternatives are presented to the actor is also modelled. This has led a number of economists, including Simon himself, to investigate “rationality as process”⁴⁸, meaning models of searching behaviour. However, it is highly doubtful whether the benefit of such models in terms of explanatory power outweighs their costs in terms of additional complexity in most circumstances.⁴⁹

A more lengthy survey of the literature on bounded rationality, satisficing or, for that matter, mixed scanning⁵⁰ is clearly outside the scope of this introductory chapter. Nor is it

⁴³ Most notably, in 1992, Gary Becker received the Alfred Nobel memorial prize in economics for “having extended the domain of microeconomic analysis to a wide range of human behaviour and interaction, including nonmarket behaviour”. Cf. Becker (1981a, 1981b, 1993), Becker & Murphy (1988) and Becker & Posner (1993). Another fascinating example is Barro & McCleary (2011), which explores the economics of beatification and canonisation in the Roman Catholic church.

⁴⁴ Cf. Lazear (2000) and Medema (2009).

⁴⁵ The article has spawned a line of research called Search Theory. Eg. Ellison & Fisher Ellison (2005). Its most useful application has been to modelling the labour market, along the lines of Stigler (1962).

⁴⁶ Cf. Simon (1947), p. 118-120 and Simon (1955).

⁴⁷ Simon (1955), p. 111. Emphasis in the original.

⁴⁸ Simon (1978).

⁴⁹ Cf. for example, the approaches collected in Rubinstein (1998).

⁵⁰ Cf. Etzioni (1967).

necessary in order to solve the problem at hand. Later chapters will occasionally return to this question, as part of a critique of specific models, or as part of their application to the railway sector. As a general starting point, however, there is no need to say much more than that bounded rationality means behaviour that is “intendedly rational, but limitedly so.”⁵¹ All versions of NIE take into account that the gathering, analysis and use of information is costly, and that this results in a situation where people’s behaviour differs in predictable ways from how it would be under perfect rationality. Within the framework of this body of theory, there is little benefit to following Simon’s suggestion and attempting to incorporate detailed psychological models about learning and information processing.⁵² Doing differently would produce a result similar to an amphibious car in motoring: a car that does not drive very well combined with a boat that does not sail very well adding up to a vehicle that is not very useful for anything.⁵³

What is left is a mixture between studying transaction costs as a *Ding an sich*, and studying transaction costs as they are perceived by economic actors.⁵⁴ From any theory within NIE one would certainly be entitled to expect some outline of the circumstances in which significant transaction costs should be expected to occur. That is why Oliver Williamson, for example, devotes so much attention to the concept of asset specificity.⁵⁵ When it comes to judging the explanatory power of the theory, transaction costs as they are perceived by actors have to take centre stage.⁵⁶ At that stage, the central questions are whether the actors involved perceive the transaction costs of their contracting problem correctly, and whether the governance structures they set up in order to minimise the perceived transaction costs are constructed in accordance with the predictions of the model.

Given these preliminary remarks, the next section proceeds with a more systematic discussion of the research questions.

⁵¹ Simon (1947), p. 88.

⁵² Cf. Simon (1955), p. 99-100.

⁵³ Cf. Robinson (1933), p. 1, defining analytical economics as “a box of tools”.

⁵⁴ With apologies to Kant (1781).

⁵⁵ Cf. section 2.4, below.

⁵⁶ Cf. North (1990), p. 22-24. In North (2005), he writes in the preface: “Economic change (...) is for the most part a deliberate process shaped by the perceptions of the actors about the consequences of their actions. The perceptions come from the beliefs of players – the theories they have about the consequences of their actions – beliefs that are typically blended with their preferences.”

1.3 Research Questions and Outline

As explained above, the central research question of this thesis is very clear: *What is the value of the framework of the New Institutional Economics for questions of government ownership and control of certain industries?* However, inevitably we will find ourselves also studying what the various versions of NIE, not to mention the most obvious alternative frameworks, say about the choices that have been made. One could say that that is the secondary research question: can we identify choices that were clearly inefficient, in the sense that they cannot be understood with the economic models that we have at our disposal?⁵⁷ In so doing, however, it is important not to forget which of these questions is the more important one. History is fixed, sometimes literally written in stone, and there is little to be gained by second guessing the choices from the past. A more careful understanding of the theory, on the other hand, makes it more likely that we will be able to avoid mistakes in the future.

The steps that are necessary in order to carry out this task are simple: 1) Develop the model, 2) apply it to the history of the railways, 3) develop the benchmark models, 4) apply them to the same problem, and 5) observe the relative merit of NIE. These tasks will be divided over the chapters of the book as follows:

In chapter 2, we will consider what the new institutional economics is, and whether it is true. As explained in the previous section, that means that we will construct one or more versions of the model that are derived by sound logic from sensible assumptions in such a way that it is most likely to be capable of doing what it was designed it to do: to study public sector privatisation by analogy to private sector outsourcing. Specifically, we will construct two versions, which we will call Transaction Cost Economics (TCE) and GHM, after its creators Grossman, Hart and Moore. Initially, the chapter will cover some definitional issues, such as the nature of law and the nature of transaction costs, before moving on to the two main lines of research in NIE: Oliver Williamson's Transaction Cost Economics and Hart and Moore's Incomplete Contracting Theory. Each will be evaluated on its own merits, as well as in comparison to the other. Finally, we will consider which elements make up the core of each version, the chain of inference that runs from its starting assumptions to its ultimate hypothe-

⁵⁷ Note that this is a very modest way of framing the question. We could envisage increasing its ambition by removing the double negative: Were the choices made by decision makers efficient, meaning that they comport with the economic models that we have at our disposal? However, neither our modeling nor our grasp of the facts is sufficient to make such an ambitious research question viable.

ses. For each version – and even more importantly for the core that they both share – the goal is to identify a set of elements that we would expect to observe if the model is to have any significant explanatory power.

Chapter 3 will discuss the history of the railway industry in Western Europe, specifically in the Netherlands, Germany, France, Belgium and the United Kingdom. Between them, these countries started out all over the public-private spectrum, ranging from the extreme reliance on the private sector in the UK and the Netherlands, to the completely government owned state railway in Belgium. Subsequently, however, they all seem to follow the same general wave pattern, converging on consolidated private ownership in the second half of the 19th century, on public ownership several decades later, and more recently back to privatisation. It is particularly this pattern that the models will be expected to explain.

In telling this story, the intention is to construct an integrated analytical narrative.⁵⁸ Rather than starting with a purportedly objective historic account, followed by a separate analysis, the analysis will be intermingled with the discussion of historical facts as much as possible.⁵⁹ The advantage of this approach is not only to avoid tedious repetition, but also to clearly show why certain facts are mentioned but not others. If case studies are to be of any use, their application has to be as transparent as possible. An artificial distinction between history and economics is not the way to achieve this. As a check on the inherently subjective nature of the case study approach, each of the four sections of chapter 3 will conclude by reviewing the basic elements identified in the end of chapter 2.

Having thus asked and answered whether the new institutional economics is “true”, and whether it has any real explanatory power in the context of the railway sector, chapter 4 moves on to consider the benchmark against which it will be judged. This benchmark consists of the two most likely alternative theories that one would normally use to explain government action and inaction from an economic point of view: public choice theory and path dependence. If the approach pioneered by Coase is to be useful here, it has to have a similar degree of explanatory power to these two alternative economic approaches. In order to allow such an evaluation, chapter 4 will briefly discuss each theory, as well as its application to the railway sector, before comparing the relative explanatory power of each. Needless to say, it would be unwise to compare NIE with non-economic theories and models, since to do so would seri-

⁵⁸ Cf. Alston (2008).

⁵⁹ Cf. the approach in Holterman (2009).

ously tax the ability of any one scholar, and would run the risk of understating the value of NIE. After all, compared with a theory that is designed “to say a lot about a little”, an economic theory will always look clumsy and unrealistic. To the extent that there are other economic theories that might shed some light on the problem of semi-public institutions, they are excluded in order to keep the analysis manageable, and because the two approaches that we will discuss should suffice in order to establish a benchmark.

Finally, and inevitably, chapter 5 concludes. That is to say, it contains a review of the main findings of this thesis with regard to the new institutional economics. The reader should be careful not to rush to judgement regarding any findings that this dissertation might seem to contain concerning the railway industry. Not only is this thesis more focused on the decision making process when changes are contemplated than on the industry per se, but it is also true that any results obtained in this dissertation about the railways are only valid within the theoretical framework in which they are cast. Ignoring the limitations of the new institutional economics, and of economics generally, is at the reader’s peril.⁶⁰ Making normative statements about the optimal governance structures for the railway industry in any detail is outside the scope of this thesis.

⁶⁰ To quote the famous historian Tony Judt: “For the purposes of mental emancipation this evening, I propose that we take a minute to study the history of a prejudice: the universal contemporary resort to “economism”, the invocation of economics in all discussions of public affairs.” Judt (2009).

2. Theory

The previous chapter already gave a first impression of the approach taken in this chapter: what the fruit fly is for the biologist, institutions, and specifically laws, are for the institutional economist. From this it follows that we cannot develop a suitable institutional economic model of the optimal governance structures governing the relationship between the state and the railways until we have developed a basic understanding of what these governance structures – or institutions as one might also call them – are. Since the notion of institutions generally is often closely tied up with the specific author’s model, we will start by discussing only the law. Taking insights from legal theory, it should be possible to make the concept sufficiently clear for further use, without drawing on institutional economic models that may or may not be helpful. Institutions that are not based in law can then be discussed later in the chapter.

In addition to providing a basic set of ideas about law, the first section will also explore the notion of incompleteness, something that will prove essential time and again throughout this thesis. Particularly noteworthy in this respect is the difference between the jurist’s notion of incompleteness and the economist’s.

Subsequent sections will develop the NIE framework one step at a time. In section 2.2, we will look at the Coase Theorem, which says that “in a world of zero transaction costs, the allocation of resources will be efficient, and invariant with respect to legal rules of liability.”⁶¹ In other words, no matter how you distribute the entitlements⁶² initially, in a world without transaction costs they will always be traded until they end up in the same hands, and the distribution of entitlements that is produced by this process of trading is efficient. Section 2.2 will explore the implications of this statement, as well as certain exceptions to the rule.

The most important premise of the Coase Theorem is, of course, the assumption of zero transaction costs. Therefore it is only appropriate that transaction costs should have a section of their own. In section 2.3, we will follow the development of this concept from Commons and Coase to Williamson and other contemporary authors. The models that these contemporary authors have created are the subject of sections 2.4 and 2.5. First Williamson’s

⁶¹ Zerbe (1980), p. 84. The author adds a qualification to his definition (“income effects aside”), which has been omitted here, but will be discussed below. The various definitions that have been proposed in the last forty years are discussed in Medema & Zerbe (1999), and in Steenge (2004).

⁶² Cf. Calabresi & Melamed (1972).

Transaction Cost Economics and then the models developed in various configurations by Grossman, Hart and Moore.

Afterwards, in section 2.6 we will discuss three so-called “alignment mechanisms”, that is to say mechanisms that should – hopefully – ensure that the alignment between the transaction and its governance structure is as efficient as it can be. Without such mechanisms the question of optimal alignment would only be of theoretical value; the alignment mechanisms are the link between the models in their normative incarnations and the models in their descriptive versions.

Finally, in section 2.7 we will attempt to discover some kind of common thread in the models discussed in this chapter, a kind of synthesis one might say. Also, we will consider which elements we would expect to see in the case study if the framework is to be validated. What is it that we should be on the lookout for? Which elements are, together, necessary and sufficient to prove the framework applicable to the problem of railway regulation?

2.1 The Law and its Incompleteness

2.1.1 *The Nature of Law*

The purpose of the law is to establish the rights and obligations of actors in any given set of circumstances, to the extent that such rights and obligations are enforced by the state. As such, the making of law is always a forward looking act, whereby those with the power⁶³ to create legal rights and obligations decide which rights and obligations would be most appropriate in any possible future situation. At the same time, the enforcement of the law is an act of looking backwards, the act of deciding which circumstances actually did occur, and which legal rules apply to that situation. The remainder of this section will formalise these insights somewhat.

A convenient place to begin is with Wesley Necomb Hohfeld, who in the early 20th century developed a framework to clear up the notions of “rights” and “obligations” in the law.⁶⁴ In this framework, rights can be either claims or privileges, while the opposite is either a duty or a no-claim. Obviously, these concepts are connected in a variety of ways. For example, my duty not to throw a rock through my neighbour’s window translates into my neighbour’s claim that I should not do so. Similarly, the rental contract of my apartment implies the privilege of access to a house that belongs to my landlord, which is equivalent to the absence of a claim that my landlord would otherwise have for me to stay out of his house.⁶⁵ For the present purposes, most of these distinctions and correlatives are unimportant. Throughout the remainder of this thesis, a claim of Y on X, a privilege of X against Y, a duty of X against Y and a no-claim of Y against X are all referred to simply as an obligation of X to Y, or the absence thereof.

The next question is the nature of such an obligation. Under Dutch law, obligations can consist of an obligation to do, an obligation to refrain from doing, or an obligation to

⁶³ Ruiters (2008) defines a power as “a legal position of a person vis-à-vis other persons having a legal relation that can, under the applicable legal rules, be changed by a fact under the volitional control of the first-mentioned person.” This concept is distinct from the sociological notion of power, as well as from the deeper question whether it is possible to imagine law in the absence of enforcement, through physical force, by a sovereign. Cf. also Hart (1972) and Weber (1920), Part Three, Chapter VI, § 1.

⁶⁴ Hohfeld (1913). There is a second framework concerning powers, liabilities, disabilities and immunities, but those are not important for present purposes. Cf. Ruiters (1993), p131-159, and Ruiters (2008).

⁶⁵ In the Hohfeld framework, this is called a no-claim. Note that in many cases such a “may” can be related only to a hypothetical “must”. Everybody has the privilege to use the pavement, for example. The corresponding possibility that one might be obliged to stay off it is only hypothetical.

give.⁶⁶ In what follows, the collection of acts and omissions that are required by law of a certain person at a certain time under certain circumstances is called the action set, i.e. the set of actions that are required.

Finally, we need to clarify what we mean by “circumstances”.⁶⁷ The more technical term for this concept is “state of the world”⁶⁸, a term that we will use throughout this thesis. Thinking about the future in this way implies a discontinuous view of the probability distribution of future events: one could imagine describing the future through an infinite series of yes or no questions, such as “Did the sun shine?” or “Did he make the investment?”. For n such questions, the number of different states of the world would be 2^n , each of which has a distinct probability of occurring.⁶⁹ It follows that the set of all possible future states of the world can be ordered⁷⁰ through the probability of each element occurring, i.e. for $s \in S, \forall i: [P(s_i) \geq P(s_{i+1})]$, where S is the set of all possible states of the world.⁷¹ These states of the world are jointly exhaustive, and, ideally, independently observable.⁷²

It follows from this formalisation that the law is a mechanism for assigning action sets to actors that describe the obligations they have in each future state of the world. For example, when the justice of the peace says “by the power vested in me by the state of X, I now declare you husband and wife”, the legal reality of that statement is that a whole body of standardised obligations come to apply to the legal situation between the spouses and between the spouses

⁶⁶ Art. 3:296(1) Dutch Civil Code. The same distinction can be found in art. 1136-1145 of the French Code Civil. § 241 of the German Civil code distinguishes between doing and omitting (“Die Leistung kann auch in einem Unterlassen bestehen.”), where doing obviously includes giving.

⁶⁷ Cf. Alchourrón & Bulygin (1971), p. 7-21, whose model is more advanced than the one given here. Their work refers not only to the concepts of state space (“Universe of Discourse”) and action set (“Universe of Actions”), but also to the set of all properties that a state of the world does or does not exhibit (“Universe of Properties”), and the set of all possible combinations of those properties (“Universe of Cases”). The latter concepts are not necessary for our purposes.

⁶⁸ Cf. Hart & Moore (1990), p. 1126.

⁶⁹ In reality, the number of such questions would necessarily have to be enormous, in order to achieve “a description of the world so complete that, if true and known, the consequences of every action would be known”. (Savage, 1954) The notion of a state of the world is related to the Many Worlds Interpretation of Quantum mechanics, cf. Everett (1957) and DeWitt (1970), an idea more commonly referred to as “parallel universes” or “alternate universes”.

⁷⁰ “A well-ordered set X is a linearly ordered set for which every nonempty subset has a minimal element.” Roitman (1990), p. 9, definition 23.

⁷¹ This notation and ordering will be maintained throughout the thesis.

⁷² Cf. Chapter 2.1.5, below. Traditionally, it is also required that “the states must be defined in a way that their likely realization not be affected by the decision maker’s choice of action” (Karni (2005), p. 3). However, this does not mean that the choices of other actors cannot be taken into account in the description of a state of the world, nor does it mean that actions that actions that lie between the present and the moment under consideration must be ignored. For example, in a classic incomplete contracting setup, the decision is made at $t=0$ whether or not to invest at $t=1$ depending on the actor’s beliefs about the probability distribution of states of the world at $t=2$, a distribution which obviously depends on whether or not the investment has been made. There is no reason why such a problem could not be analysed in terms of states of the world.

and the state. They become subject to different tax rules in many states of the world, they have the obligation to take care of each other in certain states of the world, etc.⁷³

Within the law, it is useful to distinguish between laws of general application and law that is not of general application, the latter including contracts, but also decrees⁷⁴ and decisions.⁷⁵ In what follows, a distinction is usually made between law that finds its origin in a bargain and law that originates in a constitutionally legitimated law maker. For convenience, the former will be called contracts, and the latter laws or statutes. While this may seem confusing, it should be remembered that those decrees and decisions that we will encounter in this thesis will usually either invite a bargaining process⁷⁶ or be the legal form of a bargain between the state and a citizen or company. For this reason, describing all sources of law other than laws of general application as contracts is not unreasonable.⁷⁷

In most jurisdictions, positive law recognises the power of every adult to create legal relations between himself and one or more others in the form of a contract.⁷⁸ The freedom to contract, after all, is a key cornerstone of our capitalist system⁷⁹, and this freedom presupposes the legal power to contract. In continental Europe, the legal treatment of contract formation emphasises the will of the parties, and the expression thereof. A contract is formed when parties will a contract on identical terms, and have expressed that will.⁸⁰ Whether there

⁷³ These two examples, in turn, are shorthands for more complex collections of obligations. The whole income tax code is ultimately an obligation of each citizen to give a sum of money to the government, the only difficulty being the calculation of the size of that sum. It follows that the spouses' married status only matters in those states of the world where they have taxable income.

The obligation to take care of one's spouse (eg. art. 1:81 Dutch Civil Code) only translates into an actual obligation in those states of the world where there is a need for such assistance, while the exact nature of the obligation depends on the nature of the need, and is consequently different in every state of the world.

Note that authors such as Raz (1972) argue the opposite only because they analyse the law with a different objective in mind. The choice for legal positivism here depends on its context in a thesis about law and economics. Raz is obviously right when he argues that individual laws, or individual legal rules, will often not directly impose legal obligations.

A detailed discussion of the transformation of general norms into action sets can be found in Alchourrón & Bulygin (1971), p. 13-21.

⁷⁴ Eg. art. 13 of the Constitution of the French Republic.

⁷⁵ Eg. art. 288 TFEU.

⁷⁶ Cf. chapter 2.4.3 below, which discusses the tender process for cable television in Oakland in the late 1960s.

⁷⁷ In fact, if one were so inclined, one could describe all law as contract. Cf. Wittman (1999).

⁷⁸ Eg. art. 3:32 Dutch Civil Code. Power is meant here in the Hohfeldian sense of the legal ability to create or change the legal relation between two or more persons. Cf. Ruiter (2008).

⁷⁹ Cf. Adam Smith (1776), Book 1, chapter 10. However, cf. the discussion of the history of exclusion clauses by Lord Denning MR in *George Mitchell (Chesterhall) Ltd v. Finney Lock Seeds Ltd* [1983] QB 284 at 297.

⁸⁰ Eg. art. 3:33 and art. 6:217 Dutch Civil Code or art. 1108 of the Code Civil. In some circumstances, the law also recognises the existence of a contract where the expression was defective, eg. art. 3:35 Dutch Civil Code or

is an actual trade is irrelevant; civil law also recognises contracts to give.⁸¹ In contrast, the common law tradition emphasises the nature of contract as a bargain. As a result, contracts are not enforced unless there is consideration.⁸² In recent decades, several common law jurisdictions have added the requirement that parties should have had the intention to create legal relations⁸³, thus reducing somewhat the difference between common law and civil law on this point. While the contracts that are discussed in this thesis will always be for consideration, it is important to be careful to construct a model that are not biased towards any one legal system.⁸⁴

Throughout, it is important to remember the distinction between a contract and the transaction.⁸⁵ Consider the rule of § 929 (1) BGB, which states: “*Zur Übertragung des Eigentums an einer beweglichen Sache ist erforderlich, dass der Eigentümer die Sache dem Erwerber übergibt und beide darüber einig sind, dass das Eigentum übergehen soll.*” In English: “For the transfer of ownership in a chattel, it is required that the owner hands over the object and that both are agreed that ownership should be transferred.”⁸⁶ At one end of the spectrum, the parties to a spot contract agree on terms and transfer ownership at the same time. This is what happens in supermarkets all over the world every day.⁸⁷ On the other end of the spectrum, there are contracts that establish a legal relationship for an indefinite period of time, for example the contracts between supermarkets and their suppliers. While the existence of such a relationship alone is not enough to give the governance structure in question the nature of a hybrid⁸⁸, in practice many long-term contracts do establish shared control rights.⁸⁹

art. 1110 of the Code Civil. A similar rule is established by the English case of *Smith v. Hughes*, (1871) LR 6 QB 597.

⁸¹ Eg. art. 1102 Code Civil.

⁸² Consideration is defined in Halsbury’s *Laws of England* as “either some right, interest, profit, or benefit accruing to one party, or some forbearance, detriment, loss or responsibility given, suffered, or undertaken by the other at his [the promisor’s] request”, Halsbury’s *Laws of England* (2nd ed., ed. Hailsham (1932), vol. 7, 136).

⁸³ *Balfour v. Balfour* [1919] 2 KB 571, *Merritt v. Merritt* [1970] 2 All ER 760; [1970] 1 WLR 1211. However, cf. Restatement (second) of contracts § 21.

⁸⁴ Cf. Brousseau (2001).

⁸⁵ Cf. Ruiter (2005), p. 290-291.

⁸⁶ Translation by the author. A similar rule is found in art. 3:84 Dutch Civil code.

⁸⁷ In most jurisdictions, the act of the supermarket displaying the goods is considered an invitation to treat. The customer then makes an offer when he hands the goods over to the cashier, who then accepts the offer on behalf of the store owners by entering the sum into the register and handing the goods back to the customer. The leading UK case is *Pharmaceutical Society v Boots Cash Chemists*, [1953] 1 QB 401, the leading US case is the Minnesota case of *Lefkowitz v. Gt. Minneapolis Surplus Store*, 86 NW 2d 689 (1957), and in the Netherlands there is *Hofland v. Hennis*, HR 10 April 1981, NJ 1981, 532 m.nt. CJHB.

⁸⁸ Hodgson (2002). For a discussion of this article, cf. chapter 2.4.2, below.

⁸⁹ Cf. US Supreme Court, *American Needle v. National Football League et. al.*, 560 US ____ (2010), which considered the extent to which the NFL was a “single entity” for competition law purposes.

2.1.2 Contractual Incompleteness

Having established some basic facts about the nature of the law, we must now consider how it can be incomplete. As we will discover in the remainder of this thesis, it is particularly when the law is flawed that things get interesting. After all, when contracts are flawless and costless it is difficult to see why actors would want to create more intrusive institutions to govern their transactions. While the question will be phrased in terms of contractual incompleteness, as before the notion can also be applicable to legal relations that are defined by instruments other than contracts.

To begin with, consider the definition given in Church & Ware (2000), p. 73:

“A complete contract is one that will never need to be revised or changed and is enforceable. It specifies precisely what each party is to do in every possible circumstance and for every circumstance the corresponding distribution of the gains from trade. And regardless of the circumstances a court will be able to enforce the contract – it is capable of requiring compliance and imposing damages such that both parties to the contract will honor the terms of the contract. This type of contract would provide no opportunities for renegotiation or holdup since it would contain no gaps, or missing provisions. However circumstances unfolded, the contract would unambiguously govern the exchange.”

It appears that there are three kinds of incompleteness: incompleteness due to gaps, incompleteness due to lack of detail, and incompleteness due to the fact that the contract cannot be enforced by the courts. Following Ayres & Gertner (1992), we will call the first type of incompleteness “obligational incompleteness”, it is the subject of the next section. The second type we describe as insufficiently contingent on states of the world, and it will be discussed in section 2.1.4. Problems with enforceability – which are not discussed by Ayres & Gertner – are the topic of section 2.1.5.

Before continuing, however, it is important to consider the importance of extending the concept of contractual incompleteness beyond merely the case where the contract is silent. To begin with, it should be intuitively clear that treating a contract that says “in all states of the world, no matter what happens, we will exchange widget X for \$ Y cash” as perfectly complete would be overly formalistic. What is important, though, is to consider when a lack

of state contingency merely leads to inefficiency, and when it is fair to say a contract is incomplete on the grounds that it is insufficiently contingent on states of the world.

As we will see later on in this chapter, incompleteness matters because it opens the door to *ex post* opportunistic behaviour, and to *ex post* renegotiation. In other words: these two things should not occur where there is a perfectly complete contract. Opportunistic behaviour ensues not only when the contract is silent, but also when it does not distinguish between those states of the world where a party behaves opportunistically, and those where it does not.⁹⁰ *Ex post* renegotiation would normally be expected whenever there is an *ex post* inefficiency that can be cured through further contracting. Such renegotiation will be seen, later in this chapter, to have the potential to distort parties' *ex ante* incentives.⁹¹

Since some forms of insufficient state contingency can make such problems possible in much the same way that obligational incompleteness can, it is sensible to extend the definition of incompleteness to include the former, while bearing in mind the distinction.

2.1.3 Obligational Incompleteness

Returning to the basic concept of obligational incompleteness, we can say that this is the only kind of incompleteness recognised by contract law. It occurs when parties do not specify what actions or forbearances are expected in a certain state of the world, which is notably different from the possibility that they might specify that no actions or inactions are expected. The reason why parties might choose to do so vary: it can be a decision based on the marginal costs and benefits of a certain level of completeness, but it is also possible that the information needed to achieve a greater degree of completeness is unavailable at any reasonable cost.

⁹⁰ Cf. chapter 2.1.5, below, for a discussion of the problems courts might have to observe which state of the world has been realised, even when parties are able to do so at very little cost. Where this is the case, even a contract that is written in perfect detail should still be considered incomplete, since generally only distinctions that can be enforced by the courts matter. In some states of the world, for example, it might be preferable for the seller in the earlier example to produce widget Y instead of X. If the court cannot tell the difference between the two, the contract might as well say “widget X or Y, at the discretion of the producer”.

⁹¹ In those states of the world where the transfer of widget Y instead of X would be Pareto superior, we would expect parties to renegotiate, contracting for Y instead. This may or may not hurt the *ex ante* efficiency of the transaction. Cf. chapter 2.5, below, and Hart & Moore (1999). As is explained there, the only difference between the situation where the contract is silent on which widget should be produced and a standard requirement to produce widget X in all states of the world is the threat point or outside option for the bargaining process.

Bearing in mind that we defined the state space so that it is ordered by reference to the relative probability of each element, it is easy to imagine a simple decision model that explains the optimal level of completeness for any given contract. Given the convenient assumptions that both the benefits of specifying an action set for a certain state of the world, the costs of deciding the optimal action set, and the costs of describing the state of the world to which it is to apply are all constant for every state of the world, it follows that the expected benefit of specifying one additional contingency declines as the degree of completeness increases, while the costs remain the same. This means that the costs and benefits of specifying an action set can be equated at the margin, yielding the optimal level of completeness.⁹² For states of the world beyond that limit, the probability that they will occur is simply so small that the additional costs of writing the contract so that it takes care of that possibility is greater than the expected benefit of doing so. If we assume, instead, that the benefits of specifying an action set are unique to that state of the world, or at least uncorrelated with the probability of it occurring, it follows that the cut-off point would be less clear, since there would be some states of the world that were very unlikely, but with very high stakes, making it worthwhile to account for them in the contract.

It should be noted that there is only a small difference between this scenario, and the common observation that parties failed to make arrangements for a certain eventuality because they did not realise this might actually happen, i.e. because they did not realise the probability of that state of the world was nonzero.⁹³ As explained above⁹⁴, an economic model will always need to assume a certain degree of rationality, making it difficult to account for mistakes by parties.⁹⁵ If need be, however, this gap can be bridged by introducing some reasoning from information economics, which models the choices of actors when they decide how much information to gather.⁹⁶ This is akin to saying that parties make mistakes about certain improbable states of the world, because they have estimated that it is too costly to learn more about those states of the world, including learning more precisely what the exact

⁹² Formally: let B_i be the benefit of contracting for a certain state of the world, and let B_i be constant for all i . Let TC_i be the transaction costs of describing state of the world i , and bargaining for it, and let TC_i also be constant for all i . The optimal level of completeness is then given by the number of states of the world i for which $p_i B_i \geq TC_i$.

⁹³ Consider the title of the chapter in Bloembergen et.al. (2001) where this issue is discussed: Onvoorziene omstandigheden, i.e. unforeseen circumstances. p. 344-357. Cf. Eloffson (1996) and Restatement (second) of contracts, chapter 11.

⁹⁴ Chapter 1.2.

⁹⁵ Cf. the example given by Couwenberg (2003), p. 3, where parties had apparently failed to realise that the square of a negative number is positive.

⁹⁶ Stigler (1961) and its progeny.

probability is. On the other hand, for the most part it will prove more practical to simply assume that mistakes do not happen, and leave it to other disciplines to explain why this is not true.⁹⁷

At law, omissions can happen not only because of error, but also because of the intervention of a higher rule. In a labour contract, it is in the interest of both parties that the employer should invest in the employee's human capital, so as to maximise their productivity.⁹⁸ Such investments are ordinarily part of the contract, under the heading of "schooling".⁹⁹ However, parties frequently find themselves in difficulties because they cannot assure that both will benefit from such investments. Once the investment has been made, there is no legal way to assure that the employee will not resign his position and offer his services to another employer. Even though such a term would be in the interest of both parties, the law forbids contractual terms that boil down to creating a situation of involuntary servitude.¹⁰⁰ As a result, economic theory suggests a second-best solution whereby the employer only pays for those investments in human capital that are specific to his organisation¹⁰¹, while the employee pays for those investments that could also be of value to other employers.¹⁰² This is an example of the public policy exception to the freedom of contract.¹⁰³

A different kind of "higher rule" that could potentially cause a kind of incompleteness is the common law parole evidence rule. This rule can cause incompleteness not because it blocks terms on principle, but because it makes it impossible to prove certain terms exist. On its face, it states that parole evidence, i.e. evidence that is not in the form of documents, "cannot be received to contradict, vary, add to or subtract from the terms of a written contract, or

⁹⁷ The same goes for ambiguity. Cf. Clark (1998), p. 24-32, Restatement (second) of contracts § 200-204, and *Bores v. Domino's Pizza*, Court of Appeals for the 8th Circuit, June 20, 2008.

⁹⁸ Or, to be precise, invest to the point where the marginal contribution to the productivity of the employee of an additional dollar of investment is exactly one.

⁹⁹ Cf. North-Rhine Westphalia Act of 6 November 1984, GV. NRW. S. 678, *Gesetz zur Freistellung von Arbeitnehmern zum Zwecke der beruflichen und politischen Weiterbildung*, i.e. Act Regarding the Release of Employees for the purposes of Continued Professional and Political Education. Cf. also ILO Convention No. 142 of 1975, regarding Human Resources Development.

¹⁰⁰ Restatement (second) of contracts § 367. Cf. *Sommersett's Case* (1772), 20 How.St.Tr. 1, per Lord Mansfield, the XIIIth Amendment to the US Constitution, art. 15 Dutch Constitution, art. 12 German Basic Law, art. 7(1)(c) of the Statute of Rome concerning the International Criminal Court, art. 5 of the EU Charter of Human Rights, art. 4 of the Universal Declaration of the Rights of Man, art. 4 of the European Convention on Human Rights, art. 1 of the European Social Charter, art. 8 International Treaty regarding Political and Civil Rights. Cf. also Kronman (1983). For a criticism of this argument, cf. Oman (2009).

¹⁰¹ A discussion of asset specificity, including human capital specificity, follows in chapter 2.3 below.

¹⁰² Becker (1964).

¹⁰³ Eg. Restatement (second) of contracts § 178, art. 3:40 Dutch Civil Code, art. 6 Code Civil or § 134 BGB.

the terms in which the parties have deliberately agreed to record any part of their contract.”¹⁰⁴ It is easy to see how this could make the contract more incomplete than the parties may have realised. On the other hand, when the UK’s Law Commission evaluated the rule in 1986, they found that it had so many exceptions, that “no parole evidence rule today requires a court to exclude or ignore evidence which should be admitted or acted upon if the true contractual intention of the parties is to be ascertained and effect given to it.”¹⁰⁵ It is significant nonetheless because, although it is not a rule of evidence itself¹⁰⁶, it is a well-known example of the effect that procedural and evidentiary rules can have on contractual completeness.¹⁰⁷

Finally, a contract can say nothing about an eventuality because it does not involve enough parties. Analytically, Dixit & Olsen have shown that where entry into a negotiation is costly, a form of prisoners’ dilemma ensues, where parties end up staying out of the negotiations, even though participating would bring them strictly positive results.¹⁰⁸ This conclusion stands even where these costs are very small, for example the costs of learning the details of the problem that is to be discussed. In practice, contracts often have externalities, such as the case of *Allen v Rescous* (1676), where two men contracted to commit an assault on a third.¹⁰⁹ However, in some cases, a contract that involves too few parties is not only inefficient, but also obligatorily incomplete.

An historic example of this happening is the Luxembourg crisis of 1867.¹¹⁰ In the contract of sale between the Dutch king Willem III and France, whereby title to the Grand Duchy of Luxembourg would pass from the former to the latter, no arrangement was made about whether the ordinary neutrality of the Grand Duchy would be maintained in case of any future war between France and a third country. This omission is understandable, given that neither

¹⁰⁴ *Bank of Australasia v. Palmer* [1897] A.C. 540 at 545, per Lord Morris. While this is the most common definition of the rule, it was actually not decided by the judge, but rather agreed by the parties in the case. Cf. Law Commission of the United Kingdom report No. 154 of 1986, p. 2n.

¹⁰⁵ Law Commission of the United Kingdom report No. 154 of 1986, p. 3. When Clark (1998), p. 111-117 discusses the rule, he finds the following categories of exceptions: 1) to establish the limits of the contract, 2) to explain the circumstances surrounding an agreement, 3) to explain the subject matter of the contract, 4) in case of mistake, 5) to establish the existence of consideration, 6) to prove custom and 7) where the written document is not the entire contract.

¹⁰⁶ Restatement (second) of contracts § 213, comment a, Law Commission of the United Kingdom report No. 154 of 1986, p. 1-2.

¹⁰⁷ Of course, it could be argued that there is no incompleteness in these examples, at least not in the legal sense, because the law does give an answer to what should be done, albeit possibly an inefficient one. This problem of distinguishing in practice between obligational incompleteness and insufficient state contingency is discussed further, below.

¹⁰⁸ Dixit & Olsen (2000). Cf. also Aderlini & Felli (2006).

¹⁰⁹ *Allen v. Rescous* (1676) 2 Lev. 174. Cf. Clark (1998), p. 305. Obviously, this contract was declared void for illegality.

¹¹⁰ Cf. Van der Pot & Donner (2001), p. 128-129 and Fyffe (1895), chapter XXIII.

the king nor the French government immediately benefited from any stipulation on this point. When the Prussian government under Count Bismarck objected to the sale on these conditions, the Dutch government intervened and forbade the king from going through with the sale.¹¹¹ Due to the non-involvement of the Prussian government in the contract, no arrangement was made for those states of the world where France would be at war while Luxembourg would not be. A negotiation with all interested parties might have resulted in an arrangement that was preferred by all to no sale at all.¹¹²

In many of the cases that will be discussed in later chapters, the “missing party” in the negotiation is the population at large, or at least that part of the population that has some interest in consuming the service or good in question. Usually, it will be assumed that the government acts as a perfect agent for the population, which, if true, means that parties will negotiate an obligatorily complete contract anyway.¹¹³

Having considered what one might describe as the “easy” version of incompleteness, where there is literally an omission in the terms of the contract, causing it to fail to specify an action set for a certain state of the world, we will now continue with the other version proposed by Ayres & Gertner (1992): insufficient state contingency. While contracts that are obligatorily incomplete are clearly “incomplete” no matter how you look at it, it may not be immediately obvious that contracts that are insufficiently state contingent also can also be described in this way.

¹¹¹ The Prussian objection was motivated in part by pure security concerns, and in part by the fact that formally Luxembourg was still a part of Germany, as it continued to be until the creation of the German empire three years later.

¹¹² The result that was ultimately reached was that the sale did not go ahead, and that a treaty was made, the Treaty of London, whereby the continued neutrality of the Grand Duchy was guaranteed. It is difficult to say whether this result was Pareto efficient, especially since in politics the usual economic prescription of side payments is frowned upon. Cf. Holtermann (2007), p. 48.

¹¹³ The question of whether, and in which circumstances, the assumption of perfect agency is a reasonable one is studied by a branch of economics called Public Choice theory. Cf. Mueller (2003) and Chapter 3.2, below.

2.1.4 *Insufficient State Contingency*

It should already be obvious from the examples discussed in the previous section that the difference between obligational incompleteness and a lack of state contingency is not always clear. Many of these examples can be recast as involving the same action set for a series of states of the world where more flexibility would have been called for, i.e. as involving a lack of state contingency. Others can be viewed as involving an inefficient choice by parties.

The Luxembourg crisis, for example, involved a contract of sale where France was given complete control over Luxembourg. Depending on whether this is assumed to mean that the action set is empty, or that the action set involves no obligations of any kind¹¹⁴, this is either an example of obligational incompleteness or of lack of state contingency. Based on the facts of the case, both interpretations are defensible, and the difference matters only analytically.

Nevertheless, the distinction should not be abandoned. Not only are there also plenty of examples that cannot be reinterpreted in this way, but the difference is also important in order to make sure that law scholars and economists do not misunderstand each other.¹¹⁵ In any case where a jurist considers a contract to be incomplete, the economist will normally agree, but the reverse is not necessarily true.

For an example of excessive rigidity, where only an economist would see incompleteness, consider Declaration number 30 made by the Intergovernmental Conference when it adopted the Treaty Establishing a Constitution for Europe.¹¹⁶ Quoted in full, it states: “The Conference notes that if, two years after the signature of the Treaty establishing a Constitution for Europe, four fifths of the Member States have ratified it and one or more Member States have encountered difficulties in proceeding with ratification, the matter will be referred to the European Council.” Given the political complexities of negotiating such a treaty, this is truly a unique provision; in that it explicitly considers what the action set should be in case of a

¹¹⁴ Cf. art. 5:1(1) of the Dutch Civil Code, which defines ownership as “the most wide-ranging right that a person can have with respect to a good”. Similar definitions can be found in other legal systems, for example in art. 544 Code Civil. Since right and obligation are each other’s mirror image, this implies that ownership equally involves the most limited obligation that a person can have with respect to a good. Note, however, that the Grand Duke of Luxembourg in the 19th century did not “own” the land in the sense that the medieval kings of England owned their land, Pollock & Maitland (1898), p. 69, which is why it is probably better to avoid talking about ownership in this context.

¹¹⁵ Cf. Ayres & Gertner (1992), p. 730.

¹¹⁶ OJ C310/464, dd. 16.12.2004.

problem with ratification. Arguing *a contrario*¹¹⁷, we can infer that the drafters of the treaty intended for it to be essentially “dead” in those states of the world where more than one fifth of the member states would have “difficulties”.¹¹⁸

In retrospect, we can see that this provision was insufficiently flexible in that it failed to take into account the political reality of having the ratification process spread out over an extended length of time. Instead of having a head count at the end, countries started hesitating once France and the Netherlands “encountered difficulties”. In the end, there were only two countries with real difficulties, but on the other hand only 13 member states deposited their instruments of ratification with the Italian government.¹¹⁹ Since, on further reflection, declaring the Constitutional Treaty “dead” at this point proved undesirable, the provision of Declaration no. 30 was renegotiated *ex post*, and the signatory states decided to have a “period of reflection” followed by renegotiation, instead.¹²⁰ Such *ex post* renegotiation is often the result of contractual incompleteness of either variety.

In *Saladin v. HBU* (1967), the Dutch Supreme Court ruled on a contract between a bank and its customer, and particularly on the exemption clause that it contained.¹²¹ In this case, the bank had given Saladin bad investment advice, causing him to lose a lot of money. When he sued to recover, they pointed to their exemption clause. Of course, exemption clauses can be welfare enhancing, because they allow parties to place certain kinds of risk with one party or another, depending on their attitude to risk, their ability to insure, their ability to affect the likelihood of the event occurring, or any of a number of other factors that determine the most efficient division of liability.¹²² The clause that HBU wrote, however, placed all lia-

¹¹⁷ Cf. Dutch Supreme Court ruling of 10 December 1976 (Winterthur), NJ 1978/114, Henket (1992), Kaptein (1993) or Jansen (2005).

¹¹⁸ N.B. To the extent that this reading is not justified, this example would be one of obligational incompleteness.

¹¹⁹ Of the 25 member states, 16 decided to ratify. Of those 16, Slovakia and Germany had Constitutional Court cases pending that prevented them from depositing their instruments of ratification, while Finland had to wait for a vote by the parliament of the autonomous Åland islands. The Czech Republic, Denmark, Ireland, Poland, Portugal, Sweden and the United Kingdom abandoned their ratification processes after the referendum results in France and the Netherlands.

¹²⁰ N.B. This example is obviously flawed in one important respect: the uncertainty lay not in anything that happened after this contract became binding, but rather in the question of whether it would be ratified in the first place, i.e. in the pre-contractual phase. This is the result of the existence of a distinction between the fixing of the text of a treaty and the ratification of the treaty, which is essentially unique to international law. (Cf. Bloembergen et.al. (2001), p. 75-82.) The example is discussed here anyway, because this flaw does not take away from its ability to illuminate the notion of insufficient state contingency.

¹²¹ *Saladin v. HBU* (1967), Dutch Supreme Court 19 May 1967, NJ 1967/261. This case was the basis of what would later become art. 6:248 (2) Civil Code: *bona fides* as grounds for denying a party to a contract the relief they would otherwise be entitled to under the contract. Exemption clauses are discussed in Clark (1998), chapter 7 and in *George Mitchell (Chesterhall) Ltd v Finney Lock Seeds Ltd* [1983] QB 284, per Lord Denning, MR.

¹²² Cf. Mauel (1996), Seita (1984).

bility with Saladin for a wide range of states of the world, including states of the world, such as the one that was ultimately realised, which could only be the result of the bank's mismanagement. Aside from the question of whether this bargain was unconscionable¹²³, a Pareto superior contract could have been written by limiting the exemption clause to those states of the world where the bank was not culpable.

While the previous two examples concerned contracts that were insufficiently state contingent because that was the way they were written by parties, a lack of state contingency can also occur if we analyse examples of obligational incompleteness at the level of the law in its entirety.

Where the law finds a contract that is incomplete in the legal sense, it will have to fill in the blanks somehow. To do so, the court first examines whether any contract can be said to exist at all. Where there is no agreement over something that is considered essential to the contract, the contract is said to be void for incompleteness.¹²⁴ For example, where the agreement extends to little more than the price, the courts will normally assume that parties did not yet finish negotiating.¹²⁵

If the contract is not void, the court will fill in the blanks in such a way as to follow "the presumed intention of the parties".¹²⁶ Obviously, this is an exercise that is extremely context-dependent, and therefore quite flexible. However, because this kind of inquiry can only emerge *ex post*, i.e. after it has become known which state of the world has been realised, it is unavoidable that in those cases where the doctrine of implied terms has been applied, it is possible that parties could have done better by writing a more complete contract *ex ante*. That said, a lack of state contingency is not a problem here.

¹²³ In *Gillespie Bros. & Co. Ltd. v. Bowles (Roy) Ltd.*, [1973] QB 400, Lord Denning, MR. applied the doctrine of unconscionable bargain to an exemption clause. Cf. Clark (1998), p. 160.

¹²⁴ *Pagnan Sp.A. v. Feed Products Ltd.* [1987] 2 Lloyd's Rep. 601. (English Court of Appeals), Dutch Supreme Court *Polak v. Zwolsman*, dd. 14 June 1967, NJ 1968/331.

¹²⁵ *Harvey v. Facey* [1893] AC 552. (House of Lords, in an Irish case.)

¹²⁶ *The Moorcock* [1886-90] All. E.R. 530; (1889) 14 PD 64, per Bowen L.J. This is the classic English case on implied terms. Other English House of Lords cases include *Liverpool City Council v Irwin and Another* [1976] 2 All E.R. 39; [1977] AC 239 and *Imperial Group Pension Trust Ltd. and Others v. Imperial Tobacco Ltd. and Others* [1991] 2 All E.R. 597.

For Dutch law, cf. Bloembergen et. al. (2001), p. 336-337, which discusses art. 6:248 Dutch Civil Code, the obligation of good faith, which has been used to imply terms in Dutch law. The French Code Civil contains a treatment of the interpretation of contracts, similar to the one in *The Moorcock*, in articles 1156-1164. (The general obligation of good faith is in article 1134 CC.) The German BGB contains an obligation of good faith in § 242. The situation in US law, finally, is discussed in Restatement (second) of contracts § 204.

The same cannot be said for statutory default rules.¹²⁷ Such rules, which set a kind of statutory background for contracting, can be either immutable or waivable. The *ex turpi causa* rule¹²⁸ and the rule against specific performance in employment contracts¹²⁹, which was discussed above, are examples of immutable default rules. These rules have the effect of making the contract void. Alternatively, immutable rules can insert a term into the contract, whether parties want it or not. Examples include the general obligation of good faith, and a host of rules aimed at protecting consumers.¹³⁰

Where a default rule allows for parties to waive it, the degree of state contingency of the legal relationship between them should return to its optimal level. After all, a naive model of contracting would assume parties to waive the statutory rule if there are gains from doing so. On the other hand, if that were truly the case, there would be no benefit from having such a statutory scheme in the first place, since parties could then equally be assumed to introduce the term in question when it would be beneficial to do so. A good approach to understanding the functioning of waivable default rules is probably the line of research carried out by Richard Thaler and Cass Sunstein, under the heading of “liberal paternalism”.¹³¹ Put simply, they assume that for a given proposed default rule, there are four possible circumstances: 1) the rule is not beneficial and parties are rational enough to waive it, 2) the rule is not beneficial and parties are not rational enough to waive it¹³², 3) the rule is beneficial and parties are not so rational that they would have written such a term anyway, and 4) the rule is beneficial and parties are so rational that they would have written such a term anyway. In situation 1, parties will waive the term, in situation 4, the optimal result will obtain no matter what kind of default rule the legislator specifies, and so the main question for the legislator is situations 2 and 3. Sunstein and Thaler advocate that the law maker should define a waivable default rule that maximises the total gain in those two situations. In most cases, this comes down to writing a rule that makes situation 3 as likely as possible relative to situation 2.¹³³

¹²⁷ Default rules are analysed from an economic point of view in Ayres & Gertner (1989).

¹²⁸ “No action may be founded on an wrongful cause.” For a detailed discussion of the different modalities of voidness for illegality, cf. any textbook on Contract Law, for example Clark (1998), p. 305-369 or Bloembergen et.al. (2001), p. 178-192.

¹²⁹ Cf. Kronman (1983), p. 774-786.

¹³⁰ Eg. Art. 7:13 Dutch Civil Code, which gives rules for transport costs in consumer purchase contracts. The German BGB has a system of rules for consumer purchase in §§ 474-479.

¹³¹ Sunstein & Thaler (2003, 2006), Thaler & Sunstein (2003, 2008a, 2008b). Cf. also Choi et al. (2003).

¹³² This irrationality can come in many guises; it can be a matter of a lack of knowledge, but equally a problem of simple inertia.

¹³³ To the extent that the rule comes in the form of an opt-out, this can mean choosing the rule that “minimises the number of opt-outs”. Thaler & Sunstein (2003), p. 178-179. The authors give the example of a 401(k) pen-

To the extent that waivable default rules are not waived when they should have been, the contract is Pareto inefficient; there are potential gains from trade that have not been realised. Sometimes, this inefficiency amounts to a lack of state contingency. Take the difference between a law that states that employees can only be dismissed for cause, which is common in Europe, and the US practice of allowing employees to be dismissed at will. In many circumstances, employer and employee can both prefer to exclude the possibility of dismissal at will, since that would arguably increase the employee's investment in human capital that is specific to his job. However, a blanket rule that dismissal is only possible for cause would also be inefficient; there are many possible states of the world where the employer would value the option of being able to dismiss the employee, and those are not limited to states of the world where the employee has behaved in a manner that qualifies as "cause" under employment law.¹³⁴ So either blanket rule would be insufficiently state contingent, since additional value could be realised if parties would work out in more detail in which circumstances dismissal at will should be possible.¹³⁵

Summarising the chapter so far, we can say that contracts are incomplete in the legal sense if parties neglected to specify an action set for one or more states of the world, or if a higher rule intervenes to make the contract incomplete. In the economic sense, a contract is incomplete if a Pareto superior alternative could have been achieved by writing in more detail. An important class of cases where this is the case is where we view a contract that is obligatorily incomplete in conjunction with the statute law that caused this incompleteness. In those situations, the contract in isolation is obligatorily incomplete, and the legal relationship between the parties as a whole is incomplete in the sense that it is insufficiently state contingent.

Having discussed the definitions proposed by Ayres & Gertner (1992), as well as their practical application, what is left is to consider an entirely unrelated reason why contracts can be considered incomplete: contracts may contain terms that are unenforceable because they

sion plan, where the law can either ask employees to opt-in or to opt-out. Under perfect rationality, this choice would be irrelevant, but show evidence that this is not true, causing them to prefer the opt-out approach for paternalistic reasons. Sunstein & Thaler (2003), p. 1172-1173.

¹³⁴ For the Pareto principle to favour allowing dismissal, the employer would obviously have to favour this option highly enough that he could indemnify the employee for accepting this risk and still come out ahead. In the Dutch civil code, art. 7:678 lists the reasons why an employee can be summarily dismissed. § 626 of the German BGB only gives a general standard against which to judge whether summary dismissal was justified, while the French law on this point is laid down in the Code du Travail, art. L1231 up to L1238-5.

¹³⁵ The employer might like to reserve the option of dismissal without cause in case of financial difficulties. Cf. art. 40 and 239 Dutch Bankruptcy code.

refer to facts that cannot be observed or verified. In such a case, we will generally treat them as if they do not exist, even though their effects may well be felt in ways outside the expertise of economists.

2.1.5 Observability and verifiability

New College, Oxford, was founded in 1379 by the Bishop of Winchester, William of Wykeham. In making arrangements for this new college, one of the things that he decided was that the college should have a chapel with extensive staff, whose duties included regular prayer for the Bishop's soul.¹³⁶ In the centuries since its founding, the college has gradually reduced its clerical staff to the point where there is only a single chaplain left, and prayers for William of Wykeham are held only twice a year. The question is: what can be done about this breach of contract?¹³⁷

The problem of interest here is not the question of who could have standing to sue¹³⁸, or whether a court order requiring more regular prayer would be in violation of the principle of the separation of church and state, but rather whether it is physically possible for the court to give such an order in the expectation that it will be obeyed. Is this the kind of thing that a court is capable of ordering? Where the subject matter is a full Catholic mass, the answer is surely yes, but what about prayer of the kind that was recommended in the Sermon on the Mount?¹³⁹ If prayer is done in this way, where no one can see, how can it be ordered by a court?¹⁴⁰

¹³⁶ Dawkins (2007), p. 358-359.

¹³⁷ In protestant England, a will creating a trust for the saying of (Catholic) masses was traditionally considered void on the grounds that it was for "superstitious uses". *West v. Shuttleworth*, [1835] 2 Myl. & K. 684. This rule was overruled by the House of Lords in *Bourne v Keane*, [1919] AC 815. Cf. also *Nelan v Downes*, [1917] HCA 51. (High Court of Australia) and *Sherman* (2008).

¹³⁸ Cf. US Supreme Court, *Massachusetts v EPA* (2007), 127 S.Ct. 1438 and US Supreme Court, *Lujan v Defenders of Wildlife* (1992), 504 US 555, 112 S.Ct. 2130.

¹³⁹ "And when thou prayest, thou shalt not be as the hypocrites [are:] for they love to pray standing in the synagogues and in the corners of the streets, that they may be seen of men. Verily I say unto you, They have their reward. But thou, when thou prayest, enter into thy closet, and when thou hast shut thy door, pray to thy Father which is in secret; and thy Father which seeth in secret shall reward thee openly." New Testament, King James version, Gospel according to Matthew, verse 6:5-6.

¹⁴⁰ In fact, the closest courts have come in this respect is in cases where a conflict within a church, regarding the question how services should be conducted have ended up before civilian courts. Eg. *Wylde v Attorney-General (N.S.W.) (At the Relation of Ashelford)*, [1948] HCA 39. (High Court of Australia), where the courts were asked whether the terms of a charitable trust were being obeyed.

In economic literature, it is customary to distinguish between problems due to a lack of observability and problems due to a lack of verifiability.¹⁴¹ Something is verifiable if a court, or any other independent third party, is able to verify whether it has occurred, while observability requires merely that the parties to the contract can ascertain the truth.¹⁴² A contractual stipulation that requires the periodic holding of a mass is both observable and verifiable, since the court can quite easily question witnesses to discover whether the stipulation has been obeyed. Similarly, a contract that states that parties shall pray together regularly is observable only by parties, while a contract that states that party A shall pray in private for the benefit of party B is neither observable nor verifiable.

Another distinction that is useful to make is between cases where the problem lies with the observability of the state of the world that has been realised and cases where the problem lies with the observability of the action set. In the previous example, the problem is of the latter variety. In fact, given that the bishop's arrangements were extremely inflexible with respect to states of the world, this would hardly be otherwise.¹⁴³

In statute law, it is fairly common for one of both parties, usually the government, to have difficulty proving which state of the world has been realised.¹⁴⁴ Whenever the law creates an entitlement dependent on factual circumstances, parties will have to agree on whether those factual circumstances actually occurred, i.e. whether the claimant is right in their description of the state of the world. Often, lawmakers have created bureaucratic solutions to this problem, with forms to be filled out giving information about one's employment status, for instance. In other cases, lengthy disputes arise in the courts, because a bureaucratic solution may not be feasible and the stakes may be high. In most countries, the immigration laws make an immigrant's right of asylum dependent on the circumstances in their home coun-

¹⁴¹ Eg. Hart & Moore (1988), p. 759. While it is clear why one should be interested in whether or not something can be verified by a judge, it may be less obvious why observability should matter. This is discussed at length below, but one important answer is the possibility of constructing a revelation mechanism, whereby parties are given an incentive to truthfully reveal their private information. Cf. Myerson (1979).

¹⁴² Obviously, it follows from these definitions that anything that is verifiable is also observable.

¹⁴³ According to Dawkins (2007), quoting the college's history books, the bishop stipulated that, should the income of the college fail, the chapel and its staff were to be the last to be cut from the budget. (p. 359). This shows that Wykeham tried to consider what he would want in certain probable states of the world, making his arrangements more state contingent than they otherwise would have been.

¹⁴⁴ Especially if the statute is written in terms of general standards rather than detailed rules. Cf. Ehrlich & Posner (1974). The debate as to rules vs. standards is still ongoing, cf. Mahoney & Sanchirico (2005), yet outside the scope of this thesis. The reader is merely invited to consider to what extent a standard that says nothing more than that individuals shall take reasonable care not to cause damage to another is obligatorily complete and state contingent.

try.¹⁴⁵ Particularly because of the high stakes for the immigrant, lengthy factual disputes often arise about the exact level of danger they would be in if they were to be sent back.¹⁴⁶ This factual argument reflects the lack of verifiability, for immigration purposes, of the state of the world in question.¹⁴⁷

In contract law, the most obvious area where disputes may arise due to a lack of verifiability of the state of the world is insurance law. After all, the nature of an insurance contract is an exchange between two promises: on the one hand, the promise to pay a relatively modest sum of money no matter which state of the world is realised, and on the other hand the promise to pay a more considerable sum in certain carefully defined states of the world, such as the state of the world where your house burns down or the one where you end up in the hospital. This means that insurance litigation focuses either on interpreting the terms that have been used to describe the state of the world where a claim can arise¹⁴⁸ or on whether such a state of the world has actually been realised.

Finally, bearing in mind the positivist approach to law that was adopted above¹⁴⁹, one could pose the question whether a lack of observability or verifiability leads to obligational incompleteness, or whether it should be considered a separate category of incompleteness. After all, if legal norms are only legal if they are enforced by the courts, a norm that depends

¹⁴⁵ Eg. 8 USC 1101(a)(42)(A) and 8 USC 1158 in the United States Code, the UK the Immigration Acts 1971-2006, as amended most recently by the Immigration, Asylum and Nationality Act of 2006 (c. 13), and the Dutch Vreemdelingenwet, Stb. 2000, 495. In international law, the most important treaty relating to refugees is the 1951 Convention Related to the Status of Refugees, and its 1967 amending Protocol, both negotiated under the auspices of the United Nations.

¹⁴⁶ To the extent that past actions become part of the description of the state of the world, cf. fn 11, above, an even larger part of statute law litigation gives rise to disputes that concern the state of the world. Criminal law, for example, is of the nature of a promise, by the state, that he who commits act X shall be punished in a certain way. Here, X is the premise of the promise. Whether this premise has actually been fulfilled is the subject of any criminal prosecution. The use of a description of a state of the world as a premise for a legal statement is discussed in Ruiter (1993), p. 8-11.

¹⁴⁷ Remember that, in order to keep the notion of a state of the world tractable, we only distinguish between different states of the world to the extent relevant for the problem at hand.

¹⁴⁸ The owners of the World Trade Centre in New York had insurance against damage caused by a terrorist attack, which was capped at \$ 3,5 bn per attack. After September 11, 2001, there was quite a bit of litigation over the question whether the events of that day should be considered one attack or two. Cf. *SR International Business Insurance v WTC Properties*, United States Federal District Court for the Southern District of New York, rulings of June 3, 2002 (2002 WL 1163577), September 25, 2002 (222 F.Supp.2d 385) and January 29, 2003 (2003 WL 192487), and the appeal between the same parties before the Court of Appeals for the Second Circuit, ruling of September 26, 2003 (345 F.3d 154). The litigation continued on related issues afterwards, and was only concluded with another ruling by the second Circuit on October 18, 2006, which set the total amount to be paid out at \$ 4,6 bn (467 F.3d 107).

¹⁴⁹ Chapter 2.1.1. The approach is positivist in the sense that, like most law & economics, it analyses legal validity without reference to moral principles. Cf. Hart (1961) and Werner (1995).

on an unverifiable premise¹⁵⁰ or that requires an unverifiable action could not very well be described as a legal norm. This goes *a fortiori* for the problem of observability. In legal terms, a promise to do something unobservable would be described as an illusory promise, because in such a situation performance is entirely up to the discretion of the promisor.¹⁵¹ It follows that it is quite reasonable to treat unobservable terms as if they do not exist, making the contract obligatorily incomplete. Where the problem is one of verifiability, but not observability, on the other hand, there might be some way to save it; the term may still have some effect between parties.¹⁵² Therefore, treating all non-verifiable terms as non-existent would seem premature.

In conclusion, we can say that the law is a mechanism for establishing action sets in different states of the world. Due to a variety of reasons, the law can be obligatorily incomplete or insufficiently state contingent, i.e. incomplete. Given the methodology proposed in section 1.2, above, with its distinction between the derivation of the model and its resulting “truth” on the one hand, and its applicability on the other hand, the remainder of the chapter will not always spell out the practical application of these basic concepts in detail. Only in chapter 3 will we return to them.

First, however, the next section will discuss the Coase Theorem, an idea that “has captured the attention and interest of economists and legal scholars as have few other ideas”.¹⁵³

¹⁵⁰ Ruiter (1993), p. 8-11 describes the state of the world for which a legal norm prescribes an action set as the premise of that norm.

¹⁵¹ “[The doctrine of illusory promise] is used to invalidate contracts where one party has not actually bound itself to do anything. The traditional example of an illusory promise is “I will give you ten dollars if I feel like it.”” Lowe (2007), p. 710. Since it is an aspect of the doctrine of consideration, the notion of an illusory promise is unique to common law. Cf. *Wood v Duff-Gordon*, New York Court of Appeals, December 4, 1917, 222 N.Y. 888, 118 N.E. 214, *Portland Gasoline v. Superior Marketing*, Texas Supreme Court, November 28, 1951, 150 Tex. 533, 243 S.W.2d 823. Cf. also Clark (1998), p. 24-26 and Restatement (second) of Contracts, § 77.

¹⁵² Cf. fn 75, above.

¹⁵³ Medema & Zerbe (1999), p. 836.

2.2 The Coase Theorem

2.2.1 Introduction

In 1960, Ronald Coase wrote something truly amazing: “The traditional approach has tended to obscure the nature of the choice that has to be made. The question is commonly thought of as one in which A inflicts harm on B and what has to be decided is: how should we restrain A? But this is wrong. We are dealing with a problem of a reciprocal nature. To avoid the harm to B would inflict harm on A. The real question that has to be decided is: should A be allowed to harm B or should B be allowed to harm A?”¹⁵⁴ It is hard to overstate how much this statement goes against everything lawyers are taught about law. In law, the question is usually phrased in terms of who has the better claim, the stronger right. The strongest of these, of course, are called human rights.¹⁵⁵

While it was only to be expected that an economist, looking at the law from the outside, would not accept an argument derived from dogma, i.e. from a Rule of Recognition¹⁵⁶ that has no justification outside the law¹⁵⁷, but would instead apply an economic approach to his chosen subject, the analysis that followed surprised lawyers and economists alike. Unfortunately, the article does not explicitly state what has later become known as the Coase Theorem: “In a world of zero transaction costs, the allocation of resources will be efficient, and invariant with respect to legal rules of liability”.¹⁵⁸ So what does this mean?

First of all, even though this thesis will conform to the custom of illustrating the Coase Theorem primarily with examples taken from tort law¹⁵⁹, it should be clear that “legal rules of liability” can be found in other areas of law as well.¹⁶⁰ As we will see, Coase’s aim in phrasing the question in the manner quoted above was precisely to discover in which circumstances

¹⁵⁴ Coase (1960), p. 2. According to the tally of Shapiro (1996), this is the most cited law review article of all time.

¹⁵⁵ Cf. US Declaration of Independence: “We hold these truths to be self-evident, that all men are created equal, that they are endowed by their Creator with certain *unalienable* Rights, that among these are Life, Liberty and the pursuit of Happiness.” (emphasis added)

¹⁵⁶ Hart (1961)

¹⁵⁷ Cf. Ruiters (1993), p. 15-17.

¹⁵⁸ Zerbe (1980), p. 84. The author adds a qualification to his definition (“income effects aside”), which has been omitted here, but will be discussed below. The various definitions that have been proposed in the last forty years are discussed in Medema & Zerbe (1999), and in Steenge (2004).

¹⁵⁹ Cf. Coase (1960) itself. Note that this complements nicely with the previous section, which focused almost entirely on contract law.

¹⁶⁰ This extension of the Coase Theorem is due to Cheung (1969).

the pre-existing rules that are found in tort law should be amended through contract. Similarly, entitlements are found in administrative law, where again the question is whether the principles laid down by laws of general application should be amended by some other legal instrument, in order to make the result more efficient given the circumstances of a specific case.¹⁶¹ In fact, the Coase Theorem applies to all rules that seek to establish an action set for one or more states of the world.

Having established a basic definition of the Coase Theorem, this section will first discuss Coase's arguments for his claim, which are in the form of various hypotheticals derived from tort law. Thereafter, attention will turn to some exceptions to the Theorem. However, the most important one of these – transaction costs – will be discussed separately in section 2.3, below.

2.2.2 Torts and the Coase Theorem

Analytically, the nature of tort law is quite simple: it is a collection of legal rules, each of which attaches an action set to a specific state of the world. Theoretically, at least, the rules should not overlap: for a give set of circumstances, only one rule should apply.¹⁶² The action set that is prescribed by the rule is called the remedy. In tort law, it usually consists of a sum of money that is to be paid as damages, but other remedies are also possible.¹⁶³ Since it is important that an economic analysis of the law should not be biased towards either common law or civil law, it is important to consider briefly the differences between the two systems before proceeding with the analysis.¹⁶⁴

¹⁶¹ The word “entitlements”, which is used here, is the general term for claims and privileges, in the Hohfeldian sense. Cf. 2.1.1, above, and Ruiters (2007). In the model developed by Calabresi & Melamed (1972), these entitlements can be “protected” by either property rules, which are *erga omnes*, or liability rules, which are established only between two people. It is not entirely clear whether this distinction between entitlements and protective rules is sensible within our approach. After all, these protective rules are themselves conditional entitlements, conditioned on a violation of the original entitlement.

¹⁶² In practice, it is not uncommon for different torts to be argued in the alternative. Usually, this has to do with the plaintiff's uncertainty about whether they will be able to prove all the facts, which is why a plaintiff might argue the deliberate and negligent version of a tort in the alternative. Also, the rule that applies might be a composite of a statutory or precedent-based rule, modified by a contract, for example because the contract might contain an exemption clause.

¹⁶³ For example, the plaintiff might ask for an injunction ordering the defendant to stop his tortious behaviour (cf. art. 3:296 and 6:168 Dutch Civil Code), or for a rectification (cf. art. 6:167 Dutch Civil Code).

¹⁶⁴ Cf. Brousseau (2001).

In common law jurisdictions the law of torts, which concerns itself with civil wrongs, has a structure similar to that of criminal law. Each tort has its own elements, each of which the plaintiff will have to prove, just like the public prosecutor in a criminal trial will have to prove every element of the crime that is charged.¹⁶⁵ For example, the tort of defamation is defined as “the wrongful publication of a false statement about a person, which tends to lower that person in the eyes of right-thinking members of society or tends to hold that person up to hatred, ridicule or contempt, or causes that person to be shunned or avoided by right-thinking members of society.”¹⁶⁶ The attentive reader will realise that this definition contains four elements: 1) publication, 2) falsity, 3) defamatory nature and 4) wrongfulness. For the third of these, the definition provides at least three alternative sub-definitions.¹⁶⁷

While in the common law the principle of *stare decisis*¹⁶⁸ has assured a degree of structure and predictability in the law of torts, in the civil law countries judges enjoy considerable freedom in this area, as long as they stay within the legal definitions. Central to this is the definition of an unlawful act, such as it can be found in article 1382 of the French Code Civil¹⁶⁹ or § 823 of the German BGB.¹⁷⁰ This fact makes civil law tort law more casuistic than its common law counterpart, less easily broken down into coherent lines of precedent. However, since our analysis is ultimately concerned only with the rule for an individual case, with little concern for how the courts arrive at that rule, the Coasian approach applies equally well to both systems.

¹⁶⁵ Cf. McMahon & Binchy (2000).

¹⁶⁶ McMahon & Binchy (2000), p. 882, quoting the Irish cases of *Quigley v Creation Ltd* [1971] IR 269 and *Berry v Irish Times* [1973] IR 368.

¹⁶⁷ It can be doubted whether wrongfulness should be considered a separate element. In any case, it will not have to be proven by the plaintiff, but instead it is open to the defendant to prove one of the defences that the law recognises, i.e. 1) justification, 2) absolute or qualified privilege, 3) fair comment, 4) consent, 5) apology and 6) offer of amends. Cf. McMahon & Binchy (2000), p. 917-951. Where the defendant succeeds in his proof, the wrongful nature of the statement is removed.

¹⁶⁸ “*Stare decisis* is the policy of the court to stand by precedent”, *United States Internal Revenue v Osborne*, Court of Appeals for the 9th Circuit 1996, 76 F.3d 306. Cf. US Supreme Court case of *Humphrey’s Executor v United States*, 295 US 602, at 627 (1935) and *Landes & Posner* (1976).

¹⁶⁹ “Tout fait quelconque de l’homme, qui cause à autrui un dommage, oblige celui par la faute duquel il est arrivé à le réparer.”

¹⁷⁰ “Wer vorsätzlich oder fahrlässig das Leben, den Körper, die Gesundheit, die Freiheit, das Eigentum oder ein sonstiges Recht eines anderen widerrechtlich verletzt, ist dem anderen zum Ersatz des daraus entstehenden Schadens verpflichtet.” (Paragraph 1.)

Turning now to the hypothetical cases that Coase himself used in his 1960 article, we find that the first one concerns two farmers, whose farms are adjacent to one another.¹⁷¹ One of them breeds cattle, and the other one grows crops. The claim is that, in a perfect world, it is socially irrelevant whether there is a rule of tort law that requires the owner of the cattle to pay for the damage caused by his cows straying onto the other farmer's land.¹⁷² In other words, under certain, admittedly unrealistic assumptions, social welfare will be unaffected by the existence or absence of such a rule. The proof goes as follows: In the absence of any obligation to indemnify, the owner of the crop will undertake to build a fence on his own expense, if and only if the cost of such a fence is lower than the value of the damage that is prevented by it. If a liability rule does exist, it will be the cattle farmer who considers whether to build a fence, again comparing the amount he would have to pay in damages to the price of the fence. In either case, the decision maker considers all socially relevant costs and benefits, and builds the fence only if that is the social welfare maximising choice. The only difference between the two scenarios is the effect on the income of both farmers.

Consider the same scenario again; only this time there is a third option: instead of building a fence, the damage can also be prevented by having the cattle farmer hire someone to look after the cows. If the cattle farmer is liable for the damage caused by his animals, he will obviously choose whichever option is cheapest for him: pay the damages, hire someone, or build a fence. If he is not liable, it will be the crop farmer who will consider whether it would be profitable to pay his neighbour to hire a mender for the animals. It may seem unfair for the farmer to pay someone not to cause him damage, but from a social welfare point of view, all that matters is that all costs and benefits are taken into account by whoever takes the decision.¹⁷³

A second case proposed by Coase was based on an actual English precedent, *Sturges v. Bridgeman*¹⁷⁴ and concerned a doctor who bought a house next to a confectioner.¹⁷⁵ When he tried to consult with his patients in his brand new consulting room, he found that the noise

¹⁷¹ Coase (1960), p. 2-6.

¹⁷² McMahon & Binchy (2000), p. 750-754, analyses cases involving animals under various torts, including the torts of nuisance and occupier's liability. The specific case under consideration here would be qualified as a case of cattle trespass at common law, eg. the English Court of Appeals case of *Wormald v Cole* [1954] 2 WLR 613.

¹⁷³ N.B. This conclusion rests on the assumption of an additive social welfare function, i.e. a social welfare function where total welfare is calculated by adding up the utilities of the members of society, with the additional assumption of constant marginal utility of money. If that last assumption is dropped, or if a more advanced Social Welfare Function is used, social welfare will likely change as a result of income transfers between members of society. Cf. Mueller (2003), p. 563-581, and further discussion, below, in par. 2.2.3.

¹⁷⁴ 11 Ch.D. 852 (1879).

¹⁷⁵ Coase (1960), p. 8-10.

and vibrations caused by his neighbour's machinery made it very difficult for him to do so. While a lay person, or even a lawyer, might argue that the confectioner was there first, and that the doctor should not have bought the building if he couldn't live with the noise from next door, Coase argued that it really does not matter: in a world without transaction costs, the neighbours will sort out their differences by contracting a specific level of noise in return for the payment of a sum of money, if and only if that is efficient.¹⁷⁶

Countless other examples could be given, and Coase actually gives a few more, but for now these two will suffice. The next step is to consider the exceptions to this seemingly elegant rule.

2.2.3 Exceptions to the Coase Theorem

One of the reasons why there are so many definitions of the Coase Theorem, is that different authors have different ideas about what to do with the exceptions. Should they be included in the definition, or listed separately? In fact, the definition by Zerbe given above reads, in full: "In a world of zero transaction costs, the allocation of resources will be efficient, and invariant with respect to legal rules of liability, income effects aside."¹⁷⁷ In the version given above, this final qualification was omitted for reasons of expositional clarity. After all, income effects represent only one category of exceptions. Other exceptions include the question of voluntary participation, questions about the degree of rationality that is required of actors and, of course, the existence of transaction costs.

Income effects matter for two reasons: because they can affect the social welfare calculation that underlies the definition of efficiency, and because they matter in their own right. Even though a detailed discussion of social welfare functions is outside the scope of this thesis¹⁷⁸, it is not very difficult to see how the validity of the Coase Theorem might be affected by the choice of efficiency concept: Pareto efficiency¹⁷⁹ or social welfare maximisation?¹⁸⁰ If one is going to apply the latter, is it going to be a multiplicative or an additive function? And is it going to be based on wealth or utility?

¹⁷⁶ In actual fact, the court ruled that the doctor was entitled to stop the confectioner from producing so much noise and vibrations.

¹⁷⁷ Zerbe (1980), p. 84.

¹⁷⁸ Cf. Mueller (2003), p. 563-581.

¹⁷⁹ Cf. Cowell (1986), chapter 8.

¹⁸⁰ In his specification of the Coase Theorem, Hendrikse (2003) explicitly opts for Pareto efficiency. (p. 68)

Given the explicit assumption of costless bargaining, applying the Pareto principle transforms the theorem into something that is possibly a tautology¹⁸¹ and in any event not very exciting. Even though the theorem says not only that the outcome will always be efficient, but also that the outcome will be invariant with respect to the starting point, i.e. that it will be the same Pareto optimum in every case, one could very well consider that it is begging the question, that it is almost inconceivable that it would not be true given its assumptions.

Using an additive social welfare function based on wealth, i.e. a function that simply adds up the wealth of all members of society, is identical to a Pareto criterion, and similarly immune to income effects. However, adding up individuals' utility means that a simple zero-sum income transfer could be welfare enhancing. To illustrate, consider a classic nuisance suit, which is a common type of legal problem that is analysed using the Coase Theorem. In *Hanrahan v. Merck, Sharp & Dohme* (1988)¹⁸², the plaintiffs alleged damage to their health, to the health of their livestock and interference with the enjoyment of their property, all primarily under the tort of nuisance.¹⁸³ While the plaintiffs lost because they were unable to prove that the defendant's emissions caused their damage, for the present purposes there is no reason not to assume that it did, in order to analyse the income effects involved under different definitions of efficiency. To be clear: the Coase Theorem claims that, under zero transaction costs, it would be irrelevant whether Merck, Sharp & Dohme had the right to emit whatever substances they pleased or whether Hanrahan and the other plaintiffs had an unlimited right of clean air, or something in between. Whatever rules the law lays down, they only function as a starting point for negotiations, and in a world of zero transaction costs and perfectly defined property rights, parties will simply bargain to achieve an efficient outcome to the extent that the starting point is not, of itself, efficient.

Given the reasonable assumption that the marginal utility of wealth is decreasing, any income transfer from a rich person to a poorer person would increase social welfare in case of an additive utility social welfare function.¹⁸⁴ Assuming that plaintiffs are the poorer party¹⁸⁵,

¹⁸¹ Usher (1998).

¹⁸² [1988] ILRM 629. It is a ruling by the Irish Supreme Court, focusing mainly on the issue of whether the doctrine of constitutional torts can be used to overcome the plaintiff's difficulty in proving causation. Cf. McMahon & Binchy (2000), p. 8-9.

¹⁸³ Technically, they claimed trespass, negligence and *Rylands v. Fletcher* as well, but the courts all treated the case as one of nuisance. Cf. McMahon & Binchy (2000), p. 8.

$$U = \sum u_i$$

¹⁸⁴ Formally: for all i , where, as before, $i \in I$ represents an individual, and u_i represents the total utility of that individual.

this means that starting out with Hanrahan and the other plaintiffs having an unlimited right of clean air would be preferable, from a social welfare point of view, compared to starting the negotiations with Merck having the right to pollute. Even though the amount of pollution would end up in the same, in the former case there would be a payment from the company to the plaintiffs, while in case of the latter, the plaintiffs would end up paying the company.¹⁸⁶ Because the utility that Hanrahan derive from any given euro is probably higher than the utility derived from it by the company's shareholders, social welfare would increase as a result of a payment from the company to the plaintiffs, and decrease in the opposite case. The same result holds for any multiplicative social welfare function¹⁸⁷, since in such cases social welfare increases as the income distribution becomes more equal.¹⁸⁸

Throughout, the choice of social welfare function depends on whether one wants to model certain preferences by including them in the objective function, or whether one prefers to deal with them separately. For example, in his discussion of the Coase Theorem, Nentjes (1993) starts by discussing the theorem in terms of Pareto efficiency and total welfare, two concepts that he seems to equate, before turning his attention to the relative fairness of the different possible outcomes.¹⁸⁹ As we have seen, an alternative approach is to endogenise the general preference for income equality by choosing a multiplicative social welfare function. In most circumstances, it is probably preferable to choose the former approach, since it more clearly explicitises the assumption that income effects matter in their own right, as well as allowing for a greater variety of assumptions with regard to how they matter. As noted be-

¹⁸⁵ Given that Merck, Sharp & Dohme is a public limited company, this need not necessarily be true. According to Merck's Annual Report over 2007 on form 10-K, there were approximately 173.000 stockholders of record at the end of 2007, down from 233.000 at the end of 2003. (p. 49)

¹⁸⁶ On the assumption that the optimum lies somewhere in the middle, so that a transaction takes place in either extreme. If, in fact, the company is so profitable that it could afford to completely indemnify the plaintiffs for any pollution it may wish to cause, there would be no transaction at all under the "Merck has the right to pollute" scenario, since no amount that the plaintiffs could offer would be enough to satisfy Merck.

$$U = \prod w_i \quad \text{or} \quad U = \prod u_i$$

¹⁸⁷ I.e. a function of the form $U = \prod w_i$ or $U = \prod u_i$, where w_i represents the wealth of each individual and u_i , as before, represents the utility of each individual.

¹⁸⁸ This can easily be demonstrated by observing that $1 \times 1 < 10 \times 2 < 9 \times 3 < 8 \times 4 < 7 \times 5 < 6 \times 6$. Generally, for

$\sum x_i = C$, $\max_{x_i} \prod x_i$ is $\frac{C}{n}$ for all x_i , i.e. the product is maximised if all x are the same. This insight is applied, for example in the Nash (1950) solution to the bargaining problem, which maximises the product of utilities.

¹⁸⁹ Nentjes (1993), p. 31-32.

fore¹⁹⁰, however, in practice economic analysis often has an unfortunate tendency to overlook this second step, making the results biased in a variety of ways.

Having considered how income effects affect the Coase Theorem, there are two more “exceptions” that need to be discussed briefly, before we turn our attention to the nature of transaction costs: voluntary participation and rationality. The problem of voluntary participation was already noted in chapter 2.1, above. Dixit & Olsen (2000) develop a model where each citizen first has to decide whether they will participate in the negotiations over the provision of some public good, and only in the second stage do the negotiations actually take place, at least between those who have elected to join in. Those who have chosen not to negotiate can free ride on whatever is decided, but have no opportunity to influence the outcome. In those cases where M , the minimum number of participants that is necessary in order to produce the good, is small relative to N , the size of the population, they show that in a one-shot game, the good will probably not be produced, even though transaction costs are zero, income effects are ignored and all actors are perfectly rational. When the game is played repeatedly, when many such games are played at the same time, or when there are sub-groups in society who benefit disproportionately from the provision of the public good in question¹⁹¹, the likelihood that the good will be produced increases. Also, when all else fails, there is the possibility of making participation non-voluntary, i.e. of using state compulsion. For many public goods, that is exactly what happens, for example when it comes to street lights or crime prevention. Usually, this is explained by pointing to the transaction costs of organising such things privately, but the Dixit & Olson article shows that public provision of public goods could even be necessary in a world with no transaction costs.¹⁹²

As for the third exception, we return to the observation previously made in section 1.2, that it is important to keep in mind the line between economics and psychology. However, where behavioural economics provides insight that can be modelled in the usual economic manner, it

¹⁹⁰ Chapter 1.2, above.

¹⁹¹ Particularly when the size of this sub-group is greater than M .

¹⁹² Anderlini & Felli (2006) reach a similar conclusion using a different approach.

can be useful to take such ideas into consideration.¹⁹³ One such insight that is particularly relevant to the understanding of the Coase Theorem is the so-called endowment effect.¹⁹⁴ The nature of this problem goes back to the common accusation that an economist is someone who “knows the price of everything and the value of nothing”.¹⁹⁵ In fact, economists have had a great deal to say about the notion of value, starting with Adam Smith himself.¹⁹⁶ Using arguments similar to those underlying the Coase Theorem, they tended to conclude that the value in use of a product, i.e. the value it had to someone using it, would tend to be equal to the value for which it was traded, i.e. its price. After all, in case of a difference between these two measures of value, a profit can be made by selling the good instead of using it.

Unfortunately, more recent research has shown that the value for which a good is traded is less invariant with respect to the initial distribution of property rights than originally assumed. The conventional wisdom is that a decision to buy or sell is based on a comparison between the price that is offered or asked and the value in use of the good in question, a value that would necessarily be the same regardless of whether we are talking about buying or selling. However, experimental research has shown that there is often a significant difference between the willingness to pay (WTP), i.e. the price a person would be willing to pay for something, and the willingness to accept (WTA), i.e. the price for which that same person would be willing to sell that same good if they already owned it. To see how this affects the Coase Theorem, let us return to the Hanrahan case that was discussed before. The Coasian prescription here is to define emission rights and organise a market for them.¹⁹⁷ In so doing, the Theorem claims that, ignoring the provisos that were already discussed, it is irrelevant whether you start by awarding these rights to the company or to the inhabitants of the surrounding area. Assuming perfect rationality for the company, its trading curve in this market is the same in both cases; it depends on the company’s marginal abatement costs, i.e. the costs of reducing its emissions of pollutants.¹⁹⁸ The curve of the local residents, on the other hand,

¹⁹³ This was exactly the intention of the creators of prospect theory, Kahneman & Tversky (1979): to create an alternative for the Von Neumann & Morgenstern (1944) model of decision making under uncertainty that was itself also an economic model, only one with more solid psychological footing.

¹⁹⁴ Cf. Coursey, Hovis & Schulze (1987) and Curran (1999).

¹⁹⁵ This is a misquote of something Oscar Wilde wrote. He made the accusation, in slightly a different wording, of cynics in Act 3 of *Lady Windemere’s Fan*. Cf. Liberman (2006).

¹⁹⁶ Book I of the *Wealth of Nations*, Smith (1776). Cf. also Marx (1867), *passim*, Menger (1871), Chapter III and Marshall (1890), Book III, Chapter VI.

¹⁹⁷ Cf. Geerdink & Stauvermann (2007) for a formal derivation of this result.

¹⁹⁸ Whether this is a demand or a supply curve depends on the initial distribution of emission rights. By allowing for negative prices and quantities, the company’s curve can be represented similarly to an ordinary demand curve, i.e. downward sloping, as long as the right in question is defined as a right to pollute. If it were defined as a right to clean air, the company would be the one supplying.

would have a discontinuity at the place where it crosses the vertical axis, where supply and demand meet. When the residents are suppliers, their curve would likely be less steep than when they demand emission rights.¹⁹⁹ Different starting conditions are represented by shifting the curves, and it is easy to see that this affects the locus of the discontinuity and the amount of pollution that is ultimately allowed. Formally:

$$Q_C = Q_{0,C} - \alpha P$$

$$Q_R = Q_{0,R} + \beta P \text{ if } Q_R > 0$$

$$Q_R = Q_{0,R} + \gamma P \text{ if } Q_R < 0$$

Assuming $\beta = \gamma$ yields an ordinary set of supply and demand curves, while in the presence of an endowment effect, instead we have $\beta > \gamma$. In these equations, Q_0 represents the amount each party would wish to trade if the price were zero. In the case of the company, that means buying all the emission rights that have been allotted to the residents, putting it in a position of zero abatement. In the case of the residents, a price of zero would have them buy all the emission rights that have been initially allotted to the company, putting it in a position of maximum abatement. So $|Q_{0,R}| + Q_{0,C}$ equals the total amount of emission rights that have been created, with the first term representing the rights allocated to the company, and the second representing the rights initially allocated to the residents. Graphically:

¹⁹⁹ Again, if the right were defined as a right to clean air, this statement would be reversed.

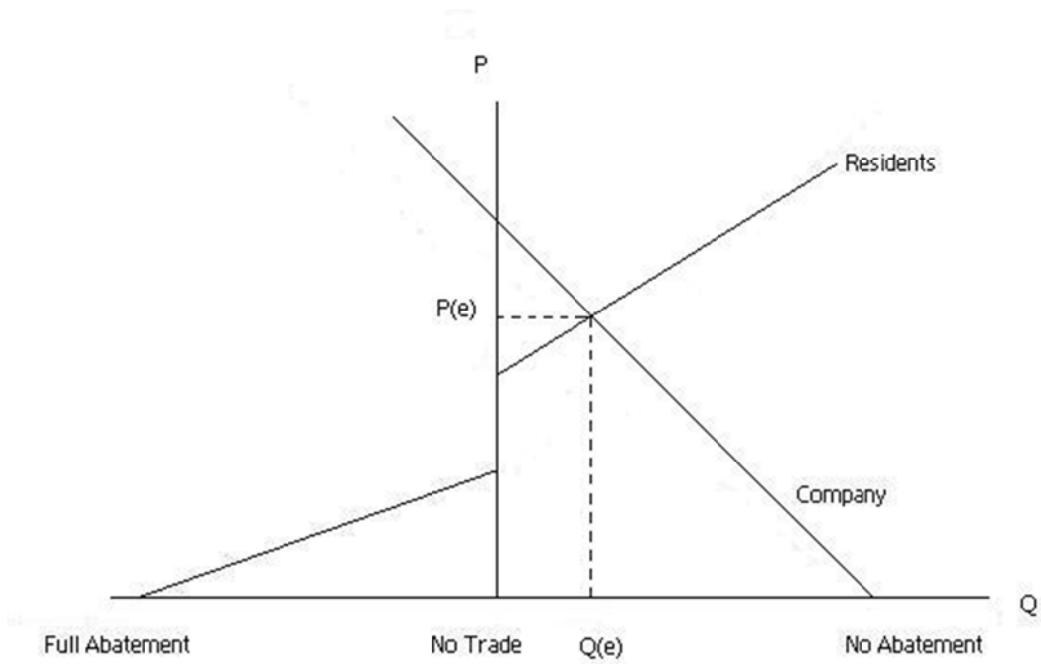


Figure 1: Emissions trading with the residents selling.

Here, the equilibrium price and quantity lie in the right quadrant, meaning that the residents are selling some of their emissions rights to the company. Since they are selling some of their initial endowment, they are relatively price sensitive.²⁰⁰ Graphically, a change in the initial allotment translates into a move of the vertical axis. Maintaining the same curves, but moving the vertical axis, and with it the locus of the discontinuity in the residents' curve, we can construct an outcome where the residents are buying. Comparing the two figures shows that $Q(e)$, the equilibrium amount of trade, and therefore also the equilibrium amount of pollution, lies further to the right in the second figure, illustrating the fact that the Coase Theorem is untrue in the presence of an endowment effect. It should be noted that otherwise, everything that was said about the Theorem still holds true in this case. Moving the axis has no effect on the amount of pollution, except when the axis is moved so far that the sellers become buyers, and vice versa. Barring such a drastic shift, the equilibrium price of an emission right does not depend on the initial endowment, but the total income transfer from the buyers to the sellers, $P(e) \cdot Q(e)$, does.

²⁰⁰ Note that, because of the algebraic form chosen above, extending the right-hand side of the residents' curve into the left quadrant has that line dissect the horizontal axis at the same point as the left-hand side of the resident's curve, i.e. at $Q_{0,R}$, the level of full abatement.

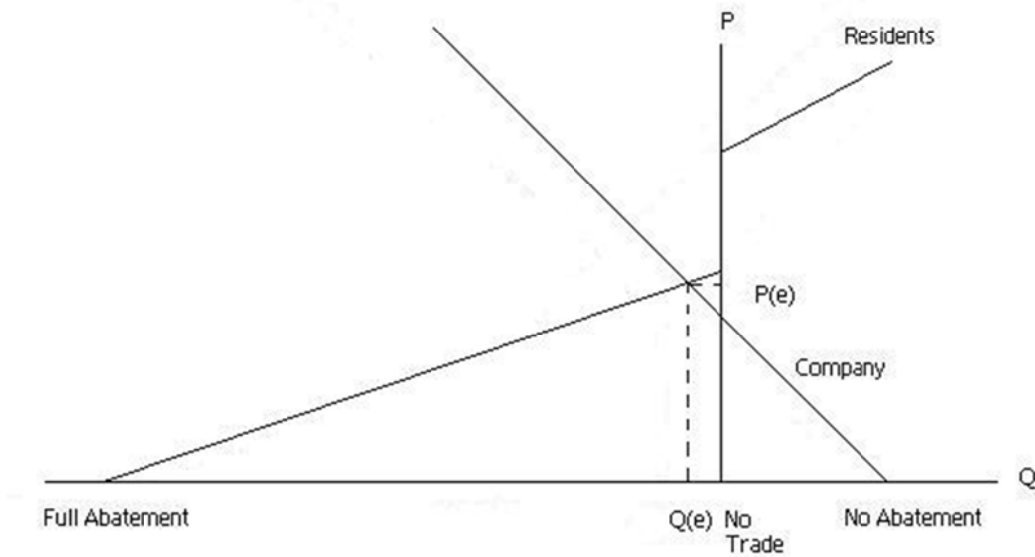


Figure 2: Emissions trading with the residents buying.

At this point, we have covered all the main reasons why the allocation of resources should not be “invariant with respect to legal rules of liability”²⁰¹, except for one: transaction costs, which will be the subject of the next section.²⁰²

²⁰¹ Zerbe (1980), p. 84.

²⁰² A related matter, which is not discussed in the present thesis, is what a world of zero transaction costs actually looks like. It turns out that this thought experiment is much less straightforward than originally imagined. Cf. for example Barzel (1982) and Hsiung (1999).

2.3 The Nature of Transaction Costs

Given that we have already seen that transaction costs – perhaps by definition – explain all contractual incompleteness, as well as that transaction costs play a central role in the world of the Coase Theorem, it is only fitting that they should have their own section. In the following pages we will discuss how various authors, from Commons and Coase to Williamson and Hart, have defined transactions and transaction costs, what the commonalities in their approaches are, and where they might have gone wrong.

It should be clear that, since transaction costs are essentially nothing more than the costs of transacting, any discussion of their nature has to begin by considering the definition of a transaction. As always when constructing a model, it is important to choose such a definition not only based on semantic and lexicographic considerations, but rather with the model already at the forefront of one's mind.

Starting with some dictionary definitions, we find that the main Dutch dictionary, Van Dale, equates a transaction with the contract that governs it.²⁰³ For a variety of reasons, that is inappropriate for present purposes. After all, the model also concerns transactions that are imperfectly governed by contracts, as in chapter 2.1, or transactions that are carried out inside a firm, using hierarchy as a governance mechanism instead of a contract. So while it may be useful to construct a definition by analogy to certain legal principles, the definition of a transaction cannot reference legal concepts, directly, since it is exactly the law that we are trying to understand.

The online version of Merriam-Webster gives just such a definition. It says: “[A transaction is] an exchange or transfer of goods, services, or funds”.²⁰⁴ Since an exchange is nothing more than a combination of two or more transfers²⁰⁵, we can omit that part of the defini-

²⁰³ Van Dale (1992), Vol. 3, p. 3146.

²⁰⁴ <http://www.merriam-webster.com/dictionary/transaction>, definition 1a. (Last visited on July 22, 2008.)

²⁰⁵ <http://www.merriam-webster.com/dictionary/exchange>, definition 2b: “reciprocal giving and receiving”. (Last visited on July 23, 2008.) Clearly such reciprocity becomes more complex in a transaction involving three or more parties. Cf. for example the well known problems involving consideration in such contexts, Restatement (second) of Contracts § 9 or the English House of Lords ruling of *McEvoy v Belfast Banking Co.*, [1935] AC 24.

tion without loss of generality.²⁰⁶ Instead, the questions remain what it is, exactly, that is being transferred, and between whom. To a non-lawyer, it may appear as if the former question is already adequately addressed in the definition itself, but that does not do justice to the legal technicalities of property law. As before, we have to decide to what extent such legal notions should be admitted to the definition.

In law, there are two distinct ways in which someone can “have” a good. And even though the dictionary definition of transfer carefully distinguishes between the two, it does admit both the possibilities of ownership and the possibility of possession.²⁰⁷ While the details of the definition of these concepts differs from one jurisdiction to another, and especially between the civil law, where the distinction is quite clear, and the common law, where this is less so, generally speaking the right of ownership is the “most absolute right a person can have with respect to a good.”²⁰⁸ Possession, on the other hand, depends on physical control²⁰⁹, which is an essentially non-legal concept, although a good can come to be in the physical control of someone other than the owner as a result of the operation of a variety of legal rules, such as the law on leasing²¹⁰ or the law on agency.²¹¹ The reason why this distinction matters is simple: in a vertically integrated company, ownership is not transferred between the different stages of the production process, but possession is. This implies that defining a transaction as a transfer of ownership means that transactions only take place in a market, or possibly in hybrid structures that closely resemble markets. If the purpose is to discuss the choice between “make” and “buy”, that might not be very convenient.

²⁰⁶ Indeed, in his inaugural lecture of September 30, 2008, prof. Steenge described the classical Coasian problem of air pollution by a factory, which was also discussed in chapter 2.2, above, as a “unilateral transaction”. This is obviously correct, since such pollution involves a transfer of a good (or rather: a bad) from the factory to the community. Steenge (2008).

²⁰⁷ <http://www.merriam-webster.com/dictionary/transfer> refers to “convey” in definition 1a, and to “possession or control” in definition 2. (Last visited on July 23, 2008.) The Restatement (First) of Property also provides a definition of transfer, in § 13, saying: “The word “transfer” (...), when applied to interests in land or in a thing other than land, means the extinguishment of such interests existing in one person and the creation of such interests in another person.” This ignores the possibility of purely factual control, allowing only the spectrum of different rights that a person can have with respect to a good, be they possessory rights or ownership.

²⁰⁸ Paraphrasing and translating article 544 of the French Code Civil, article 5:1(1) of the Dutch Civil Code and § 903 of the German BGB.

²⁰⁹ Cf. § 854 of the German BGB. Note that Dutch law, art. 3:107 c.a. Civil Code, distinguishes between someone who possesses a good “for himself” and someone who possesses a good “for another”. The former would usually be the owner of the good (art. 3:112 BW), but can also be the possessor non-owner when there is a defect in the transfer of ownership (art. 3:118 BW), or in certain cases of theft (art. 3:121 BW). This distinction, which goes back to Roman law (Book 41 of the Digesta), appears in many legal systems (eg. Article 2228 CC). However it is not very interesting for present purposes, and therefore omitted from the text.

²¹⁰ Eg. the entirety of the Restatement (Second) of Property, which deals with the legal relations between landlords and tenants.

²¹¹ Cf. Restatement (Third) of Agency § 8.05, which discusses the use of the Principal’s property by the agent.

Taking this into consideration, it appears that the notion of a transaction can be defined in a way that admits certain legal terms, or in a way that decidedly seeks to avoid them.²¹² The choice between these approaches cannot be made on *a priori* grounds alone. Instead, what matters is which approach makes the model work better. However, no matter how the concept of a transaction is defined, any model based on Coase's work is going to distinguish between transaction costs and all other costs²¹³, in such a way that all costs can be assigned to one of these categories, so that the ultimate institutional choices can be made by comparing the benefits and costs of the different alternatives.

The first economist to consider transactions as “the basic unit of analysis”²¹⁴ was in fact John R. Commons, in 1924. In his book *The Legal Foundations of Capitalism*, he defined a transaction as “two or more wills giving, taking, persuading, coercing, defrauding, commanding, obeying, competing, governing, in a world of scarcity, mechanism[s] and rules of conduct.”²¹⁵ However, in the chapter of the book that is devoted to the concept of a transaction, his focus is mainly on developing the Hohfeldian framework into a basis for institutional economics²¹⁶, without even making clear whether or not he means to distinguish between a transaction and the contract that governs it.²¹⁷

Ten years later, Commons returned to the subject. Again, he argued for a bigger role for law in economic analysis, saying “Transactions (...) are not the “exchange of commodities”, in the physical sense of “delivery”, they are the alienation and acquisition, between individuals, of the *rights* of future ownership of physical things, as determined by the working rules of society.”²¹⁸ In other words, a transaction is the transfer of ownership, at least when it concerns goods rather than services. The consequence of such a definition is that a transaction does not occur when the good is transferred within a single firm. In order to consider this question more carefully, Commons introduced two more categories of transactions, in addition to the bargaining transaction that involves a transfer of ownership: managerial transac-

²¹² Allen (1999), for example, discusses two different possible definitions, a general one: “the costs [of] establishing and maintaining property rights” (p. 898), and a more legal one: “the costs resulting from the transfer of property rights” (p. 901). The first he associates with what he describes as the “property rights” version of transaction costs economics, and the second with the “neo-classical approach”.

²¹³ For convenience, we will usually call these costs “operating costs”.

²¹⁴ Williamson (1985), p. 41.

²¹⁵ Commons (1924), p. 7. Cf. similar statements in art. 3:33 Dutch civil code and art. 1108 CC.

²¹⁶ Commons (1924), p. 65-132. The Hohfeldian framework is briefly discussed in chapter 2.1, above.

²¹⁷ In discussing the concept of will, for example, on p. 69-74, in order to explore somewhat the definition given on p. 7, as quoted above, one could very easily get the impression that the transaction consists of a meeting of wills, instead of in carrying out that what is willed.

²¹⁸ Commons (1934), p. 58. Emphasis in the original.

tions and rationing transactions. He defines the latter as “the negotiations of reaching an agreement among several participants who have authority to apportion the benefits and burdens to members of a joint enterprise.”²¹⁹ Obviously, such negotiations are only distinct from a bargaining transaction if they occur *ex post*, i.e. after the “joint enterprise” has already been created. The consequences of dividing up the gains from trade *ex post* instead of *ex ante* will be explored further below.

As for the managerial transactions, Commons gives no clear definition, but we can distil from his discussion of the subject that what he has in mind is the transfer of a good within a hierarchy, under the direction of a superior.²²⁰ As he correctly notes, such a transfer is still a manner of exchange, and is fraught with similar difficulties as a bargaining transaction. To illustrate this point, Commons quotes a management scholar of his time, who divides the job of a manager into eight distinct tasks: Observing, Evaluating, Conceiving, Analyzing, Contriving, Directing, Teaching and Inducing.²²¹ A moment’s reflection makes clear that each of these is necessary in bargaining transactions and managerial transactions alike, although the decision as to make or buy will certainly affect the nature of these tasks, as well as their relative importance.

While Commons provided us with a useful analysis of the notion of a transaction, he did not develop any rigorous models to link the costs of transacting with governance choices. This step was taken by Ronald Coase in his 1937 article *The Nature of the Firm*, where he said, quite obviously in retrospect, that “The main reason why it is profitable to establish a firm would seem to be that there is a cost of using the price mechanism.”²²² In the remainder of that article, he sketches out some of the costs of organising transactions through a firm or through the market, such as “discovering what the relevant prices are”²²³ and “the costs of negotiating and concluding a separate contract for each exchange transaction”²²⁴ on the one hand and “decreasing returns to the entrepreneurial function”²²⁵ and “waste of resources”²²⁶ on the other hand. However, he, too, like Williamson after him, imagines transactions as occurring within firms as well as between them: “A firm becomes larger as additional transac-

²¹⁹ Commons (1934), p. 67-68.

²²⁰ Commons (1934), p. 64-67.

²²¹ Commons (1934), p. 67, quoting Dennison (1924).

²²² Coase (1937), p. 390.

²²³ *Ibid*, p. 390.

²²⁴ *Ibid*, p. 390-391.

²²⁵ *Ibid*, p. 394.

²²⁶ *Ibid*, p. 395.

tions (...) are organised by the entrepreneur and becomes smaller as he abandons the organisation of such transactions.”²²⁷

Naturally, in his first endeavours to develop a theory of transaction costs, Coase was unable to treat every aspect of the model exhaustively. That is why his 1937 article does not contain a definition of transaction, much less a defence of the proposition that transaction can be said to occur within a firm. Neither does his later work, which is concerned less with transactions than with rules of liability.²²⁸ This left the task of discussing the nature of transactions to his principal heir, Oliver Williamson. In 1981, Williamson wrote:

*“A transaction occurs when a good or service is transferred across a technologically separable interface. One stage of activity terminates and another begins. With a well-working interface, as with a well-working machine, these transfers occur smoothly. In mechanical systems we look for frictions: do the gears mesh, are the parts lubricated, is there needless slippage or other loss of energy? The economic counterpart of friction is transaction cost: do the parties to the exchange operate harmoniously, or are there frequent misunderstandings and conflicts that lead to delays, breakdowns, and other malfunctions? Transaction cost analysis supplants the usual preoccupation with technology and steady-state production (or distribution) expenses with an examination of the comparative costs of planning, adapting, and monitoring task completion under alternative governance structures.”*²²⁹

The main problem with this definition is the notion of “technologically separable”. This could be construed as a question about modularity, or task definition.²³⁰ “How should tasks be assigned to different departments and individuals?” Obviously, it is in theory possible to do this in an infinity of ways, but does that necessarily make such tasks “technologically separable”? To put it simply: if two tasks are separable in the sense that they can theoretically be carried out by two different people using different equipment, but only at a prohibitively high cost, such that no sane manager would ever dream of separating them in this way, is the “transfer” between the two tasks still a transaction under Williamson’s definition? Or does he mean to

²²⁷ *Ibid.*, p. 393. Cf. Allen (1999), p. 894-896. Similarly, in his 1991 Nobel Prize lecture, Coase discussed transaction costs as follows: “[T]here [are] costs of using the pricing mechanism. What the prices are has to be discovered. There are negotiations to be undertaken, contracts have to be drawn up, inspections have to be made, arrangements have to be made to settle disputes, and so on. These costs have come to be known as transaction costs. Their existence implies that methods of co-ordination, alternative to the market, which are themselves costly and in various ways imperfect, may nonetheless be preferable to relying on the pricing mechanism”. Coase (1991)

²²⁸ Cf. Chapter 2.2, above.

²²⁹ Williamson (1981), p. 552-553. Incidentally, the definition given in the quoted text is repeated in the glossary of the website of the Ronald Coase Institute, at <http://www.coase.org/nieglossary.htm>. (Last visited at 28 August 2008.)

²³⁰ Cf. Baldwin (2008), Suematsu & Makabenta-Ikeda (2006).

take a more pragmatic approach and consider something a transaction only if at least someone somewhere has in fact separated the tasks?

To see how this matters, consider an automobile factory in the time of Henry Ford and at present day. The former is clearly what Williamson had in mind when he wrote the section quoted above; it involves clearly recognisable transfers. Worker after worker adds to the product before passing it on to their colleague who is next on the assembly line. Each time the product is transferred, a transaction takes place. Compare this picture to a present day factory, where robots do most of the actual assembly, while the employees of the company are simply there to supervise the machines. Taking Williamson's definition literally, the number of transactions has not changed, since each robot could conceivably be owned and run by a separate entrepreneur. Presumably, that is not what he meant, but then, how many transactions do take place inside a car factory?²³¹

Clearly, the most realistic description of such a factory is that the employees work as a team, working together to produce cars using the factory as a tool.²³² But this means that no actual transactions take place during the production process. In other words, the number of transactions that make up a production chain would not depend on the governance choices that are being made, but it would depend rather directly on technological developments. More troublesome is that the number of transactions also ends up depending on the range of governance choices that are being applied in the market. In the car industry, where a relatively small number of conglomerates compete in a highly sophisticated manner, there is no reason to believe that any set of tasks that can sensibly be described as separable, has not in fact been separated somewhere. So the number of transactions can be studied by looking at the task assignments in different companies. However, in markets that are being controlled by a monopolist, the only tasks that end up being described as "separable" are those that have, in fact, been separated. It follows that a literal reading of what is a "technologically separable interface" leads to too many transactions, and a pragmatic reading leads to having too few.

Finally, after Commons, Coase and Williamson, the most recent development in the history of transaction cost economics is the development of incomplete contracting theory, in

²³¹ Consider this definition by Hendrikse (2003), who takes the notion of a "basic unit of analysis" literally and defines a transaction as "the largest unit of economic activity which cannot be subdivided and performed by several different people." (p. 469)

²³² Cf. Alchian & Demsetz (1972) for a discussion of the unique difficulties of modelling teamwork analytically.

a number of papers, by Grossman, Hart and Moore (hereafter: GHM).²³³ While this model will be discussed in detail in section 2.5, it is appropriate to discuss their ideas about the nature of transactions here.

What stands out about GHM is that they have little difficulty applying the concept of transactions equally within firms as between them, exactly because they assume that very little changes as a result of vertical integration. “[G]iven that it is difficult to write a complete contract between a buyer and seller and this creates room for opportunistic behavior, the transactions cost-based arguments for integration do not explain how the scope for such behavior changes when one of the self-interested owners becomes an equally self-interested employee of the other owner.”²³⁴ In response, they deal with the question of integration as one of residual control rights; in each case, many of the same transaction costs are present, but in the case of integration, there is only one owner who gets to decide what should happen if the state of the world that is realised is one for which no action set is specified.²³⁵ Regardless of the governance choices made, there are two assets, each with an actor controlling them, where the first is used to produce an intermediate product, which is then transformed by the second asset into a finished product. It follows that the implicit definition of transaction here is one similar to Williamson’s: a transfer of control between two stages in the production process, however defined.

In the end, the question of what exactly changes as a result of integration, of what exactly the difference is, if any, between bargaining transactions and managerial transactions, is an empirical one, and one that is far from resolved.²³⁶ For now it is probably best to continue to work with three categories of costs: production costs, which are mostly exogenous to the model, and which vary with governance choices only as a result of economies of scale and scope, management costs, which are the costs of organising management transactions, in other words: the costs of organisational failure²³⁷, and transaction costs narrowly defined, which are the costs of market failure. With such a distinction, the defining characteristic of bargaining

²³³ The most important ones are Grossman & Hart (1986), Hart & Moore (1988, 1990, 1999), Hart (1995) and Hart, Shleifer & Vishny (1997).

²³⁴ Grossman & Hart (1986), p. 692.

²³⁵ Insufficient state contingency, which we have defined based on *ex ante* Pareto efficiency, will often imply *ex post* inefficiency as well. In the GHM model, all cases of *ex post* inefficiency lead to renegotiation.

²³⁶ Cf. Cheung (1983)

²³⁷ Apparently, this term was first used in Williamson (1971). Cf. Pessali (2006). For a general discussion of market failure and government failure, cf. Winston (2006).

transactions, as with Commons, is that they involve a transfer of ownership in the legal sense.²³⁸

Before continuing to the detailed description of the two main modern approaches to transaction cost economics, Williamson's model and the GHM models, it is important to stop for a moment and consider the implications of what has been said so far. What are these transaction costs, when you get right down to it? Can we really classify all costs incurred in the economy as either production costs or transaction costs, broadly defined?

In the 1980s, Wallis and North set out to calculate the relative size of each of these two categories in the economy.²³⁹ For the US over the period of 1870 to 1970, they calculated a steadily rising share of the transaction sector, increasing from 24,19% of GNP in 1870 to 46,66% in 1970.²⁴⁰ This exercise was later repeated by a number of other authors, with similar findings.²⁴¹

While there are many interesting things to say about these results²⁴², what concerns us here is which kinds of costs these authors did or did not include in their tallies. Taking the later Dollery & Leong study as an example²⁴³, we have such things as finance, insurance and real estate, wholesale and retail trade and two categories they call public and private non-transaction industries, i.e. the transaction costs incurred by the managers, lawyers and accountants who work in such industries. On the government side, they included the entire core of the government, reflecting the fact that the social contract²⁴⁴ is the ultimate transaction costs saving governance structure, as well as the police and the military.²⁴⁵ Finally, they count the postal services in its entirety as a transaction industry.²⁴⁶

²³⁸ Or, in the case of services, a service being provided by one legal person to another.

²³⁹ Wallis & North (1986, 1988).

²⁴⁰ Wallis & North (1986), table 3.13, which is replicated in table 1 in Wallis & North (1988).

²⁴¹ Cf. Dagnino-Pastore & Farina (1999) and Dollery & Leong (1998), concerning Argentina and Australia, respectively.

²⁴² One could ponder, for example, that while a reduction in transaction costs is supposed to lead to an increase in social welfare, in every one of these studies the share of the transaction sector in the economy increases over time.

²⁴³ Dollery & Leong (1998).

²⁴⁴ Hobbes (1651), Locke (1689), Rousseau (1762), Rawls (1971). In the absence of transaction costs, the only purpose of government would be to enforce contracts. Cf. Coase (1974), which discusses the transaction costs associated with lighthouses.

²⁴⁵ In the absence of transaction costs, people could costlessly contract for their personal safety, and such contracts could costlessly be enforced by the courts. To be sure, that would likely involve paying someone in return for their promise not to set your house on fire. However, while such a contract might be unfair, it is not in any

All things considered, it turns out that it is relatively unimportant how certain costs are classified, as long as one is careful to put everything in its proper place. In the Williamson and GHM models, which will be discussed in the next two sections, the absolute level of transaction costs is irrelevant, since the only question that these models seek to answer is the question of which alternative governance solution maximises efficiency. As long as one focuses on comparing discrete structural alternatives²⁴⁷, a degree of common sense will usually suffice to make sure that no costs are overlooked, even in the absence of a generally agreed and unambiguous definition of transaction costs.

way irrational from an economics point of view. For an interesting experimental study of how fairness enters into economic decision making, cf. Fehr, Kirchsteiger & Riedl (1993). Concerning the proper classification of “protective services”, cf. also Allen (1999), p. 911.

²⁴⁶ Clearly postal services can be privatized. The reason why the authors argue that they fall in the transaction sector is that the purpose of the mails is to allow the transfer of information. Without transaction costs, all information would be costlessly available, so there would be no need for postal services. Given this rationale, it is difficult to understand why the authors classify communications as a non-transaction sector. Cf. Dollery & Leong (1998), p. 209, table 1.

²⁴⁷ Williamson (1991).

2.4 Williamson

The key contribution that Oliver Williamson made to transaction cost economics is that he took the Coasian hypothesis that the manner in which transactions are governed depends on the characteristics of the transaction, and attempted to define which transaction characteristics are relevant and how different transactions are governed. The present section discusses his three transaction characteristics: asset specificity, frequency and uncertainty²⁴⁸, as well as his operationalisation of governance structures. We will also briefly touch on the issue of probity, which Williamson offered as a fourth transaction characteristic when he discussed public bureaucracies in 1999.

2.4.1 Characteristics of Transactions

According to legend²⁴⁹, in September 1919, General Motors made a highly incomplete contract with Fisher Body, a manufacturer of car bodies.²⁵⁰ Under the terms of this contract, for the next 10 years, GM would exclusively buy its car bodies from Fisher, at a price of cost + 17,6%.²⁵¹ That is, GM was obligated to buy its closed care bodies from Fisher to the extent that Fisher was able to meet their demand, but Fisher was free to use any excess capacity it might have to sell to GM's competitors. By 1925, GM was becoming increasingly unhappy with this arrangement. In an ever growing market, Fisher was taking an ever increasing share of the total profit, while resisting any attempts to increase efficiency where that would have the effect of making it depend on GM in the same way as GM depended on Fisher. For example, the companies discussed the construction of a new factory for Fisher. GM asked that they

²⁴⁸ Cf. Milgrom & Roberts (1992), p. 30-33, who distinguish five "dimensions of transactions": specificity, frequency and duration, complexity and uncertainty, the difficulty of measuring performance and the connectedness of the transaction to other transactions involving other parties. Some of these we will return to later.

²⁴⁹ This case was first discussed in Klein, Crawford & Alchian (1978), pp. 308 – 310, and has since become the customary example of a hold-up. (Eg. Williamson (1985), p. 114-115, Klein (1988) and, more recently, Coase (2000, 2006).) Since then, many of the legal facts of the case have been disputed by Goldberg (2007). However, since the veracity of this account is fairly unimportant for expositional purposes, it is discussed here regardless.

²⁵⁰ In that same contract, GM also took a 60% equity share in Fisher Body, but the latter continued to be run as a separate company, seeking to maximise its own profits, since at least until 1924, those shares were placed in a voting trust. Coase (2000), p. 22-23, discusses the organizational arrangement that was set up in 1919 to protect Fisher's independence, based on the court documents from the 1950s US antitrust case against DuPont De Nemours, GM and a number of other companies, which resulted from the attempted takeover by Du Pont of GM. *US v. Du Pont*, US District Court for the Northern District of Illinois, December 3, 1954, 126 F. Supp. 235, and October 2, 1959, 177 F. Supp 1.

²⁵¹ Coase (2000), p. 20. On p. 19-21, Coase discusses the histories of the two companies.

build it in Flint, Michigan, the home town of GM. But Fisher preferred to build it in Detroit, where it would have ready access to a number of customers. Finally, in 1926 GM acquired complete control over Fisher Body, choosing hierarchy over market as a governance structure.

For Oliver Williamson, the only reason for vertical integration of this kind is to deal with the kind of opportunistic behaviour that would otherwise ensue where one party is “stuck” to the other the way GM was stuck to Fisher.²⁵² As it appears, GM was stuck to Fisher mainly because of the ill-considered terms of their contract, which is a possibility that does not normally fall within the scope of economic theory.²⁵³ Fisher Body, on the other hand, quite rightly avoided the similar mistake of locking itself into a permanent relationship with GM by building its factory next to the GM facilities.

In Williamson’s terminology, the problem with building your factory right next to your main customer is asset specificity, or, to be precise, site specificity.²⁵⁴ Quoting Williamson (1991):

“Asset specificity has reference to the degree to which an asset can be redeployed to alternative uses and by alternative users without sacrifice of productive value. Asset specificity distinctions of six kinds have been made:

1. *Site specificity, as where successive stations are located in a cheek-by-jowl relation to each other so as to economize on inventory and transportation expenses;*
2. *Physical asset specificity, such as specialized dies that are required to produce a component;*
3. *Human-asset specificity that arises in learning by doing;*
4. *Brand name capital;*

²⁵² In Williamson’s jargon, there is a fundamental transformation when a specific investment causes a situation of *ex ante* competition to transform into one of *ex post* bilateral monopoly. Cf. Williamson (1985), p. 61-63. This is a topic that we will return to in later chapters, below.

²⁵³ Behavioural economics, which does concern itself with behaviour that is, strictly speaking, irrational, is probably better considered a branch of psychology. Another possibility is to construct a model that shows that behaviour that seems irrational, is, in fact, a rational response to a consistent preference set, eg. Becker & Murphy (1988). As for carelessness, it can only be explained by economists along the lines that the costs of being more careful outweighed the benefit of avoiding the damage done, an explanation that does not seem to be applicable here.

²⁵⁴ Williamson (1991), p. 281. Cf. also Joskow (1985, 1987, 1988).

5. *Dedicated assets, which are discrete investments in general purpose plant that are made at the behest of a particular customer;*
6. *Temporal specificity, which is akin to technological nonseparability and can be thought of as a type of site specificity in which timely responsiveness by on-site human assets is vital.*^{255 „256}

Finding further examples of each category is left as an exercise for the reader. The important thing to note here is simply that Williamson's notion of asset specificity goes well beyond the physical asset specificity and site specificity that are most commonly envisaged.

The second characteristic of transactions is the frequency of the transaction.²⁵⁷ How often does the good or service in question get transferred? Since obviously creating a governance structure to police the transaction has costs as well as benefits, the pertinent question is over how many transactions the fixed portion of these costs can be divided. After all, some transaction costs occur for every instance of the transaction, for example the costs of making sure that you received what you purchased, while others only occur the first time you transact, such as the costs of finding someone to transact with. The costs of setting up a governance structure, on the other hand, are almost completely fixed. The result is that, all else equal, a higher frequency of transacting causes the cost-benefit analysis to shift more towards integration.

Something else that bears mentioning at this point is the connection between frequency and transaction costs. Although the underlying "Folk Theorem" is much older²⁵⁸, the basic argument comes from Robert Axelrod's 'The Emergence of Cooperation among Egoists', which employed a tournament pitting different strategies for a repeated Prisoners' Dilemma game against each other to study which strategy would prove superior.²⁵⁹ The Tit-for-Tat strategy, which begins with "cooperate" and subsequently repeats whichever move was made by the opponent in the previous round, came out as the winner. Since a game with two Tit-for-Tat players leads to cooperation in every round, this was truly a case of "cooperation

²⁵⁵ Masten, Meehan & Snyder (1991).

²⁵⁶ Williamson (1991), p. 281-282. Reference in the original. Asset specificity is also discussed in Williamson (1985), p. 52-56.

²⁵⁷ Cf. Williamson (1985), p. 60-61.

²⁵⁸ Cf. Gintis (2000), chapter 8.5. An early proof is by Rubinstein (1979).

²⁵⁹ Axelrod (1981), Cf. Siegfried (2006), chapter 4.

among egoists”.²⁶⁰ Even though later research from evolutionary game theory has nuanced this result somewhat²⁶¹, the principle still stands: repeated interaction makes it easier to use reputation instead of formal governance structures, reducing somewhat the correlation between frequency and degree of vertical integration that is hypothesised by Williamson’s model.²⁶²

Thirdly and finally, transactions differ with respect to the amount or degree of uncertainty involved.²⁶³ Indeed, sometimes uncertainty can be stated as a measurable amount, in which case economists usually call it “risk”. This is the perspective that was adopted earlier, where uncertainty was modelled as a range of possible states of the world, each with its own probability: The greater the number of states of the world whose probability exceeds some threshold minimum, and the more even their probability distribution, the greater the risk.²⁶⁴ This is a situation where, in some real sense, rational behaviour is still possible, where decision making paradigms can be applied to make a recommendation that can be said to be “correct” or “incorrect”.²⁶⁵ True uncertainty, on the other hand, cannot be quantified in this manner.²⁶⁶

In Williamson’s model, uncertainty matters because it causes incompleteness. Remember, when writing a contract, parties only write action sets for those states of the world where doing so has benefits that exceed the costs of negotiating and writing the contract to cover that state of the world. All else equal, greater risk means that there are more states of the world that “matter”, that need to be covered in the contract, which means that the total costs of negotiating and writing will be higher – costs that can be avoided by vertically inte-

²⁶⁰ Cf. also Fudenberg & Maskin (1990), Greif (1991) and Dixit (2004), for applications in advanced game theory, economic history and the study of alternative modes of governance, respectively.

²⁶¹ Cf. for example Imhof, Fudenberg & Nowak (2005).

²⁶² Unless, of course, one were to see the use of reputation mechanisms itself as a form of vertical integration. Cf. below in chapter 2.4.2 for a discussion of the nature of governance structures in Williamson’s model.

²⁶³ Cf. Williamson (1985), p. 56-60. For a general treatment of uncertainty, cf. Cowell (1986), Chapter 10.

²⁶⁴ As the recent discussion about the Large Hadron Collider in Geneva made clear, virtually anything has a nonzero probability of happening. Cf. US District Court for Hawaii, *Sancho v US Department of Energy*, September 26, 2008, 2008 WL 4370009, yet to be published in the F.Supp.2d. (Order granting motion to dismiss for lack of subject matter jurisdiction) and *Der Spiegel*, 29.08.2008, ‘Gericht weist Eilantrag gegen Superbeschleuniger ab’ (Reporting that the European Court for Human Rights declined to allow an expedited procedure or grant interim measures in a case started by a group called LHC Concern. Cf. <http://lhc-concern.info> (last visited on October 11, 2008), for the group’s statement of reasons.)

The risk involved in a certain project can be gauged by reference to the number of states of the world that can be distinguished meaningfully, but only if those outcomes that are extremely unlikely are excluded. Otherwise, the number of different possible states of the world would always be infinite.

²⁶⁵ The classical treatment of utility maximisation under uncertainty is Von Neumann & Morgenstern (1944), section 3.3.

²⁶⁶ The distinction is due to Knight (1921), cf. particularly his chapter VII.

grating – while the temptation will be greater to ignore a state of the world that really ought not be ignored, leading to problems *ex post*. Uncertainty in the Knightian sense implies that there are chunks of state space where the parties are ignorant about the characteristics of the states of the world, as well as their probabilities of occurring. Obviously, such ignorance immediately leads to greater contractual incompleteness.²⁶⁷

It is important to realise that this third transaction characteristic also includes all sorts of information asymmetries.²⁶⁸ Akerlof’s so-called Lemon Problem is characterised by uncertainty about the present: the seller knows whether he is selling a “lemon” or a good car, and the prospective buyer does not.²⁶⁹ The same goes for any case of adverse selection, for example in the insurance market, where insurers are having difficulty distinguishing between good risks and bad risks. Moral hazard, on the other hand, is caused by the foreseen inability to tell

²⁶⁷ When considered formally, the amount of risk is proportionate to the sum of squares of the probabilities of the different states of the world. This means that the amount of risk is highest when there are a lot of states of the world with similar levels of probability. Fixing the number of states of the world at n , for convenience, and assuming a set of states of the world $S = \{s_1, s_2, \dots, s_n\}$ with corresponding probabilities $\{p_1, p_2, \dots, p_n\}$ which are non-increasing in n , as before, and a total payoff for all parties combined of X , which depends on the action set A , if any, for that state of the world. In a low risk situation, considering only the question of obligational incompleteness, p_i decreases relatively quickly as n increases, and an action set is specified for a certain state of the world iff $p_i X_i \geq TC_i$, the cost of writing a contract for that state of the world. As in chapter 2.1, above, this yields a subset of the state space, for which an action set will be defined. Now increasing the level of risk, or

$$\sum [p_i^2 \mid \sum p_i = 1]$$

increasing $\sum [p_i^2 \mid \sum p_i = 1]$, means that p_i will be lower for some of the most probable states of the world, and higher for some of the states of the world that were not very probable before. This might cause some of the more probable states of the world to no longer satisfy the inequality of $p_i X_i \geq TC_i$, while others will be contracted for that did not have an action set specified in the low risk scenario. The net result will normally be that the contract will end up being more complete than before, in the sense that it becomes less likely that a state of the world will be realised that does not have an action set, though it is possible, for sufficiently high TC , that the opposite will happen, or that in fact the contract becomes unfeasible due to prohibitively high transaction costs. An increase in uncertainty, on the other hand, means that the subset of S for which it is completely impossible to specify an action set becomes larger, which necessarily means an increase in incompleteness. Adding the question of the level of state contingency to the story is analytically tricky, but qualitatively the result is the same: an increase in risk can lead to an improvement in contractual completeness, but an increase in uncertainty never can.

$$\sum (X_i - \bar{X})^2 p_i^2$$

Another possible modification is to model risk in more detail, as $\sum (X_i - \bar{X})^2 p_i^2$. Under this specification, a decrease in completeness as a result of an increase in risk is even less likely.

²⁶⁸ Cf. Parkin (2008), p. 444-447 for a basic treatment, and Cowell (1986), chapters 10-12 for a more advanced discussion.

²⁶⁹ Akerlof (1970). Note that this situation can easily be transformed into one of uncertainty about the future: The seller gives a warranty, guaranteeing that he will fix the car if, and only if, it breaks down due to a flaw that already exists at the time of sale. The buyer realises that this does not solve the problem, since he will not be able to distinguish those states of the world that are characterised by a breakdown caused by a pre-existing flaw, and those states of the world that are characterised by a breakdown caused by something else. The only alternatives are for the seller to give an inefficiently generous warranty, or no warranty at all.

Regardless, uncertainty about the present is perfectly acceptable in transaction cost economics, since some form of bounded rationality is always assumed. Cf. Williamson (1985), p. 44-47.

apart future states of the world, which are distinguished by different actions by one of the parties to the contract, for example excessive risk taking. The interesting question, which we will return to later²⁷⁰, is whether such information asymmetries are a sufficient condition for vertical integration, even in the absence of asset specificity. Williamson's model, on the other hand, states that asset specificity, uncertainty and a reasonable frequency are each necessary to explain vertical integration, though not necessarily jointly sufficient.²⁷¹

Finally, it is important to note that Williamson (1999) added a fourth transaction characteristic, specifically for the context of public bureaucracies: probity, meaning "the loyalty and rectitude with which the (...) transaction is discharged."²⁷² Obviously, Williamson writes, this transaction characteristic applies to all transactions, not only public ones.²⁷³ This is true not only because of the extreme importance that he ascribes to the concept, but also because a contrary specification would make the model useless: whether or not something should be a "public bureaucracy" is exactly what the model is meant to decide, so having one of the variables depend on the public or private nature of the transaction makes it circular. Instead, Williamson argues that probity only "surfaces" in "extreme instances".²⁷⁴ Which instances that might be – other than the foreign affairs transaction that he discusses in that 1999 article – remains unclear.

In order to clarify, Williamson distinguishes between vertical, horizontal and internal probity, giving examples of each. In so doing, he seems to borrow ideas from economic sociology²⁷⁵ and management literature, many of which appear to contradict or at least duplicate what he would normally say under the headings of "behavioural assumptions" and "the attributes of different governance structures". Using the area of foreign affairs as an example, he introduces vertical probity by asking: "Does the president have confidence in the information and assessments that are provided by the foreign affairs agency?"²⁷⁶ Applying traditional institutional economics, the answer would seem to be that the president would have no reason to

²⁷⁰ For example in section 2.4.3, below.

²⁷¹ Cf. Williamson (1991), p. 281, where he explicitly declines to claim joint sufficiency: "Without purporting to be exhaustive, [the critical dimensions with respect to which transactions differ] include the frequency with which transactions recur, the uncertainty to which transactions are subject, and the type and degree of asset specificity involved in supplying the good or service in question."

²⁷² Williamson (1999), p. 322.

²⁷³ Williamson (1999), p. 322.

²⁷⁴ Williamson (1999), p. 322.

²⁷⁵ Cf. Smelser & Swedberg (2005).

²⁷⁶ Williamson (1999), p. 322-323.

have confidence in the information provided by a privatised foreign affairs agency unless the veracity of the information provided has a sufficient likelihood of becoming known to all *ex post*. If that is the case, the high frequency of the interaction between the president and the foreign affairs agency allows for the use of reputation mechanisms to assure honest performance. If not, the agency will provide inaccurate information whenever doing so is cheaper than honest performance. In other words, the problem is the same as in any other situation where the customer has difficulty verifying the value of the goods or services he has purchased even *ex post*. It is difficult to see why the answer to this question should call for the introduction of any kind of “probity hazard”. Instead, straightforward recourse to agency theory seems more appropriate.²⁷⁷

The other questions he uses to clarify the concept of vertical probity are similarly unhelpful: “Is the agency perceived to be adventurous?”²⁷⁸ “Does the agency comply with directives in a timely and efficacious way?”²⁷⁹ Similarly, the concepts of horizontal probity, which seems to refer to the agency’s ability to deal with its foreign counterparts, and internal probity, add little to the existing framework. In his discussion of Williamson’s approach, Ruiter (2005) places probity in a place where it seems more at home: among the characteristics of governance structures. He argues that the concept is similar to loyalty within organisations and to good faith in contract law.²⁸⁰ However, this still leaves the troubling implication that actors are apparently less opportunistic in some governance structures than in others.

In the end, probably the best way to understand Williamson’s hazard of probity is to analogise from the case he discussed: foreign affairs. Some transactions need to be governed in a manner that meets society’s desire for legitimacy, a desire that can be categorised as “probity”. However, we should be reluctant to consider probity a separate variable for the purposes of economic modelling. Rather, we should leave issues of legitimacy to the political scientists and philosophers who are better equipped to handle them. In any event, this is not a question that is likely to come up in the context of railway regulation, since few people will

²⁷⁷ Cf. Hendrikse (2003), chapter 5-7, discussing the basic principal-agent model, the hidden action problem and the hidden characteristic problem, respectively.

²⁷⁸ Williamson (1999), p. 323.

²⁷⁹ Williamson (1999), p. 323.

²⁸⁰ Ruiter (2005), p. 300. Contrary to Ruiter, and as discussed in section 2.1.3, above, the part of contract law devoted to good faith is probably better viewed as providing an additional set of rules that are incorporated into contracts as a matter of law, rather than as affecting parties’ opportunism.

Certain elements are added to parties’ action sets under the banner of good faith. Most of the time parties know what these elements are because they have been specified in past litigation. Sometimes, however, an element is only made explicit *ex post*, when the parties themselves litigate. All of this acts as additional legal constraint on the freedom to contract, but there is no reason to assume that it will have any effect on parties’ attitudes.

insist that the transactions that take place in the railway industry are in any sense “sovereign transactions”.

2.4.2 Differences between Governance Structures

Compared to his analysis of the characteristics of transactions, Williamson’s discussion of the differences between governance structures is much less well-organised and insightful. What it comes down to is that he describes a one-dimensional range of governance structures, with hierarchy on the one side, the market on the other, and “hybrids” in the middle, based on the instruments they use, their performance attributes and the extent to which they involve contract law.²⁸¹

It goes without saying that different types of governance structures can be characterised by the type of law they use. Markets run on contract law, hierarchies run on employment law, while hybrids use a variety of different areas of law, such as antitrust law, securities law, as well as, most importantly, the non-law of relational contracting, which was already discussed above. Since Williamson works from a one-dimensional perspective, however, he merely observes that different governance structures use contract law to a varying degree.²⁸²

The second way Williamson ranks different governance structures from market to hierarchy is by asking which instruments they rely on in order to produce value: incentives, which are typical of markets, or administrative controls, which are typical of hierarchies? Again, this metric is mostly unproblematic, except that the idea of hybrids involving a moderate amount of both incentives and administrative controls causes difficulty. Many real life hybrids involve neither incentives nor administrative controls, which is exactly what makes them suspect.²⁸³ A poorly designed hybrid is isolated from the incentives generated by the market, while at the same time being independent enough to be able to resist administrative controls.

²⁸¹ Cf. the table on page 281 of Williamson (1991), which sums up the entire framework.

²⁸² Cf. Williamson (1991), p. 280-281. For hierarchies, the argument is that associating hierarchies with employment law is incorrect, since the key feature of this governance structure is that the courts will generally refuse to second-guess instructions given within a hierarchical organisation. Cf. Williamson (1996), p. 153. However, this argument is clearly flawed, since it is employment law that determines when an employer is allowed to dismiss her employee for disobedience (eg. art. 7:678 Dutch civil code), as well as setting rules for penalties that an employer might like to use that are less drastic than dismissal, such as a fine. (eg. art. 7:650 Dutch civil code)

²⁸³ Eg. Van Genugten (2008).

Thirdly, and finally, there is the question of performance attributes. How do different governance structures adapt to changing circumstances? Is this adaptation coordinated, as in a hierarchy, or autonomous, as on a market? As before, the question arises why Williamson presumes that the adaptation of a hybrid is best described as one half of each.

The question of what to do, analytically, with firm-market hybrids was best discussed in a 2002 article by Geoffrey Hodgson, where he considered the “myth of the firm-market hybrid”.²⁸⁴ In this article, he argues for a law-based definition of the firm, along the same lines as the argument in chapter 2.3, above, for at the very least distinguishing between bargaining transactions and managerial transactions: by basing the definition of a firm, or a transaction, on law, the resulting model obtains a degree of clarity and immediate applicability that it would never otherwise have. When applied to the notion of a firm-market hybrid, the conclusion is that it is a myth – legally the parties to the transaction are vertically integrated or they are not. They are part of the same company or they are not.

It follows that the space that is left for hybrids is not so much half-way between firms and markets, but rather somewhere off the continuum altogether. For example, “different degrees of relational exchange, including the possibility of networks.”²⁸⁵ Regardless of its lack of precision, this suggestion is exactly right: hybrids are as different from markets and firms as markets and firms are from each other. Depending on their design, they are sometimes the worst of both alternatives, sometimes the best, and sometimes something different completely.

One of the authors that Hodgson takes umbrage with is Cheung (1983), who described the network of contractors, sub-contractors and sub-sub-contractors in a building process as a quasi-firm.²⁸⁶ Hodgson writes: “However, contrary to Cheung, being ‘vertically integrated by contracts’ is not the same as vertical integration *within* a firm.”²⁸⁷ In my view, Cheung is right to ascribe a degree of vertical integration to such a construction process, although many such projects, while complex, are taxonomically not very problematic. However, in emphasising “the contractual nature of the firm”, he ends up obscuring the nature of the firm without any corresponding benefit, except to enable the author to claim that “[i]t is not quite correct to say that a “firm” supersedes “the market”. Rather, one type of contract supersedes another

²⁸⁴ Hodgson (2002).

²⁸⁵ Hodgson (2002), p. 49.

²⁸⁶ Cheung (1983), p. 11.

²⁸⁷ Hodgson (2002), p. 49. Emphasis in the original.

type.”²⁸⁸ If all governance structures are about contracts, there is no objection to viewing hybrids as nothing more than a halfway-house. But clearly that does not do them justice.

While the Hodgson’s problem with Cheung was that he argued from the correct idea that a firm can be viewed as a nexus of contracts²⁸⁹ to the idea that all hybrids are therefore to some extent a firm, his objection to Ménard was the latter’s use of the market metaphor to describe transactions that are clearly located within a firm.²⁹⁰ The offending statement is that “organizations can be internally structured as quasi-markets.”²⁹¹ To support this statement, Ménard discusses two examples of such quasi-markets: internal labour markets and franchising. The notion of an internal labour market is due to a study by Doeringer & Piore from 1971, and refers to the idea that, for a variety of reasons, firms will prefer to consider only candidates for certain vacancies who already work for the firm. The result is that, when a desirable vacancy opens up, a number of employees compete for the promotion, a situation that can sensibly be described as an internal labour market. However, this nexus-of-contracts perspective of the firm should be kept quite separate from the perspective of a firm as a party to transactions. After all, the pertinent question is how the firm transacts, not how it obtains the labour it needs to do so. That is a question for another day.²⁹²

Ménard’s example of franchising, on the other hand, is relevant enough, bearing in mind his definition of an organisation, which is definitely broad enough to make the collaboration between the franchisor and the franchisee into a single organisation.²⁹³ However, given how broad his definition is, it is understandable that Hodgson should wonder where firms and/or hierarchies are in Ménard’s story. As much as it is defensible to consider McDonald’s as a single organisation, it is important to note that it is an organisation that consists of a great number of different firms, whose connection depends on their ongoing consent.

Having thus taken a slight detour from the work of Oliver Williamson, in order to achieve additional clarity about the notion of a hybrid, a final thing to consider is his 1976 case study of the city of Oakland’s tendering process for cable television within their city. While this article predates the bulk of Williamson’s work on TCE, it summarises and clarifies

²⁸⁸ Cheung (1983), p. 10, emphasis removed.

²⁸⁹ This suggestion is originally by Jensen & Meckling (1976), the pioneering article on agency theory.

²⁹⁰ Hodgson (2002), p. 50, citing Ménard (1995). Ménard (1996) is in the same vein.

²⁹¹ Ménard (1995), p. 176.

²⁹² Or rather, for agency theory such as Jensen & Meckling (1976).

²⁹³ Cf. Ménard (1995), p. 172: “an organization is an institutional arrangement designed to make possible the conscious and deliberate coordination of activities within identifiable boundaries, in which members associate on a regular basis through a set of implicit and explicit agreements, commit themselves to collective actions for the purpose of creating and allocating resources and capabilities by a combination of command and cooperation.”

his work in much the same way that a legal ruling summarises and clarifies a body of statute law. Just like it is almost impossible to discuss the meaning of legal rules in the abstract, without reference to specific cases the rules have been applied to, it is similarly helpful to see the application of an economic model to a practical case in order to clarify its meaning.

2.4.3 Cable television in Oakland, California

On June 19, 1969 the city council of Oakland passed an ordinance clearing the way for a 20 year contract for a Community Antenna Television system to be tendered.²⁹⁴ As we will see again in the empirical chapters below, the problem with such tenders is the winner's curse.²⁹⁵ Or, to be more precise, the question of whose curse it shall turn out to be. Normally, in any kind of auction process where the bidders are largely similar in properties and knowledge, the winner of the auction is likely to have overestimated the true value of the lot, a phenomenon known as the winner's curse.²⁹⁶ However, if the tender process involves a project that requires considerable investments by both parties in transaction specific physical assets and human capital, the winner of the tender might be able to avoid the winner's curse by renegotiating *ex post*. In fact, the optimal strategy might very well be to submit a loss-making bid on purpose.

In the Oakland case, five companies indicated they were interested in bidding, and on April 30 1970 the city published its terms, with a deadline for bids of July 1, 1970. A key variable in the bids was the charge the company proposed imposing on its customers for the basic package of channels. The three main contenders, Focus Cable, Cablecom and TelePrompTer, bid \$ 1,70 per month, \$ 3,48 per month and \$ 5,95 per month, respectively. In other words, the winning company underbid the second lowest bid by a factor of two. So unless Focus would be able to work significantly more efficiently than the others, which seemed unlikely given the evidence, it would have to rescue its profitability either at the expense of its customers or at the expense of the city.

The plot thickened when, on July 16, TelePrompTer and Focus Cable set up a joint venture that would carry out Focus' obligations under the contract. Williamson writes: "Why TelePrompTer was prepared to do this at a monthly charge less than 30 percent of its own bid

²⁹⁴ Facts taken from Williamson (1976), p. 92-101.

²⁹⁵ Cf. Meyerson (1991), p. 132-136 for a game theory treatment of the winner's curse, and Bazerman & Samuelson (1983) for an application to hostile takeovers.

²⁹⁶ Cf. Athias & Nuñez (2008) for an interesting study of the concept in the context of tenders for toll roads.

was not disclosed.”²⁹⁷ Surely, part of the answer was the potential profit on system B, the extended package of channels, but then again, it is difficult to see how that alone would make up for the missing 70%. Despite the cause for concern, the city asked no questions, and in fact approved of this “ideal marriage of local investors, CATV expertise, and over-all financial strength”.²⁹⁸ In September, the companies started raising money for the project that they were officially awarded in November.

Construction was supposed to be completed by the end of 1973, but Focus did not even come close to making that deadline. In early 1974, it requested a number of modifications, including an extension of the deadline with two years, a reduction of the percentage of households that they had to cover, the waiver of the liquidated damages that were specified for the case of delay, and a reduction of the number of channels that had to be offered in the extended package. Consider the city’s options at this point, as listed by Williamson²⁹⁹:

1. Try to enforce the original contract. This would be extremely difficult and costly, and would likely result in expensive litigation.
2. Try to find another operator to take over the contract. This was rejected because it was judged impossible to find another operator who would offer better terms than what Focus was now asking for.
3. Transfer the project to public ownership. This alternative was rejected for “philosophical and financial reasons”, i.e. because public ownership was judged to be too expensive, and generally not desirable.³⁰⁰
4. Renegotiate.

So they renegotiated, reaching a result that gave Focus some, but not all, of the modifications it had asked for, in a deal that was finalised on May 30, 1974. By November 1974, some 11.000 subscribers were connected.³⁰¹

So what does this all mean? Well, first of all it is important to realise that the purpose of this study was not primarily to develop a new theory, but rather to critique an earlier one. In 1970-

²⁹⁷ Williamson (1976), p. 94.

²⁹⁸ Williamson (1976), p. 95.

²⁹⁹ Adapted from Williamson (1976), p. 96.

³⁰⁰ The notion that sometimes governance alternatives are rejected for “philosophical reasons”, rather than for their lack of efficiency, is one that we will return to later.

³⁰¹ According to the 1970 Census, Oakland had approximately 360.000 inhabitants during the relevant period. Source: <http://en.wikipedia.org/wiki/Oakland>, last visited on October 16, 2008.

1972, the later judge Richard Posner wrote a number of articles about cable television, which were based on an idea developed by Harold Demsetz called “franchise bidding”.³⁰² Following in Coase’s footsteps, Demsetz had tried to construct an institutional setup for utilities regulation that involved as little as possible actual regulation. The idea was for the government to decide in advance which services it wanted the franchisor to offer to consumers at which prices, and then to invite bids, which could be in terms of an amount paid or received, depending on the level of generosity of the government towards its citizens. As Williamson showed in his article, that does not work. The problem is primarily one of information: the government cannot very easily prove in a court of law that any breaches of contract occurred, that they are material, or that they are attributable to the company.

Applying Williamson’s model to this case study, we can start by noting that this type of moral hazard problem can, with some creativity, be considered under the uncertainty prong of the model. Asset specificity and frequency, on the other hand, are a bit of a problem. The assets that Focus invested in are clearly specific to the transaction, but then, they are not the ones being held up. The only specific investment that the city of Oakland made was in the contracting process itself.³⁰³ The transaction costs they incurred in setting up and administering the bidding process were sizeable, but one wonders whether they were sizeable enough for the purposes of the model.

The frequency, strictly speaking, is only one. Only once did the city grant Focus a set of prerogatives, such as the right to dig up the streets for construction, and the right to sell cable television subscriptions in Oakland, and only once did Focus, in return, agree to contract with its customers on the terms specified by the city, which is what constituted its consideration. The only transfer of goods or services that occurred more than once was the annual payment Focus was required to make to the city.

Looking at the underlying logic of Williamson’s model, one would say the point is to establish the stake parties have in setting up a possibly expensive governance system.³⁰⁴ How much value is generated by the transaction, and how does this amount compare to the costs of governance under various alternative governance structures? Different forms of vertical integration, or the absence thereof, carry with them their own exposures, which consist of a prob-

³⁰² Demsetz (1968). The Posner articles include Posner (1972).

³⁰³ In other words, one could say that the only specific asset that the city owned was the contract.

³⁰⁴ Notice how Milgrom & Roberts (1992) call their second transaction characteristic “frequency and duration”. (p. 30-31)

ability of opportunistic behaviour by the counterparty, multiplied by the damage caused by such behaviour.³⁰⁵ Ultimately, we might conclude that Williamson, too, imagines governance choices to be made by comparing these exposures under different arrangements with their governance costs.

Finally, it is useful to consider what would happen to the Williamson model if we took out the fundamental transformation. What if both parties are monopolists who need each other in order to do business? What if Fisher Bodies has no one else to do business with but Ford and vice versa? In that case, a transaction still occurs, which has to be governed by some kind of contract. Also, the parties still have the alternative of vertically integrating. However, all of a sudden each party's investment in transaction-specific assets is completely unimportant. After all, such investments do not lock the parties together any more than they already are. If the only outside option is to go out of business altogether, the relevant question is not whether each party has invested in specific assets, but rather to what extent they have incurred sunk costs generally.

Similarly, the question of frequency becomes difficult to handle; after all, for all intents and purposes the parties are locked into a perpetual series of transactions, where it is even more difficult than usual to see where one transaction ends, and another one begins. Moreover, the question of frequency, like asset specificity, is much less important in this kind of bilateral monopoly setting. It is no longer a question of dividing the fixed costs of setting up a governance structure over a given number of transactions, but rather one of comparing the relative efficiency of alternative institutions.

In terms of exposures, there are two sides to this coin. On the one hand, the risk of opportunistic behaviour is much smaller, since it is no longer possible to do a "hit and run". A threat of hold-up, for example, will normally not be credible in a bilateral monopoly setting. How can Ford hold up Fisher Bodies if it has no alternative source of parts? However, on the other hand, it is also much more difficult to punish opportunistic behaviour. Normally, a sufficiently serious case of opportunistic behaviour simply means that the parties will no longer do business together in the future. In a bilateral monopoly setting, the parties do not have the

³⁰⁵ One could imagine a Coasian model that simply aims to describe and map out the different sources of exposure, such as exposure due to asset specificity, exposure due to moral hazard, and possibly many others. While such an effort is outside the scope of this thesis, the case studies of the different reforms in the railway industry, which can be found in chapter 4, will frequently be cast in terms of exposures.

option of punishing each other in this way. Instead, all they can do is renegotiate, a topic that is not covered by Williamson, but that has a prominent place in the models of Grossman, Hart, and Moore which are the topic of the next section.

2.5 Incomplete Contract Theory

Having discussed Oliver Williamson's answer to Ronald Coase's challenge, we must now turn our attention to the other main approach, the framework developed in a series of articles by Grossman, Hart and Moore (hereafter: GHM).³⁰⁶ While Williamson's model is eminently useful for scholars looking to do empirical research³⁰⁷, the GHM model is the preferred starting point for those interested in developing sophisticated formal models.³⁰⁸ After all, the original articles themselves also contain some fairly advanced formal modelling. Since the present thesis does not aim to engage in such sophistry, we will discuss these articles mostly in non-mathematical form.

As noted in section 2.3, above, in this line of models, all transactions are essentially bargaining transactions. In every one of the articles that we will discuss, the consequence of integration is that the veto points are distributed differently, changing the nature of the *ex post* renegotiation. However, there is always a bargain.

2.5.1 Grossman & Hart (1986): *The Costs and Benefits of Ownership*

In this first article, Grossman and Hart construct a two-period model of two actors transacting in an environment with a great degree of incompleteness. The benefit that they each derive from the transaction, which is noncontractable, depends on their respective relationship-specific investments, which are noncontractable, and the other actions taken by themselves and their counterparty, which are also noncontractable. While these variables are not verifiable, they are observable, meaning that both parties have identical information about them.³⁰⁹ After they contract, at date 0, they each choose their specific investment. At date 1, the owner of each firm chooses the action set for that firm, meaning that in case of vertical integration, they still each chose their own specific investment, but the owner of the integrated firm chooses the action set for both. In case of non-integration, the action sets can be renegotiated costlessly *ex post*, i.e. at date 1, ensuring that they are *ex post* Pareto efficient.

³⁰⁶ The most important ones are Grossman & Hart (1986), Hart & Moore (1988, 1990, 1999) and Hart, Shleifer & Vishny (1997), each of which will be discussed briefly in this section. A recent extension is Volokh (2007). For an introduction, cf. Hart (1995).

³⁰⁷ Such as Joskow (1985).

³⁰⁸ Such as Lülkesmann (2007).

³⁰⁹ Cf. chapter 2.1, above.

The result of this setup is a thrilling series of negotiations and other noncooperative Nash games. The parties first bargain over their respective action sets, then they choose their levels of investment noncooperatively, followed by renegotiation and finally exchange. The question is whether they will invest the optimal amount, and whether they will end up choosing the *ex ante* optimal action sets. Under nonintegration, the problem is that the parties will underinvest, since they will have to bargain over the value generated *ex post*.³¹⁰ To be precise, the manager puts 50% weight on the influence of additional investment on the *ex post* surplus, and 50% weight on its influence on the *ex post* non-cooperative outcome, which will never actually happen since it is inefficient, but which serves as the threat point in the *ex post* Nash bargaining game. Under vertical integration, the problem is that the action sets will be chosen to benefit only the owner, meaning that there is still no reason why they should be Pareto efficient. *Ex post* renegotiation will still be necessary, so the only consequence of integration is to change each side's outside option, i.e. the non-cooperative outcome that serves as the threat point in their renegotiation. Each side still underinvests.

The interesting question now becomes: Which integration alternative leads to the least amount of underinvestment? The answer is given in Grossman & Hart's Proposition 1, which is worth quoting in full, substituting only our symbols for theirs:

- (A) Suppose that $X_{1,2}$, [the benefit of the transaction to each party,] depends primarily on $A_{1,2}$ [, that party's own action set,] in the sense that $X_1(A_1, A_2) = \alpha_1(A_1) + \varepsilon_1 \beta_1(A_2)$, $X_2(A_1, A_2) = \alpha_2(A_2) + \varepsilon_2 \beta_2(A_1)$, where $\varepsilon_1, \varepsilon_2 > 0$ are small. Then non-integration yields approximately the first-best, while firm 1 and firm 2 control generally do not.
- (B) Suppose that X_2 hardly depends on A_1 and A_2 in the sense that $X_2(A_1, A_2) = \alpha_2 + \varepsilon_2 \delta_2(A_1, A_2)$, where $\varepsilon_2 > 0$ is small. Then firm 1 control yields approximately the first-best, while non-integration and firm 2 control generally do not.
- (C) Suppose that X_1 hardly depends on A_1 and A_2 in the sense that $X_1(A_1, A_2) = \alpha_1 + \varepsilon_1 \delta_1(A_1, A_2)$, where $\varepsilon_1 > 0$ is small. Then firm 2 control yields approximately the first-best, while non-integration and firm 1 control generally do not.³¹¹

In simple terms, the key question is how sensitive each side's benefit is to the action set of the other side, which, remember, cannot be contracted *ex ante*. Where such sensitivity only exists for one of both firms, the firm that is most sensitive should control the vertically integrated

³¹⁰ Bargaining is assumed to result in a Nash-type 50/50 split. Cf. Nash (1950).

³¹¹ Grossman & Hart (1986), p. 705.

firm. Alternatively, the question can be asked in terms of whose underinvestment is a bigger problem, which, at least in theory, amounts to the same thing.

To illustrate, the authors discuss vertical integration in the insurance markets, where the question arises whether companies should have their own in-house sales force or whether they should work with independent agents, who “own” their own client list. While it goes too far to discuss their entire analysis here, one can get a sense of it by considering their conclusion: “in products in which the renewal is not guaranteed and is sensitive to the agent’s actions, the agent will be more likely to own the list, whereas in products in which the renewal is more certain and is less sensitive to the agent’s actions, the company will be more likely to own the list.”³¹²

2.5.2 Hart & Moore (1988): *Incomplete Contracts and Renegotiation*

The starting point for this article is even simpler. No more externalities, just the assumption that for the seller the cost price of the product depends on the seller’s own specific and unverifiable investment and on the state of the world, while the buyer’s valuation of the product depends on the buyer’s specific investment and, again, on the realisation of the state of the world.³¹³ The first best is any outcome where a sale occurs whenever the buyer’s valuation of the product is equal to or higher than its cost price, but if the sales price is cast in stone, that might not happen. After all, there is a chance that the buyer’s valuation will end up higher than the cost price, but lower than the previously agreed sales price. In such a case, renegotiation is necessary in order to “save the deal”, but, as before, such renegotiation distorts the parties’ incentives to invest.

If such renegotiation could be limited to those cases where it really is necessary to avoid a breakdown, it would not be much of a problem. But since the realisation of the cost price and the buyer’s valuation cannot be verified by a court, we would expect renegotiation to occur in every state of the world. So, as in the previous model, whenever a party invests, he can expect to share part of the value generated by that investment with the other side, thus

³¹² Grossman & Hart (1986), p. 714.

³¹³ Their equation 2.1: $v = v(\omega; \beta), c = c(\omega; \sigma)$, where v is the buyer’s valuation, c is the cost price, ω is the state of the world, and β and σ are the buyer’s and the seller’s investments, respectively. It follows that the goal is to maximise $W(\beta, \sigma) = E_{v,c}[\max\{v - c, 0\} | \beta, \sigma] - h_b(\beta) - h_s(\sigma)$, where the last two terms represent each side’s cost of investment. (Their equation 2.3.) Hart & Moore (1988), p. 757-759.

reducing the incentive to invest in the first place. To solve this problem at least partly, Hart & Moore develop a mechanism revolving around two prices, instead of just one. These two prices are made dependent on the only thing that a court can observe: whether a trade has occurred or not. In what follows, the authors make both the price in case of trade and the price in case of no trade dependent on messages that are being sent by the parties. Such revelation mechanisms have since been dealt with extensively elsewhere, and are, in any event, outside the scope of this thesis.³¹⁴

All that matters is that they conclude that the first-best is possible when: 1) there is a price for which there is no chance that it will not lie between the buyer's valuation and the cost price, or 2) if the buyer's valuation does not depend on her investment, or 3) if the cost price does not depend on the seller's investment, or 4) if both the buyer's valuation of the good and its cost price are the same in all states of the world.³¹⁵

2.5.3 Hart & Moore (1990): Property Rights and the Nature of the Firm

This third article is best summed up by an example that the authors give to illustrate:

“Suppose that two workers 1 and 2, in conjunction with some specific asset, can provide a service to consumer 3 at date 1. To aid the exposition we shall imagine that the asset is a luxury yacht, worker 1 is a chef, worker 2 is the skipper, the service is gourmet seafare, and consumer 3 is a very important customer, for example, a tycoon who entertains frequently. For this service to be useful, the chef must, at cost 100 to himself, take an asset-specific action at date 0, for example, acquire a skill (the skill might be the preparation of a particular cuisine). (...) For simplicity, we suppose that it is not transferable at all (there are no other yachts cruising nearby). Suppose also that (i) the value of the service to the tycoon is 240 and, moreover, no other consumer values the service at all (only the tycoon can afford to fly to these waters); and that (ii) there are many substitutes for the skipper's skills at date 1, in the sense that the

³¹⁴ Eric Maskin, Roger Myerson and the late Leonid Hurwicz received the 2007 Nobel Prize in Economics for their work on mechanism design. The purpose of this literature is to study the possibility of designing a mechanism to give parties incentives to truthfully reveal information that is not observable or at least not verifiable, so that a contract can nevertheless be written that refers to such information. The result usually requires highly sophisticated actors, which is why such mechanisms can only be used when the stakes are sufficiently high, for example with auctions for mobile telephone concessions. Cf. for example Maskin & Moore (1999) and Maskin & Tirole (1999), which are also part of the incomplete contract theory literature that is the topic of the current section, Myerson (1981, 1982, 1983), where the emphasis is more on advanced game theory, and Baron & Myerson (1982), a pioneering article on the theory of regulation. Binmore & Klemperer (2002) discuss the role of economists in the design of the UK UMTS auction.

³¹⁵ Hart & Moore (1988), p. 772, proposition 3. The argument for the first and the fourth of these possibilities should be obvious. Cases 2 and 3 allow for the first best by giving the party whose investment does matter a claim to the entire *ex post* surplus.

*chef and tycoon can easily go to the date 1 spot market and replace the skipper if necessary. Suppose in addition that transaction costs prevent the writing of any long-term contract at date 0 and that the discount rate is zero.*³¹⁶

So who should own the yacht?

It turns out that the chef will only invest if he himself owns the yacht, or if the tycoon does. The reason for this is that if the skipper owns the yacht, the consent of three people is necessary to carry out the transaction, meaning that the gains from trade will have to be split equally between them.³¹⁷ Since 240 divided by 3 is 80, which is less than 100, the chef will not invest if the skipper owns the yacht. Under any other ownership configuration, the gains from trade only have to be divided between two people, creating a net gain to the chef of $240/2 - 100 = 20$. Intriguingly, the only aspect of ownership that the authors need in order to reach this conclusion is the ability to exclude others³¹⁸; if the skipper owns the yacht, he has the power to exclude the chef or the tycoon from its use, and that alone is enough to entitle him to an equal share. No other assumptions about the consequences of integration are necessary.

For this setup, Hart and Moore derive twelve propositions about who should own what in different circumstances. Throughout, the asset is a physical asset, and the specific investment is in human capital. For example, if only one agent has an investment, they should own all assets³¹⁹, if more than one agent has an investment, underinvestment is unavoidable.³²⁰ Also, there should be no shared control³²¹ and complementary assets should be owned by the same person. Most of these propositions are intuitive, so the proofs are not given here.³²²

³¹⁶ Hart & Moore (1990), p. 1122-1123.

³¹⁷ This follows from the Nash bargaining solution which is assumed. Nash (1950).

³¹⁸ Hart & Moore (1990), p. 1121.

³¹⁹ Proposition 2, p. 1131.

³²⁰ Proposition 1, p. 1130.

³²¹ Proposition 3 and 4, p. 1132.

³²² Some of the other propositions concern the straightforward situation where there is a special connection between an asset and one of the agents, for example because the asset only affects the marginal benefit of one of the agents or because the agent is indispensable to the asset. In both cases, the agent should own the asset. (Propositions 5 and 6, p. 1133.) Another interesting proposition is proposition 9, which excludes the possibility of control by an outside party, i.e. an agent whose participation does not affect the benefit to the others. This proposition is interesting because it requires that stochastic control should be necessary. If stochastic ownership is impossible, giving control to an outside party can help balance the incentives to the other agents.

2.5.4 Hart & Moore (1999): Foundations of Incomplete Contracts

In this final article of the series, Hart and Moore go into the details of incompleteness. Instead of postulating that a long-term complete contract is impossible, they construct a model that shows why that should be the case. The idea is that parties are uncertain about the exact product that they will trade in the future. They do not know, at the time of contracting, which of the N possible widgets will turn out to be the special widget. What they do know is that the production costs of this special widget will take one of two values, with probabilities depending on the seller's investment, and that the production costs of the generic widgets are in a range between those two extremes. Also, the value of the special widget to the buyer is always higher than the production costs, so ideally a trade should always occur. For this setup, they prove that the first best can be achieved, even without a complete contract, as long as renegotiation can be excluded.³²³ Without renegotiation, parties can simply agree a price equal to the benefit of the product to the buyer, which is fixed, making the seller's incentive to invest depend on the full size of the gains from trade. If renegotiation cannot be avoided, on the other hand, no contract, complete or incomplete, can do better than the "no trade" contract, since parties' *ex post* renegotiation will always dissipate the seller's incentive to invest.³²⁴ Taking all things together, the only way that contractual completeness, *in casu* the ability of parties to describe fully which widget they will want to trade, matters, is if specific performance is possible as a remedy.

Remember, specific performance, the contract law remedy of having the court compel the party in default to do as contracted, is a less than obvious remedy to authors, such as Hart and Moore, who are writing from a common law perspective.³²⁵ Under civil law, any default in the execution of a contract entitles the other party to ask the court to order specific performance.³²⁶ In common law countries, on the other hand, traditional common law only entitled the creditor to money damages. As a result, the doctrine of specific performance was only developed in the courts of equity, which means that it comes with the same strings attached as all equitable doctrines. For example, equitable remedies are only available if they are appro-

³²³ Hart & Moore (1999), proposition 1, p. 119.

³²⁴ Hart & Moore (1999), proposition 2, p. 121. Note that this result depends critically on the assumption that the cost of the generic widgets lies in the interval between the high cost and the low cost possibility for the special widget. Without this assumption, at least some of the seller's incentive to invest can be maintained.

³²⁵ This is one of the moments where the complaint of Brousseau (2001) about a common law bias in institutional economics is particularly valid.

³²⁶ Cf. art. 3:296 Dutch Civil Code, art. 1184 Code Civil and § 241 BGB, all three of which express the, to the authors almost self-evident, principle that a creditor can go to court to claim what he is entitled to under the contract.

priate given all the circumstances of the case, and only as long as the party requesting the remedy has acted in good faith. These conditions are usually maintained in statutory versions of the remedy, for example § 2-716 of the American Uniform Commercial Code.³²⁷ Throughout the common law, specific performance is an alternative to money damages, when the latter is somehow not appropriate.³²⁸

Regardless of the legal niceties, though, it should be clear that specific performance is only possible to the extent that the contract actually specifies what each party is expected to do. In terms of Hart & Moore (1999), if the parties cannot, *ex ante*, specify how to tell which widget is the special widget, specific performance is impossible. Instead, they would have to rely on a revelation mechanism, which necessarily distorts the seller's incentive to invest.³²⁹ This means that the general conclusion of the article, which is that the possibility or impossibility of avoiding *ex post* renegotiation is more important than the completeness of the contract, has one important exception for the case where specific performance is possible.

2.5.5 Hart, Shleifer & Vishny (1997)

Hart, Shleifer & Vishny (1997), finally, is interesting because it is the first article that applies the GHM model to the privatisation problem that is the topic of this thesis. As it turns out, the setup can stay largely the same as a result of this change in application. While theoretically most privatisation questions concern three parties, i.e. the government, the would-be contractor and the citizen/customer, this last party can be abstracted away without loss of generality.³³⁰ This leaves a simple problem of a transaction consisting of a service being provided by a private party to the government, and the question of whether the efficiency of the transaction can be improved by having the buyer and the seller vertically integrate.

³²⁷ Consider the language of this article: "Specific performance *may* be decreed if (...). The decree for specific performance may include such terms and conditions (...) as the court may deem just." (Emphasis added.) Similarly, s. 50 of the UK Supreme Court Act of 1981, c. 54, gives the High Court blanket authority to award damages instead of specific performance.

³²⁸ For more information on the doctrine of specific performance in various jurisdictions, cf. Catalano (1997), Clark & Clarke (2000), p. 1140-1165, Cuncannon (2004), Schwartz (1979), and Szladits (1955).

³²⁹ A sample of literature in the field of revelation mechanisms was cited above, in a footnote to the discussion of Hart & Moore (1988).

³³⁰ Note that this same approach was used in the discussion in par. 2.4.3 of the Oakland CATV case as well. Cf. also the discussion of incompleteness due to missing parties in par. 2.1.3, above.

It is worth going into the nuts and bolts of the Hart, Shleifer & Vishny version of the GHM model, because in many ways it is the most basic specification of the GHM approach in general. What they need to make it work is the following: Let the benefit of the good to the buyer be expressed as $B = B_0 - b(e) + \beta(i)$ and the cost price to the seller as $C = C_0 - c(e)$, where e represents the seller's uncontractible effort to innovate in order to reduce costs, i represents the seller's equally uncontractible effort to innovate in order to improve quality, and $b(e)$, $\beta(i)$ and $c(e)$ are functions that express the influence of the seller's efforts on the buyer's valuation of the product and the seller's cost price. These two forms of effort cannot be written into a contract because only the seller can observe them. Also, they cannot be described indirectly, because of the uncertainty relating to each type's influence on B and C .

Under the first-best outcome, the marginal costs and benefits of each kind of effort are the same, where the marginal benefit of effort is defined with respect to its influence on the total gains from trade and the marginal cost of effort is, by definition, one. Mathematically: $-b'(e^*) + c'(e^*) = 1$ and $\beta'(i^*) = 1$.

Under privatisation, the seller clearly benefits from reducing his cost price. In fact, to the extent that such a reduction hurts quality, it is likely that the seller will overinvest in cost savings. Improvement in quality, though only benefit the seller if he can renegotiate the cost price. Since the assumption is, as usual, a Nash bargaining solution³³¹, the seller receives only half the benefit created by his investment in quality improvement. So the investment in effort by the seller is characterised by the following equations: $c'(e_M) = 1$ and $\frac{1}{2}\beta'(i_M) = 1$.

Under vertical integration, on the other hand, when the state owns the critical asset, all the innovation requires the consent of both parties, meaning that, under Nash bargaining, the gains are divided equally, *ex post*. As a refinement, Hart, Shleifer & Vishny assume that the seller will receive at best one half of the gains from trade, and less if he is not completely irreplaceable. The degree to which the seller is essential is measured by the parameter λ . Mathematically, this means that under nationalisation the seller's effort is defined by:

$\frac{1}{2}\lambda\{-b'(e_G) + c'(e_G)\} = 1$ and $\frac{1}{2}\lambda\beta'(i_G) = 1$. Note that when the seller is completely expendable, he has no incentive to exert any effort whatsoever, since he would lack the bargaining power to claim a share of the benefits.

³³¹ Nash (1950).

Taking it all together, and given the usual assumptions about the functions involved, we see that the stylised facts we would expect to find do, in fact, follow from the mathematical expressions. Since the term $-b'(e)$ does not appear in the expression for the seller's effort under privatisation, it follows that his effort for cost cutting is higher than under the first best, while the multiplication by $\frac{1}{2}$ in the other equation leads to less than optimal investment in quality.³³² Comparing the expressions for public ownership with the first best, we see that the entire left-hand side of the equation for cost cutting is multiplied by $\frac{1}{2} \lambda$, implying lower incentives for this type of effort, while the seller's incentives for quality improvement are, at best, as good as under privatisation.³³³

Comparing public and private ownership to each other, instead of to the fictional first-best, we see that privatisation is preferred if $b(e)$ is sufficiently small, meaning that cost cutting hardly hurts quality, or generally if cost cutting effort is largely ineffective, when quality is all that matters.³³⁴ Public ownership is clearly better when the net effect of effort in cost cutting on social welfare, $c(e) - b(e)$ is sufficiently small or when effort in quality improvement is sufficiently unimportant.³³⁵ Altogether, the model implies that costs are always lower under private management, while quality could be either lower or higher.³³⁶

2.5.6 Evaluation

Comparing the GHM approach to the Williamson approach, we notice first and foremost the different starting points of both models. Where Williamson started by carefully studying case studies of real life contracting problems, GHM focused on constructing models that were simple enough to be mathematically tractable, yet rich enough to yield falsifiable hypotheses. From that point of view, it makes sense that those economists who are more mathematically inclined should favour the GHM models, while those who are interested in doing empirical studies should favour Williamson. In fact, one could very well wonder whether it is possible at all to empirically test the GHM framework without significant additional work.

³³² This is Hart, Shleifer & Vishny (1997) proposition 1, p. 1139.

³³³ Hart, Shleifer & Vishny (1997), proposition 2, p. 1139.

³³⁴ Hart, Shleifer & Vishny (1997), proposition 3, p. 1141.

³³⁵ Hart, Shleifer & Vishny (1997), proposition 4, p. 1142.

³³⁶ Hart, Shleifer & Vishny (1997), proposition 5, p. 1143.

From the point of view of pure theory, on the other hand, the most interesting difference is the role of administrative controls.³³⁷ What kinds of control does a manager have over his employees that contractual parties do not have over each other? In the work of Oliver Williamson, administrative controls have always been a bit of a weakness. Judging from his work, parties' tendencies towards opportunistic behaviour seem to magically disappear after integration. Would one not expect that the agency problems after integration would be of a similar nature to the contracting problems before? It was this very criticism, articulated for example in Grossman & Hart (1986), that first prompted GHM to explore a different approach. As explained above, they constructed models where, even under integration, each party still had a lot of independence.

It should be obvious that this is a question not without importance. If Williamson's model really does critically depend on a principal's ability to overcome any agency problems and control his organisation, that would take away much of its presumed usefulness. After all, even a casual examination of many of the make-or-buy situations that these models are meant to examine shows that managers retain significant independence in either case. In the private sector, the kinds of activities that might be outsourced tend to involve specific human capital. This goes for the accounting department, or the in-house counsel, but also for the maintenance of office computers and photocopiers. It is inherent in these activities that senior management cannot control them except in very general terms.³³⁸ Similarly, in the public sector the kinds of activities that are privatised, liberalised, or at least organised into an autonomous quango³³⁹ tend to display a high degree of independence no matter how they are organised. Even when public utilities are completely owned and controlled by the government, they pretty much tend to do as they see fit.³⁴⁰

Williamson's only explicit reply to Grossman & Hart (1986) is found in Williamson (1988). Unfortunately, he mischaracterises their position somewhat, saying: "The Grossman

³³⁷ For a good study of control from a management science point of view, cf. Eisenhardt (1985). A recent attempt to set up a model explaining the nature of administrative controls is Zanarone (2008). His approach uses credible commitment literature to explain what changes as a result of integration.

³³⁸ Ironically, this is also the tenor of Williamson's dissertation, which is about managerial discretion and the theory of the firm. Cf. Williamson (1967).

³³⁹ Cf. Greve, Flinders & Van Thiel (1999) for an attempt at definition.

³⁴⁰ De Witte & Saal (2008), for example, discuss the recent history of the Dutch water sector. They find that in the late 1990s all Dutch water companies raised their prices in anticipation of privatisation and/or liberalisation. At that point, no actual decision had yet been made. "Anticipating their own privatization, the Dutch drinking water utilities strove to increase their financial strength so as to be better prepared for the expected changes (...) (several annual accounts cite this). Indeed, if the government would decide to privatize the sector, the utilities realised that they would need to be 'attractive' to draw investment." (p. 5) The authors find an increase in total economic profit, in 1995 prices, from € 4,3 million in 1992 to € 177,4 million in 1996.

and Hart treatment of vertical integration assumes that the manager of each stage is also the owner. This is a simplification, one consequence of which is that incentive intensity is assumed to be unaffected by vertical integration. The application of the argument to the case where the manager of each stage has a negligible ownership position is not developed.”³⁴¹ As we have seen above, that is not actually correct. Instead, Grossman and Hart argue that under vertical integration, the manager whose assets are now owned by the other firm still retains an essential position, allowing him to extract part of the gains from trade, while making it difficult for the owners of the vertically integrated firm to get their subsidiary and its manager to invest at the optimal level. They write: “We define integration in terms of the ownership of assets and develop a model to explain when one firm will desire to acquire the assets of another firm. (...) To the extent that there are benefits of control, there will always be potential costs associated with removing control (i.e. ownership) from those who manage productive assets.”³⁴² How the “incentive intensity” of the manager is affected by vertical integration is exactly the question that Grossman and Hart posed. Note that this question is tackled explicitly only in Hart, Shleifer & Vishny (1997), published several years later, when they introduce the parameter λ to capture the strength of the manager’s position after integration.

A related question is the role of opportunism in both approaches. Because Williamson is focused on the possibility of various hold-ups, he needs the assumption that actors act opportunistically.³⁴³ In this context, opportunism is defined as “self-interest-seeking with guile”³⁴⁴, meaning that actors are assumed to do anything they can get away with in order to promote their self-interest.³⁴⁵ GHM, on the other hand, approach the problem in a similar vein to agency theory; it is not necessary to impute the actors with any particular selfishness. The model simply shows how their incentives are misaligned. However, this difference should not be exaggerated. Much of what Williamson describes as opportunistic behaviour, GHM assign the more neutral-sounding label of “*ex post* renegotiation”. For example, analysing Joskow’s study of mine-mouth electricity plants³⁴⁶ using the GHM framework, we would conclude that the timeline looks as follows: 1. The owner of the mine and the would-be owner of the plant write a contract, 2. The plant owner chooses a location and builds a plant, 3. They renegotiate, 4. The mine starts delivering coal. Since the bargaining at moment 3 would be over all the

³⁴¹ Williamson (1988), p. 583.

³⁴² Grossman & Hart (1986), p. 693.

³⁴³ Cf. Williamson (1985), p. 64-67.

³⁴⁴ Williamson (1985), p. 65.

³⁴⁵ To be precise, it is assumed that some actors will, and that it is impossible to distinguish them from those actors that would not. Williamson (1985), p. 64.

³⁴⁶ Joskow (1985). Cf. chapter 2.4 for details.

gains from trade given the investment, the plant owner's incentive to invest at point 2 is distorted – some of the value generated by the investment would be transferred to the mine owner *ex post*. In other words, applying the GHM framework to this situation yields superficially similar conclusions, even though they do not talk about “opportunism”.

Finally, an interesting fact about the GHM approach is that it actually matters which of the two parties buys the other. Should the mine buy the electricity plant or should the electricity plant buy the mine? In Williamson's model, that is irrelevant, but for GHM it is not. Williamson, on many occasions, discusses what he calls “M-Form corporations”, i.e. multidimensional organisational designs.³⁴⁷ While such an organisational form is undeniably common, within the GHM framework it is inefficient, since it takes away the incentives of both managers; the CEO and shareholder of the company owns all the assets, but the manager of each division has to use those assets, combining them with the manager's transaction specific human capital. Since each manager has to renegotiate *ex post* with the CEO, both their incentives to invest in human capital are lower than optimal.³⁴⁸ For one of them, the incentives can be improved by having him own the assets, but the question is which of the two that should be.

We should note from the outset that it is not the location of the factory, or any related investment in physical capital, that is the problem. Under all versions of vertical integration, the electricity plant will be built next to the mine if that is the efficient choice, since obviously that decision is within the sole discretion of whichever of the two managers ends up owning the integrated company. Instead, the question to focus on concerns the incentives of the manager who does not own the assets he works with to invest in human capital. Then there is each manager's incentive to exert effort into innovation, along the lines of Hart, Shleifer & Vishny (1997). Obviously such questions cannot be settled without more empirical research, but we could argue for example that effort into cost saving is more important in this transaction than effort into quality improvement. Assuming that fact, it would be more efficient for the mine to own the factory than the other way around.³⁴⁹ Similarly, if empirical examination found that

³⁴⁷ Eg. Williamson (1985), p. 273-297, referencing Chandler (1962).

³⁴⁸ The inefficiency of having someone without an essential investment in human capital own an asset is shown in the yacht example of Hart & Moore (1990), and formalised in that article's Proposition 9. Cf. par. 2.5.3, above. However, the principle can be demonstrated with most of the other models as well.

³⁴⁹ Only ownership by the buyer was analysed in Hart, Shleifer & Vishny (1997), since having the seller own the government might prove a bit problematic. However, developing their model to accommodate both possibilities, and dropping the quality improvement term, we would have a first-best of $-b'(e) + c'(e) = 1$, an expression

the mine manager's investment in human capital is both more specific and more important than the plant manager's, then, again, the former should own the integrated company.

All of this begs an important question. Given that, in reality, we would surely expect the vertically integrated company to be organised as a multidivisional corporation of the type discussed by Williamson, with a CEO and a board, and a separate manager for each division, where is the error? In the GHM model, or in reality?³⁵⁰

for buyer ownership of $\frac{1}{2}\lambda\{-b'(e) + c'(e)\} = 1$, and an expression for seller ownership of $(1 - \frac{1}{2}\lambda)\{-b'(e_G) + c'(e_G)\} = 1$. After all, it is the seller who chooses e . Under buyer ownership, he only gets to keep $\frac{1}{2}\lambda$ of the value generated by his investment, but if he, himself owns the assets, he only has to give the buyer $\frac{1}{2}\lambda$, where, in each case, λ represents the bargaining power of the manager who does not own the assets. Since λ can be no higher than 1, and assuming it is lower in both cases, seller ownership is more efficient than buyer ownership.

³⁵⁰ Note that this is a serious question, no mere rhetoric. If one is persuaded by the GHM models, one may very well conclude that the normal way that such companies are organised is inefficient. Attempts to extend the GHM model to make it apply better to companies that are not owner-managed include Holmstrom (1999) and Hart & Holmstrom (2008).

2.6 Alignment Mechanisms

2.6.1 Introduction

In the previous two sections, we discussed what the optimal alignment between transaction and governance structure would look like. When there are significant exposures, or at least significant misalignments when it comes to incentives to invest in transaction specific assets, transaction costs can be reduced by vertically integrating. If the goal were simply to advance a normative theory, that would be enough. However, since the purpose of this thesis is equally to develop a descriptive model, we need one more element: We need to explain why the optimal solution is the one that we would expect to see in reality. Is there any reason why the optimum is more likely than any other conceivable setup?

The neoclassical theory of the market, for example, not only proves that a perfect market leads to a Pareto efficient outcome, but also that this outcome is the only possible outcome given the starting assumptions.³⁵¹ All that is needed in order to achieve the Pareto efficient equilibrium is for all the – perfectly rational – actors to use their private information about their preferences and costs and the publicly available information about the market price to trade whenever it is beneficial to do so. When there are no more possible trades that are mutually beneficial, the optimum is achieved. It follows that the model works both as a normative “policy advice” and as a positive “description of reality”.³⁵²

Of course, one could very well postulate a similar mechanism for neo-institutional economics (NIE): actors will always change the governance structures governing their transactions when it is mutually beneficial to do so. However, that is not an acceptable solution, since NIE emphatically denies perfect rationality.³⁵³ Indeed, as we have seen, if actors were perfectly rational, there would be no contractual incompleteness. It is difficult to imagine how actors could be rational enough to design perfect governance structures, but not rational enough to write complete contracts.

What this means is that neo-institutional models cannot work both as descriptive and proscriptive models quite so easily as most of economic theory. What is needed is an addi-

³⁵¹ Eg. Martin (1993), p. 14-38 or Cowell (1986), *passim*.

³⁵² Whether it succeeds as such is a different story. Cf. Friedman (1953) for a defence.

³⁵³ Eg. Williamson (1985), p. 44-47 or Klein (1999), p. 467.

tional element, an explanation why reality should be expected to drift towards the optimum. What is needed is an explanation of why we would expect the “alignment” between transaction and governance structure to actually occur. What is needed is a discussion of alignment mechanisms.

Oliver Williamson gives this question precious little attention. Of course, his study of CATV in Oakland, which was discussed in section 2.4, above, can be viewed as an analysis of the practicability of what we will call rational design as an alignment mechanism. After all, that article essentially studied whether a rational lawmaker, sitting down to write the best possible governance structure for a given problem, will end up anywhere near the optimum. But nowhere in his work does he systematically analyse why a “discriminating alignment”³⁵⁴ would be more likely than any of a hundred possible misalignments.

To fill this gap somewhat, the present section will discuss four possible alignment mechanisms: evolution, competition within the market, competition with rival products, and rational design. Each will now be described in turn.

2.6.2 Evolution

In 1905, Justice Oliver Wendell Holmes of the United States Supreme Court dissented from that court’s landmark ruling in *Lochner v. New York*³⁵⁵, writing: “The 14th Amendment does not enact Mr. Herbert Spencer’s Social Statics.”³⁵⁶ What he was referring to was an 1851 book by the English philosopher Herbert Spencer, who applied the theory of evolution to social science and argued that “the survival of the fittest” should be given free rein in human society, both for moral and for practical reasons.³⁵⁷ He interpreted this as implying a libertarian approach to law making, where citizens should be left as free as possible to contract as they saw fit. Which, in turn, is exactly what the majority did in *Lochner*. In this way, evolu-

³⁵⁴ Cf. Williamson (1998), p. 37: “transactions, which differ in their attributes, are aligned with governance structures, which differ in their cost and competence, so as to effect a (mainly) transaction cost economizing result.”

³⁵⁵ *Lochner v. New York* (1905), 198 US 45, 25 S.Ct. 539, 49 L.Ed. 937.

³⁵⁶ 198 US 45, at 75.

³⁵⁷ Spencer (1851). This book is commonly credited with introducing the phrase “survival of the fittest”. However, while the book does spend ample time discussing the concept of fitness, “survival of the fittest” is not introduced by Spencer until Spencer (1864), Vol. 1, p. 444.

tion became widely accepted in the social sciences long before it ever became commonplace as a biological theory.³⁵⁸

As a matter of practicality, there is some light between the version of evolution first proposed by Lamarck in 1809, where individual animals can change their properties, and Darwin's evolution by natural selection, where they cannot.³⁵⁹ It has often been said that in the social sciences the model of evolution used resembles more the Lamarckian variety than the Darwinian, at least in practice if not in theory.³⁶⁰ In the present thesis, however, we will use evolution strictly in the natural selection sense, similar to how it is used in evolutionary game theory.³⁶¹ Evolution is the change a population through time due to natural selection. This means that individuals never evolve, but rather that the mix of individuals that make up the population changes because unfit individuals reproduce at a slower rate than their fitter counterparts. Since the subject of this thesis is governance structures, evolution is the alignment mechanism that works by successively removing the worst misalignments from the population of all governance structures, moving the average closer to the optimum. For evolution to work in this way, we would have to show that there is a correlation between the degree of alignment of a governance structure with the characteristics of its transaction and the fitness of this governance structure in the biological sense.

A classic example of evolution at work is described in Alchian (1950). He imagines a city where cars drive around randomly, stopping for petrol whenever they pass by a petrol station.³⁶² None of the cars have any intelligence whatsoever. They drive around at random. After a while, some of them start running out of fuel. If you wait long enough, the only cars left running are the ones who occasionally stop for petrol. Yet, even though the average car stops for fuel every few hundred kilometres, on average, this does not mean that one can infer that the average car is in any way rational. That was, in fact, the argument that Alchian was trying to make, arguing for neo-classical economic models against criticism by economists who had surveyed entrepreneurs and had found, to their apparent surprise, that businessmen do not, as a matter of fact, equate marginal costs and marginal benefits when they set their

³⁵⁸ It should be recalled that the famous Scopes Monkey Trial only took place in 1925, and that ended with evolution losing, a result upheld by the Tennessee Supreme Court in *Scopes v. State*, 152 Tenn. 424, 278 S.W. 57. Only in 1968 did the US Supreme Court rule that a ban on teaching evolution was a violation of the establishment clause, cf. *Epperson v Arkansas*, 393 US 97, 89 S.Ct. 266, 21 L.Ed.2d 228.

³⁵⁹ Lamarck (1809), Darwin (1859).

³⁶⁰ Cf. Baskerville (2007).

³⁶¹ Eg. Gintis (2000), chapter 9-11, and Gintis (2007), chapter 12-14.

³⁶² Alchian (1950), p. 214-215. Cf. Vromen (1950) for a detailed discussion.

prices.³⁶³ What's more, not only can we not argue against the occurrence of the theoretical optimum by pointing out the absence of rational behaviour, we can in fact confidently predict that the optimum will be achieved as long as we can point to something equivalent to cars running out of fuel if they do not stop for petrol. As long as suboptimal individuals are removed from the population at a faster rate than their optimal counterparts, the average will tend towards the optimum.

Another alleged example of evolution at work is found in the debate between proponents of common law and civil law. A famous early version of this argument is found in John R. Commons' *Legal Foundations of Capitalism*. In Chapter 2 of this book, Commons discusses the gradual change in meaning given to the word "property" in the Due Process clause of the Fourteenth Amendment to the US Constitution between its enactment after the Civil War and the early 20th century.³⁶⁴ In early cases, the Supreme Court adopted a restrictive reading of the term, limiting its meaning to physical property.³⁶⁵ However, in a series of cases, the initial minority view of Justice Field that the word property should cover anything that had value in trade, became the official rule.³⁶⁶ Commons argued that this was an example of (Lamarckian) evolution at work, since the broader reading of the clause more adequately reflected economic realities, and arguably strengthened the economy by fortifying the protection of the freedom to contract.

In other words, the hypothesis is that common law is a more flexible and decentralised system, and that this causes common law rules to be "better", in some sense, than their civil law counterparts. Whether this is true, as such, need not concern us at this point. The more interesting question is in what sense legal rules can be said to "evolve", and what the meaning of "better" would be in this context.³⁶⁷

A recent empirical study by Niblett, Posner and Shleifer (2008) found no evidence of the alleged superiority of common law. Or rather, they found that "the law did not converge to

³⁶³ Eg. Lester (1946).

³⁶⁴ Commons (1924), p. 11-46. The Due Process clause reads in its entirety: "nor shall any State deprive any person of life, liberty, or property, without due process of law." Because the preceding Privileges and Immunities clause has been interpreted almost out of existence, the Due Process clause is the most important constitutional protection against government interference with a variety of traditional liberties that are not explicitly guaranteed elsewhere.

³⁶⁵ *Slaughter-House cases*, 83 US (16 Wall.) 36, 21 L.Ed. 394 (1873). The ruling concerns three consolidated cases, all three coming from New Orleans.

³⁶⁶ The earliest such case is the *Minnesota Rates case* of 1890, *Chicago, Milwaukee & St. Paul Ry. Co. v. Minnesota*, 134 US 418.

³⁶⁷ Particularly given the thoughtless ease with which evolution is often invoked in economics, it is important to be clear about what is assumed to be evolving how. Cf. Baskerville & O'Grady (2007).

any stable resting point, but evolved differently in different states” and that “legal evolution is influenced by plaintiffs’ claims, the relative economic power of the parties, and nonbinding federal precedent.”³⁶⁸ Being economists, these authors defined “better” as “more efficient”, which they defined further with a short economic analysis of what the optimal rule would be for the “economic loss rule”, the rule they were focusing on. Taking this optimum, they hypothesised that the law in different states should converge towards it. When this was found not to occur, the evolution of the common law was declared dead, at least in the area under consideration.

Obviously, this is not an approach that would work in all circumstances. In many areas of law, there is no more reason that the laws of different states should converge than that the design of the limbs of whales and monkeys should. Convergence is only a reasonable hypothesis when the circumstances are largely the same in the different states under consideration. In construction law, where Niblett, Posner & Shleifer got their data, that is probably a reasonable assumption, whereas in transportation law, which we will discuss in the next chapter, it probably is not.

Another thing that is notable about this paper is the conceptual confusion that ensues if one talks about rules evolving. It may seem like somewhat of a metaphysical question, but if a rule is changed, is it not more precise to say that it has been replaced by a different rule? The answer is, of course, that these alternatives are the Lamarckian and Darwinian formulation of exactly the same phenomenon. The two versions of evolution have become indistinguishable. In fact, following down this path leads to a field of study called memetics, which was pioneered by Richard Dawkins, almost as an afterthought, in the concluding chapter of his classic book on evolution, *The Selfish Gene*.³⁶⁹ What he proposed is that it should be possible to analyse the evolution of certain cultural items, which he called memes by analogy to genes, using Darwin’s theory of evolution by natural selection. Unfortunately, after 30 years of debate about this idea, it is still unclear whether such a model would contribute to our understanding of human society, or even whether such a model could ever rise above the level of an amusing yet ultimately unscientific analogy.³⁷⁰

A better solution is to restrict the evolution metaphor to things that have an existence over and above their definition. Just like biologists do not talk about genes evolving, social

³⁶⁸ Niblett, Posner & Shleifer (2008), p. 1.

³⁶⁹ Dawkins (1976).

³⁷⁰ Cf. Distin (2005), who surveys the first 30 years of debate.

scientists should not talk about rules evolving.³⁷¹ Instead, the thing that evolves is an institution, which is defined by its rules, just like a species in biology is largely defined by its genes. An institution, however, is a regrettably poorly defined concept, which is why it has not been used yet in this chapter so far. Defined at its broadest, institutions include rules as well. Schumpeter, for example, defined institutions as “all the patterns of behavior into which individuals must fit under penalty of encountering organized resistance, (...) not only legal institutions, such as property or the contract, and the agencies for their production or enforcement.”³⁷² While this definition has much to commend itself, including the fact that it explicitly extends beyond the law and the fact that it does not refer to any other concepts from institutional economics, it is clearly broad enough to include any and all legal rules.

Another common approach is to distinguish between institutions, narrowly defined, and organisations. Douglass North, for example, wrote that institutions are “the humanly devised constraints that shape human interaction”³⁷³, which is a definition similar to the one Schumpeter proposed, but then he added the concept of an organisation, i.e. a “group of individuals bound by some common purpose to achieve objectives.”³⁷⁴ In other words, institutions are “the rules of the game”³⁷⁵ and organisations are the players.³⁷⁶ Even though it may be somewhat counterintuitive to decide that organisations are not institutions, this solution allows for the closest possible analogy to the biological original, at least in theory. Organisations are equivalent to species, defined by their institutions in the same way that species are defined by their set of genes.

It follows that evolution occurs when the median institutions that are used in the population of organisations tends to move towards the optimum because those organisations that have suboptimal institutions are more likely to cease to exist in any given period. Rather than searching for evidence of evolution by looking for convergence, we will search for it by considering whether there is sufficient “evolutionary pressure” to make a noticeable difference.

³⁷¹ There is another possibility. For example, when Greif (2008) begins his article by writing that “little is known about why institutions gradually evolved to support a larger set of exchange relations in some economies and not others”, he does not mean to invoke any specific biological analogy, but rather uses “evolved” in its meaning of “developed gradually”. Cf. <http://www.merriam-webster.com/dictionary/evolve>, definition 2c. (Last visited on January 9, 2009.)

³⁷² Schumpeter (1983), p. 191.

³⁷³ North (1990), p. 3.

³⁷⁴ North (1990), p. 5. “Organisations include political bodies (political parties, the Senate, a city council, a regulatory agency), economic bodies (firms, trade unions, family farms, cooperatives), social bodies (churches, clubs, athletic associations), and educational bodies (schools, universities, vocational training centres).”

³⁷⁵ North (1990), p. 3.

³⁷⁶ North (1990), p. 4: “The purpose of the rules is to define the way the game is played. But the objective of the team within that set of rules is to win the game.”

2.6.3 Competition

Much of what was said above applies equally to competition. For the purposes of the present paper, competition is defined similar to Lamarckian evolution, meaning that the organisations that compete in the market are forced by competitive pressure to change their institutions so as to maximise efficiency. In other words, compared to evolution proper, competition causes the median institutions to change not by removing organisations from the population, but by having them change. As before, competition between different possible institutions is a meaningless notion. Instead, the cause of the change is competition between organisations. We distinguish between competition within a market by organisations producing essentially the same product and competition between related products.

Consider the classic GM/Fisher Body case study that was already discussed in section 2.4.1, above. Why was GM concerned about the design of their contractual relations with Fisher Body? Because they were trying to compete with Ford, and to do so meant offering the best possible product for the lowest possible price. In order to be able to do that, they had to make sure that they, in turn, were getting the best possible product for the lowest possible price from their suppliers, which meant designing an efficient governance structure for their transactions with Fisher Body.³⁷⁷ A similar explanation can be advanced for the increasing tendency of supermarkets to backward integrate into product marketing by selling own-label products: separating the marketing of products from their physical production is an example of innovation of governance, which supermarket chains were forced to explore due to the competitive pressures in their industry.³⁷⁸

Competition between producers of related products, on the other hand, yields ever weaker competitive pressure as the degree of relatedness of the products, as measured for example by their cross-price elasticities of demand, becomes lower.³⁷⁹ While competition law insists on clear boundaries between markets, in reality many products that are clearly different

³⁷⁷ Absent this competitive pressure, the management of GM may well have preferred to avoid conflict rather than maximize profit at all costs.

³⁷⁸ Supermarkets' transactions with their suppliers were already discussed in section 2.4.1, above. For a discussion of the rise of own-label products, cf. BCG (2007).

³⁷⁹ Cf. Commission Notice on the definition of relevant market for the purposes of Community competition law, OJ C 372/5, d.d. 09.12.1997, par. 15-19.

still serve to satisfy the same need.³⁸⁰ Whether two products belong to the same market is a matter of degree. Do all vitamin supplements make up a single market, or does each type form a separate market?³⁸¹ Perhaps it would be better to say that vitamin supplements compete to some extent with other foodstuffs that are known to be healthy. And what about railways?³⁸² Is every line a separate market? Or an entire network of lines? Or even the entire transport sector?³⁸³ Despite this problem, it is still useful to make the distinction between these two forms of competition, if only for analytical clarity.

Finally, there is one aspect of competition that requires further consideration. In ordinary circumstances, the concept is fairly well understood by economists. However, since much of the remainder of this thesis is devoted to state action, it is necessary to make sure we understand in what sense states compete.

The most basic answer is that states do not compete.³⁸⁴ Unlike companies, states cannot go bankrupt in any real sense, and as such they are spared the usual punishment for a lack of competitiveness. Unlike companies that compete, states are not looking to supply the same market, but rather they buy and sell from each other, making their economic relations a positive sum game. Everyone benefits from more international trade. When one country has economic success, in the sense of having a relatively high GDP growth rate, this will increase its spending power on the world market, causing part of its success to be spread to other countries. Because of this, for most modern capitalist democracies, GDP growth is positively correlated. And even when a country is not very “competitive” on the world market, most of its products are sold to its own citizens, isolating it from whatever competitive pressures may otherwise exist.

One exception is when states find themselves in winner take all situations. The most important such situation is war. In 1688, the Dutch were worried that they would find themselves fighting a war against France and England simultaneously again, as they had 16 years prior. Despite being the richest country in the world at the time³⁸⁵, the 1672-1678 French-

³⁸⁰ Cf. Commission Notice on the definition of relevant market for the purposes of Community competition law, OJ C 372/5, d.d. 09.12.1997, par. 7.

³⁸¹ Case 85/76, *Hoffmann-La Roche & Co v Commission*, [1979] ECR 461, par. 21-30.

³⁸² Eg. Case T-229/94, *Deutsche Bahn v Commission*, [1998] ECR II-1689, par. 53-59.

³⁸³ Cf. Chapter 3, below.

³⁸⁴ Cf. Krugman (1994, 1996).

³⁸⁵ The data collected by Angus Maddison of the Groningen Growth & Development Centre of the University of Groningen (www.ggd.net) show that the Dutch GDP per capita was 152% of the average of the 12 Western European countries in the sample in 1600, and 206% in 1700. The present day UK, on the other hand, stood at 107% in 1600 and 121% in 1700, making it the second richest country in the sample. Calculated from:

Dutch War was one they could not win. So what the Dutch Stadtholder, William III, decided to do, was to leverage his unique position and his country's uniquely efficient institutions in order to put his country in a better position. While the history of the Glorious Revolution can safely be assumed to be well known, and is in any event outside the scope of this thesis, it is important to realise that none of it would have been possible without the Dutch state's ability to borrow a truly breathtaking amount of money in almost no time at all.³⁸⁶ The reason why they were able to do so was that Holland, with its more or less democratic system, was governed by the same people who were now lending it the money, the wealthy regent and merchant classes. This, combined with the presence of a relatively independent court system and other institutional safeguards against moral hazard on the part of the government, greatly enhanced the Republic's ability to raise money for war.³⁸⁷

Both to solidify their reign and to enhance England's ability to fight the war against France³⁸⁸, William and Mary agreed to a number of reforms, such as the Bill of Rights³⁸⁹, the Act of Toleration³⁹⁰ and the Bank of England Act 1694³⁹¹, which turned England into a constitutional monarchy, with all the concomitant economic benefits.³⁹² In later centuries, these reforms allowed England to become the world's dominant military force.³⁹³

Another example of political and military competition forcing a country to change its institutions is the Soviet-Union under Gorbachev. In a totalitarian state, the mere fact that communism manifestly did not work is not enough of an explanation for the fact that it was reformed.³⁹⁴ Instead, the explanation is that the inefficiency of the system had gotten so bad that it was interfering with the Soviet-Union's ability to keep up with the West in the Cold War arms race.³⁹⁵ In trying to reform the system, Gorbachev ended up destroying it, making him an example of evolution as well as competition.

www.ggdc.net/Maddison/Historical_Statistics/horizontal-file_03-2007.xls. (Last visited on January 13, 2009.)

Cf. North & Thomas (1973), chapter 11.

³⁸⁶ Cf. Israel (1991).

³⁸⁷ Cf. Stasavage (2007), particularly p. 503.

³⁸⁸ The War of the League of Augsburg started with a French declaration of war against The Netherlands in September 1688. In May 1689, the newly Williamite England joined the war on the Dutch side.

³⁸⁹ 1 Will. & Mar. session 2, c. 2.

³⁹⁰ 1 Will. & Mar. session 2, c. 18.

³⁹¹ 5 & 6 Will. & Mar., c. 20.

³⁹² Cf. North & Weingast (1989).

³⁹³ Cf. Kennedy (1987).

³⁹⁴ Eg. Kennedy (1987), p. 635-636, explains why the Soviet leadership would want to maintain a system of agriculture that was so inefficient that the Soviet Union actually had to import food.

³⁹⁵ Cf. Kennedy (1987), p. 631-664, particularly p. 647-650, Naylor (1988), p. 15-19, and, between the lines, Gorbatsjov (1987).

Analytically, the difference between economic competition and military competition is, of course, that the latter is a winner-take-all game.³⁹⁶ In war, as in rent-seeking, there is no benefit in coming in second. Instead, increased investments only lead to higher probabilities of success.³⁹⁷ In the end, this mechanism tends to produce arms races, both of the military hardware and of the institutional variety, which tend to force the outcomes for both closer to the point of maximum efficiency.

Another winner take all situation that states might find themselves in concerns investors' choice of where to invest, at least given free movement of capital and goods. Geerdink & Stauvermann (2006), for example, analyse the transition of two autarkies into one integrated economy, and show that at least one of the countries will have great difficulty maintaining its labour market regulations.³⁹⁸ This immediately shows the problem, for present purposes, of this type of competition: while the states involved certainly compete, their competition moves them in the direction that pleases the more mobile factor of production, which is not the same thing as a move towards greater efficiency, *per se*.³⁹⁹

Before continuing to discuss rational design as an alignment mechanism, it is useful to note that competition can also work as an alignment mechanism absent any "real" competition between states insofar as politicians incorrectly perceive their countries to be competing.⁴⁰⁰ Since it violates our assumptions regarding rationality⁴⁰¹, we will ignore this possibility for now, but we will return to it in chapter 5.

2.6.4 Design

Even 300 years after Bernard de Mandeville published his fable of the bees, most people are still unable to shake the illusion that the existence of order implies that there is someone doing

³⁹⁶ Cf. Hirschleifer (1988) and Stauvermann (2002).

³⁹⁷ Cf. Hirschleifer (1989).

³⁹⁸ The same goes for many rules of company law, such as capital adequacy. Cf. ECJ Case C-167/01, [2003] ECR I-10155, dd. 30 September 2003, Inspire Art (ruling that the Netherlands could not enforce its rules on capital adequacy against a company registered in another Member State, even though the company in question was only active in the Netherlands).

³⁹⁹ A model of this kind of regulatory competition was first developed by Tiebout (1956). Cf. Davies (2006) for a legal analysis of the concept of regulatory competition. Cf. Geerdink & Stauvermann (2009) for a model of regions competing for investments by using subsidies.

⁴⁰⁰ Cf. Krugman (1994).

⁴⁰¹ Cf. section 1.3, above.

the ordering.⁴⁰² In the book of Genesis, as in Greek mythology, in the beginning there is only chaos, “the earth was without form and void”.⁴⁰³ Only when God begins by separating the light from the darkness, the heaven from the sea, the sea from the land and so on does order emerge.⁴⁰⁴ In reality, however, some of the most important breakthroughs in many fields of science have come from the realisation that spontaneous order is, indeed, very well possible. Most notably, this goes for Adam Smith’s invisible hand and Charles Darwin’s theory of evolution, both of which we have already encountered in this chapter. What is left is to examine the reverse question: in what sense can we expect the existence of a designer to lead to order?

To begin with the obvious, any design is necessarily going to result in order, in the meaning of “a regular or harmonious arrangement”.⁴⁰⁵ As we have seen in section 2.1, any design by a boundedly rational lawmaker is inevitably going to result in an incomplete law.⁴⁰⁶ As we have seen, such a law lumps categories of people and categories of situations together even though a theoretical case might be made for treating them with a higher level of detail. While a “complete law” would be extremely detailed, or even infinitely so, any real law is simpler than that, and from this it follows that any design imposes order.⁴⁰⁷ In fact, it follows that in many cases the design imposes too much order.

However, instead of considering whether a design would impose the optimal level of order, the question here is whether it would impose the optimal type of order. What do we know about lawmakers’ ability to properly align governance structures with transaction characteristics? Traditionally, the place to start looking for the answer to such questions is Hayek’s article ‘The Use of Knowledge in Society’.⁴⁰⁸ Hayek argued that the notion of a centrally planned economy, as advocated by Marxism, was fundamentally flawed because it required the planner to take into account a large amount of information about personal preferences that

⁴⁰² *The Fable of the Bees, or Private Vices, Publick Benefits* was first written in 1705, and originally published in 1714. The poem is available on-line at <http://pedagogie.ac-toulouse.fr/philosophie/textes/mandevillethefableofthebees.htm>. (Last visited on January 17, 2009.) Cf. Bick (2008).

⁴⁰³ Genesis 1:2, King James Version. The Greek-Roman story of origin is found, for example, in book I of Ovid’s *Metamorphoses*. Ovid, in turn, took it from the poet Hesiod, who lived around 700 B.C.

⁴⁰⁴ Genesis 1:4, 7 and 9, respectively.

⁴⁰⁵ [http://www.merriam-webster.com/dictionary/order\[2\]](http://www.merriam-webster.com/dictionary/order[2]), definition 5b. (Last visited on January 17, 2009.)

⁴⁰⁶ N.B. it is not strictly necessary that the designer should be a lawmaker. However, it has to be someone with some kind of authority over the market for the outcome to be properly called a design. This includes rules created by professional bodies, covenants agreed by market participants and many similar phenomena. However, for ease of reference, we will focus on laws and lawmakers in the remainder of this section.

⁴⁰⁷ Obviously, this also follows from the definition of design, which necessarily involves an element of planning. Cf. <http://www.merriam-webster.com/dictionary/design>. (Last visited on January 17, 2009.)

⁴⁰⁸ Hayek (1945).

was not readily observable and that the government could not possibly get its subjects to reveal truthfully.

In the same vein, we might wonder whether the key variables that determine the optimal governance structure are readily observable to the lawmaker. Recalling the distinction made in section 2.1.5 between the observability of certain facts and their verifiability, it clearly matters a great deal whether the lawmaker is called upon to legislate a transaction that the state, itself, is a party to, or whether the legislation concerns transactions that are wholly private. When it comes to the latter, lawmakers have generally stepped back and focused on facilitating the choices of private parties instead of establishing the governance structures itself. It is widely believed that the retreat from the government away from legally established monopolies and legally protected guilds towards civil codes that gave citizens legal alternatives to choose from is the main reason for the staggering increase in the standard of living since the Middle Ages.⁴⁰⁹ In the present day, there are many laws that impinge on the citizens' freedom to contract, but only competition law affects private governance choices directly.

In the last few decades, many economists have expressed their views on competition law. Oliver Williamson, for example, devoted ample space to the topic in both of his major theoretical books.⁴¹⁰ Judge Posner has returned to the issue throughout his career⁴¹¹ and the same goes for many other economists affiliated with the University of Chicago.⁴¹² Usually, they have been critical, arguing for less government intervention. However, since one may well feel that many of these authors have ignored the practical difficulties involved in the application of institutional economics or law & economics to real-life policy problems in order to make an ideological case, we will not review their criticisms at this point.⁴¹³⁴¹⁴ Suffice it to say that the Warren Court's use of per se illegalities, where certain behaviours were declared to be automatic violations of the Sherman Act, is justly criticised for banning many welfare-

⁴⁰⁹ Cf. North & Thomas (1973) and Smith (1776). While guilds are mostly associated with the Middle Ages, they continued to exist in most European countries until their abolition in the wake of the French revolution.

⁴¹⁰ In Williamson (1975), most of the later chapters have a section called "policy implications" or "antitrust implications", and in Williamson (1985), all of chapter 14 is devoted to the issue. Cf. also Williamson (1972, 1977, 1987).

⁴¹¹ Eg. Posner (1969, 1974, 1979, 1981, 2001, 2005) and Landes & Posner (1981), as well as Posner (1973), chapters 9-13.

⁴¹² Eg. Bork (1966), Easterbrook (1984), Friedman (1999), Stigler (1966), Stigler (1968), chapters 20-22, as well as the entire special issue on antitrust law of the University of Chicago Law Review, Vol. 72, No. 1. For a discussion, cf. Jones & Sufrin (2008), p. 23-30.

⁴¹³ Cf. Schumpeter (1949), p. 346: "[I]n itself scientific performance does not require us to divest ourselves of our value judgement or to renounce the calling of an advocate of some particular interest."

⁴¹⁴ Of course, much of this anti-interventionist criticism can also be explained by the fact that it was written against the backdrop of the highly interventionist Warren Court. Cf. Page (2008) and Peritz (1996).

enhancing practices.⁴¹⁵ However, as the theory would predict, there are significant information problems associated with the more flexible rule of reason approach⁴¹⁶, which is why in certain fields the use of per se rules may well be the least inefficient solution. This is not a question that needs to be resolved here.⁴¹⁷

Obviously, the information problems involved in deciding the optimal competition laws far exceed the problem of deciding what the most efficient governance structure would be for a transaction the government itself is involved in. Nevertheless, in combination with the problem of deciding which policy goals the government should pursue in a given field, and how these goals should be balanced⁴¹⁸, the net result is a task of similar complexity to the task of writing and enforcing the competition laws, involving many of the same questions. As a result, many of the same concerns voiced in relation to competition law also apply to state intervention in Services of General Interest.⁴¹⁹ However, since such services tend to differ greatly from one to the other, analysis tends to be limited to a single market. We have already discussed Oliver Williamson's case study of cable television in Oakland⁴²⁰, where the city council's ability as a designer of efficient governance structures left somewhat to be desired. Another thoughtful study is Ronald Coase's 1959 article about the US Federal Communications Commission.⁴²¹

Unlike the other two, design critically depends on the ability of the designers to form correct beliefs about the transaction and production efficiency of certain governance structures. But what if the relevant actors are unable to do so, either because of their ideology gets in the

⁴¹⁵ Many of those per se rules have since been repealed. For example, the per se rule against restricted distribution based on territory, which was announced in *US v Schwinn*, 388 US 365, 87 S.Ct. 1856, 18 L.Ed.2d 1249 (1967) was repealed nine years later in *Continental TV v GTE Sylvania*, 433 US 36, 97 S.Ct. 2549, 53 L.Ed.2d 568. Cf. Posner (1981).

⁴¹⁶ Eg. when the District Court issued its findings of fact in the US antitrust case against Microsoft, its findings condensing thousands of pages of submissions still ran more than 100 pages in the F.Supp. *US v Microsoft*, 84 F.Supp.2d 9 (Distr. of Columbia, November 5, 1999). The 207-page pdf version is available at <http://www.usdoj.gov/atr/cases/f3800/msjudge.pdf>. (Last visited on January 20, 2009.)

⁴¹⁷ Winston (2006), p. 13-21 reviews a body of empirical literature on antitrust law and policy, concluding that there is little evidence of a clear benefit to society. One might note, however, that Mr. Winston was affiliated with the AEI-Brookings Institute.

⁴¹⁸ In other words, deciding what the government's objective function is. Cf. chapter 1.3, above, and Mueller (2003), chapters 23-27.

⁴¹⁹ Cf. Communication by the Commission of the European Communities of September 26, 1996, OJ C281/3: "Services of General Interest (...) are market and non-market services which the public authorities class as being of general interest and subject to specific public service obligations." Cf. also art. 16 and 90 EC and section 4.3, below.

⁴²⁰ Section 2.4.3, above.

⁴²¹ Coase (1959).

way? Among political scientists, ideology is often defined as a coherent set of ideas about society⁴²², these ideas being beliefs and values.⁴²³ Solidarity, for example, is not an ideology in itself, but it is a value present in different degrees in different ideologies.⁴²⁴ In economic language, we can treat it simply as the amount of weight given in person X's utility function to the utility of all other persons, i.e. as a preference. The same goes for other values, although it is not always easy.⁴²⁵

For all preferences goes that they are inevitably traded against each other. Fulfilling one preference comes at a cost with respect to all other preferences a person holds.⁴²⁶ The relative importance of preferences can change, and change in a manner that is exogenous to the model. Since the Coasian approach to institutional economics still relies on a neo-classical logic, i.e. the logic of optimising behaviour subject to constraints, it can deal with values and their changes as well or as poorly as its neo-classical cousins.

On the other hand, there is the problem with beliefs, taken here as statements of fact to which a truth value could theoretically be assigned.⁴²⁷ Particularly beliefs about the connection between ends and scarce means – the traditional core subject matter of economic science⁴²⁸ – are problematic. Consider Friedrich Hayek's *The Road to Serfdom*.⁴²⁹ This book essentially does three things at once: it argues that state intervention in the economy is undesirable, which is a value; it argues that state intervention in the economy is inefficient, which is a belief; and it provides economic evidence for said belief. Of course, to the extent that Hayek's economic argument is correct, there is no problem.⁴³⁰ However, where this belief in the power of free markets is not justified by the facts, it should be considered an ideological belief rather than a scientific one.

⁴²² Cf. Kennedy (1979).

⁴²³ Cf. Schumaker (2008).

⁴²⁴ Cf. Sternø (2004).

⁴²⁵ Cf. Berlin (1969) on liberty.

⁴²⁶ The economic modelling of preferences and consumer choices is discussed, for example, in Cowell (1986), chapter 4.

⁴²⁷ This is a much more dubious definition than might appear. Cf. Arthur v. Offit, District Court for the Eastern District of Virginia, 10 March 2010, unpublished, where the statement "she lies" was found to be a statement of opinion not susceptible to a libel prosecution.

⁴²⁸ Cf. section 1.2, above.

⁴²⁹ Hayek (1944).

⁴³⁰ While arguments from authority have no place in science, it is perhaps useful to remember that Friedrich Hayek received the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel in 1974 for, amongst other things, his "penetrating analysis of the interdependence of economic, social and institutional phenomena."

In fact, even this is too simplistic. Hayek's book is a milestone in the history of economic science, containing an argument in support of the author's beliefs that reflected a careful understanding of the body of economic knowledge at the time of writing. If some of his claims have subsequently been proven wrong, this is because both scholars and practitioners used his writings as a foundation for their work. Surely it would be unfair to describe those of his beliefs that have subsequently been proven to be unfounded as "ideological", even if they represented the best economic science could do when they were written?⁴³¹

This is not a merely hypothetical question. As we shall see in the next chapter, both the initial arguments for organising the railways privately and the subsequent arguments for their consolidation and nationalisation were based at least in part on the scholarly writings of the time. When the Labour Party wrote in its 1945 election manifesto that "its ultimate purpose at home [was] the establishment of the Socialist Commonwealth of Great Britain-free, democratic, efficient, progressive, public-spirited, its material resources organised in the service of the British people", this was at least in part a statement of belief, based on a not unreasonable understanding of the economic theory of the time.⁴³²

What is left is to consider the boundaries between design and competition or evolution. When a lawmaker designs governance structures for a market that he is not otherwise involved in, it is clear that we are talking about design. The antitrust laws are an example of this. However, when the state acts both as lawmaker and as market participant, the difference may well be less clear. Moreover, since the alignment mechanisms usually operate side by side, the differ-

⁴³¹ Cf. North (2005), p. viii.

⁴³² Earlier in the same manifesto, they wrote for example: "the whole of the national resources, in land, material and labour must be fully employed. Production must be raised to the highest level and related to purchasing power. Over-production is not the cause of depression and unemployment; it is under-consumption that is responsible. It is doubtful whether we have ever, except in war, used the whole of our productive capacity. This must be corrected because, upon our ability to produce and organise a fair and generous distribution of the product, the standard of living of our people depends.

(...) A high and constant purchasing power can be maintained through good wages, social services and insurance, and taxation which bears less heavily on the lower-income groups. But everybody knows that money and savings lose their value if prices rise, so rents and the prices of the necessities of life will be controlled.

(...) Planned investment in essential industries and on houses, schools, hospitals and civic centres will occupy a large field of capital expenditure. A National Investment Board will determine social priorities and promote better timing in private investment. In suitable cases we would transfer the use of efficient Government factories from war production to meet the needs of peace. The location of new factories will be suitably controlled, and where necessary, the Government will itself build factories."

This is, of course, wrong, but it is less clear whether these arguments, when they were made, reflected beliefs or values. The debate over the "Socialist Calculation" took place mostly in the 1920s and 30s, but the question was by no means settled in 1945. Cf. Backhouse (2003) and the literature cited there.

ence is not very important. Nevertheless, we can start by observing that design is a unilateral act that does not require the consent of all actors involved in the transaction.⁴³³ Either a lawmaker or some other body with sufficient authority sets the rules for all to follow. Also, design implies careful study of the problem, usually leading to non-incremental changes. Both evolution and competition are supposed to lead to incremental changes, where someone tries something, which then turns out to work or not.⁴³⁴ Design, on the other hand, is capable of leading to a drastic overhaul of the sector, when this is found to be necessary.

The confluence of design and competition can occur when the jurisdiction of the lawmaker is smaller than the geographical market. In that case, there is competitive pressure on the lawmaker to design efficient laws, or at least laws that help it or its citizens compete. In this sense, Gorbachev's reforms in the 1980s are an example of design as well as competition and evolution. Similarly, the need to attract foreign investors may influence the lawmaker's design, as it did for the many countries that enacted pro-market reforms in recent decades as a requirement for IMF aid.⁴³⁵

2.6.5 Conclusion

The reason why it is so useful to carefully consider the alignment mechanisms that are necessary for NIE to work as a descriptive model is illustrated by Holterman (2009). That paper discusses the first introduction of rail transport in the German Ruhr area in the 1830s, a subject we will return to in section 3.1, below. Rather than considering directly whether the governance structure chosen in that case was efficient – something that is nigh on impossible to do – the approach chosen in that paper was to examine whether there was reason to believe that the alignment mechanisms that could be identified would be strong enough to push the outcome towards the theoretical optimum. The conclusion was that they were not, a conclu-

⁴³³ Cf. section 2.1, above.

⁴³⁴ In biology, there is still debate on whether evolution involves only gradual change, or whether it also works through “jumps” from time to time. Eldredge & Gould (1972) developed a model of punctuated equilibrium, which involves occasional rapid yet gradual evolution. It is discussed for a lay audience in Dawkins (1986), chapter 9.

⁴³⁵ Cf. Jeanne, Ostry & Zettelmeyer (2008). Obviously most of such reforms do not concern “governance structures” in any direct sense, but a substantial minority do, such as the abolition of limits on foreign ownership of factories and other real estate, or the privatisation of government-controlled corporations.

sion supported by the subsequent history of frequent reform.⁴³⁶ That way, it was possible to apply the model to the case at least tentatively even though the case was essentially unique, pre-empting the usual approach of conducting empirical studies of neo-institutional economics by comparing – either by statistics⁴³⁷ or by case study⁴³⁸ – a number of similar cases. A similar approach will frequently be used in the remainder of this thesis.

⁴³⁶ Cf. Chapter 3, below.

⁴³⁷ Eg. Ménard & Saussier (2002).

⁴³⁸ Eg. Van Genugten (2008).

2.7 Conclusion

The previous six sections have presented the different elements that make up a viable transaction cost theory of government intervention in the economy, specifically intervention through ownership and similar control rights. What is left is to consider how they all fit together and which elements that uniquely identify these models might be observable in the case studies of the next chapter. The goal is to provide first of all some kind of synthesis, before asking how the synthesis version of the new institutional economics, as well as its component parts, connects to the historical narrative of chapter 3: If the models work as advertised, what would we expect to see? This section concludes with a checklist that we will refer to on several occasions in the next chapter.

The model begins with the inevitable shortcomings of human beings, as summarised in the notion of bounded rationality. As explained in chapter 1 as well as in the present chapter, the approach to bounded rationality in this thesis is a relatively modest one. Rather than adopting Herbert Simon's recommendation to incorporate insights from psychology⁴³⁹, bounded rationality is defined here as behaviour that is "intendedly rational, but limitedly so"⁴⁴⁰ in the sense of being fully rational with respect to limited information. The information people use to take decisions is limited because gathering and using information is costly. Those costs are, of course, transaction costs.

Because gathering and using information is costly, all contracts are incomplete. They either do not define any action set for certain future states of the world, or they define an action set that is (*ex ante*) inefficient. While the degree of incompleteness of a contract is almost by definition impossible to ascertain for any third party, including a researcher, it is often possible to get a first impression by considering the level of complexity of the problem.⁴⁴¹ Defining it loosely, we can say that complexity is increasing in the number of actors involved, in the number of relevant states of the world and in the evenness of their probability distribution.⁴⁴² From this it follows immediately that, *ceteris paribus*, incompleteness is increasing in

⁴³⁹ Eg. Simon (1947, 1978).

⁴⁴⁰ Simon (1947), notably repeated in Williamson (1985).

⁴⁴¹ Cf. Bajari & Tadelis (2001).

⁴⁴² A formal definition along these lines is given in Bajari & Tadelis (2001), p. 392-393.

complexity. Unfortunately, all else is rarely equal, since the costs of gathering and using information varies greatly depending on the context. Also, operationalising complexity is just as impossible as operationalising incompleteness, since in practice the only thing that can be observed is the complexity of the contract⁴⁴³, which differs from the true complexity of the problem in proportion to the degree of incompleteness of the contract. In other words, talking about complexity instead of incompleteness is not very helpful.

Nevertheless, when the goal is to describe vertical integration, incompleteness and complexity are a good place to start. First of all, in many situations parties and even outsiders are fairly well aware of what it is that they do not know. If the parties are aware, for example, that the states of the world that they have dealt with in the contract have a total probability of 80%, that gives them a first approximation of its degree of incompleteness, even if they have no idea about what might happen the other 20% of the time.

Secondly, in its descriptive incarnation, the model will necessarily rely on the perceptions of actors in order to explain their decisions. Their chosen governance choices are explained by reference to the exposures they perceive, which, in turn, are explained by reference to the perceived areas of incompleteness of any contractual solution. In other words, the argument goes from perceived incompleteness to perceived exposures, and from perceived exposures to governance decisions.

Regarding these exposures, it will be useful to distinguish two categories: Exposures due solely to a lack of verifiability of certain information, and exposures that are also caused by a lack of observability. We will call the former hold-ups, and the latter moral hazard. In case of hold-up, such as the infamous Fisher Body case, there is little or no information asymmetry between the parties; the only problem is that third parties, such as judges and juries, cannot verify certain information. In a moral hazard situation, at least one of the parties does not know exactly how much effort the other is investing⁴⁴⁴, or generally whether or not they are behaving in the joint value maximising manner.⁴⁴⁵

Given such exposures, the hypothesis is that actors should and will respond to greater exposures by setting up a more integrated governance structure. “More integrated”, of course, suggests a one-dimensional perspective on integration, which we have already shown in sec-

⁴⁴³ Cf. Eggleston, Posner & Zeckhauser (2000).

⁴⁴⁴ Holmström (1979).

⁴⁴⁵ The other classic example of moral hazard, other than employee shirking, is excessive risk-taking by people who have bought insurance. Eg. Crosby (1905).

tion 2.4.2, above, to be flawed. So instead a more careful formulation would be that actors will set up a governance structure that is tailored to the exposures they perceive, balancing transaction efficiency against production efficiency.

In agency theory, that means that a greater exposure due to moral hazard is mitigated by giving the agent a greater residual claim on his output.⁴⁴⁶ However, giving him too large a claim, as a proportion of the agent's total remuneration, is inefficient due to the agent's risk adversity; the greater the uncertain component in the agent's remuneration package, the greater the total value of the package will have to be in order to meet the agent's reservation wage. In that way, classic agency theory is an example of this balancing of transaction and production efficiency: increasing the size of the agent's residual claim improves the former while hurting the latter.⁴⁴⁷ Because agency theory can so easily be cast in neo-institutional terms, it has become common to view it as a sub-field of NIE, even though its origins are somewhat different. It is for this reason that, when we consider alternatives to NIE in chapter 4, agency theory will not be among them.

Looking finally at the role of the state in our framework, we see that the model as such can be applied to the state without difficulty. It contains no variables or assumptions that, by their nature, can only apply in the private sector. As already discussed in section 1.2, above, the unique nature of the state compared to other actors enters into the analysis firstly in the objective function, which describes what the actor wishes to accomplish and what their priorities are, and secondly in the decision what to do with the recommendation of the normative economic model, which by its nature only considers efficiency.⁴⁴⁸ The government is more likely than a private actor to let other values than efficiency play a role in the decision making, but this does not mean that a model that focuses only on efficiency is worthless, or that it would necessarily improve, either as a normative model or as a descriptive one, if it took into account other "public interests".⁴⁴⁹

Given this summary, we can identify the following elements as particularly significant. To the extent that these elements are present, the new institutional economics is validated as a useful

⁴⁴⁶ Cf. Hendrikse (2003), p. 87-102 and Milgrom & Roberts (1992), p. 166-204.

⁴⁴⁷ Note that this solution involves a "separation of ownership and control", Fama & Jensen (1983), meaning that it is not simply a case of partial integration. Instead of advocating worker-controlled organisations, the theory suggests a governance structure that is tailored to the problem. Cf. Klein (1983) and Williamson (1983).

⁴⁴⁸ Cf. section 1.3, above.

⁴⁴⁹ Cf. Van Genugten (2008), p. 4-6 and WRR (2000).

explanatory framework for semi-public institutions, that is to say that these elements are necessary and sufficient in order to make application of the models possible in the first place. The further question of their explanatory power can then be answered by confronting their predictions with reality and comparing them to each other and to the theories of chapter 4 in terms of efficacy.

Commons and Coase

- If transactions are the “basic unit of analysis”⁴⁵⁰, that fact must be obvious from the historic narrative. Given that we have ruled out transactions occurring within an organisation, transactions must occur only between the government and one or more private parties. These transactions must be properly so called, in the economic sense, meaning that they must be between (boundedly) rational actors, and based on negotiations.⁴⁵¹
- These actors must be concerned with optimising the overall efficiency of their part of the value chain: Choices that reduce production efficiency must be defended by pointing to increased transaction efficiency, and vice versa. Ideology plays a role to the extent that an ideology is a set of values, i.e. preferences.
- Information is costly. If the costs of searching, negotiating and enforcing⁴⁵² are to drive governance choices, the variation in these transaction costs between alternative governance options must be of the same order of magnitude as the variation in production costs, if not greater.
- A concrete example of the previous point is moral hazard: If enforcing contracts is costly, we would expect to see governance choices explained by fears of moral hazard.⁴⁵³

Williamson

⁴⁵⁰ Williamson (1985), p. 41.

⁴⁵¹ In other words, they must be “bargaining transactions”. Cf. Commons (1934) and section 2.3, above.

⁴⁵² Coase (1937). The transaction costs incurred after contracting consist of monitoring and bonding costs, since each party also has an incentive to demonstrate its own compliance, at least to some degree. Cf. Jensen & Meckling (1976).

⁴⁵³ Cf. Williamson (1988), where he discusses the commonalities and differences between transaction cost economics and agency theory. At their respective cores, of course both models have a lot in common, dealing as they do with efficient contracting. It is only in their later operationalisations, by Williamson and by Jensen & Meckling (1976), for example, that they start to diverge.

- A fundamental transformation takes place when the parties to a contract start carrying it out: through their investments in specific assets, they become locked in, changing the game from a one-to-many or many-to-many format to a one-to-one. However, if the actors are stuck with each other regardless of what choices they make, it is difficult to see how the Williamson model could be applicable.
- At least one of the parties has to have a significant investment in a specific asset.

GHM

- Many of the GHM models turn on uncontractible investments. The key point is not to see whether such investments are present at all, since the definition is wide enough to include investments in effort and innovation, but to see whether their uncontractibility was understood at the time when the governance decisions were made.

Alignment mechanisms

- As noted, all NIE models need alignment mechanisms to work. The ones distinguished in the previous section are evolution, competition and design. If the new institutional economics is to be credible, these forces need to be present.

3. History of the Railways

Now that our examination of the theory of the New Institutional economics is complete, it is time to turn our attention to the history of the railways. As explained in section 1.1, above, the choice of the railways as our topic of study is to some extent arbitrary; it could have just as easily been any other semi-public network industry, i.e. an industry characterised by network economies of scale that can be and has been governed with varying degrees of direct government involvement. The benefit of choosing the railway industry is that it has a long history, longer than most other network industries, and that it has, during that time, been the subject of various reforms, both in the direction of more government involvement and in the direction of less government involvement.

This last point is particularly important to remember: we are looking at reforms, in order to see whether the reform is in the direction predicted by our theoretical framework. At no point will we discuss whether the level and type of regulation that existed at a given moment was optimal; we will only talk in terms of reforms that are in line with the theory, and reforms that are not. As we will see, the history of the railways involved four notable reforms, that can be identified in each of the five countries we will look at. In Great-Britain, France, Belgium, Germany and the Netherlands, the railways were first introduced, then consolidated, followed by nationalisation and, most recently, a partial or total liberalisation and privatisation. The question is whether the theory analysed in the previous chapter might not have led us to expect a different pattern of reforms.

A second – and even more basic – way to test the NIE against the facts is to look for the elements listed in section 2.7. This will do nothing to confirm any part of any theory, but the more difficult it is to recognise those elements in the historical narrative, the poorer the “fit” between the theory and the facts that it is supposed to explain. That is why the description of each of the four phases will be concluded by going over the elements of section 2.7, to see whether the necessary conditions for the New Institutional Economics to have any explanatory power have been fulfilled. In section 3.5, we will examine whether NIE survived this most basic of challenges. The question of the relative explanatory power of NIE cannot be examined yet in this chapter. That will have to wait until we have considered the alternatives, which we will do in chapter 4.

3.1 Beginnings

As noted in section 1.3, above, it is important to be transparent about the construction of case studies, so as to avoid the accusation that the result is obtained through rhetorical trickery instead of scientific evidence. The choice of where to begin the telling of the tale of the railways immediately brings this to the forefront. It seems natural to simply trace the technology back to its beginnings, which is what most authors do.⁴⁵⁴ They start by pointing out that wooden rails were already used in the late middle ages to guide wagons in the mining industry. In the 17th and 18th centuries, this practice was slowly improved by placing the wagon on top of the rails instead of between them, and by making the rails out of iron instead of wood. However, they continued to be pulled by horses.⁴⁵⁵ After James Watt invented the steam engine in 1765⁴⁵⁶, essentially all the technical elements were in place, ready for George Stephenson to put the pieces together.⁴⁵⁷ By taking such a technology-based perspective, the author will have no difficulty making his reader understand the revolutionary nature of the introduction of the railways in the 1830s. Naturally, the reader will expect the historical actors involved to approach this new industry as something completely novel, writing the governance structures on a blank slate. In this, the reader would be disappointed.⁴⁵⁸

However, since it is governance structures that we are currently interested in, there is no particular reason why we should look for the beginnings of the railway industry by following the technology back to its roots. We can just as easily take a functional perspective and describe the turnpikes and canals of the 18th century as the ancestors of the railways instead.⁴⁵⁹ This perspective is no more or less valid than the previous one, which is why most historians tend to combine the two.⁴⁶⁰ This second approach would begin by noting that the Romans used public tendering to construct their famous *viae*, meaning that the construction of

⁴⁵⁴ Eg. Holterman (2009) and Wolmar (2007).

⁴⁵⁵ Wolmar (2007), p. 1 describes how Louis XIV had what can only be described as the world's first roller coaster. This construction, known as the Ramasse, was built in 1691 at Marly-le-Roi. It still had wooden rails and relied on human rather than horses for power. Cf. Mayor (1998).

⁴⁵⁶ Cf. James Watt's Fire Engines Patent Act of 1775, 15 Geo 3, c. 61, the Act of Parliament that extended the original patent, which Watt obtained in 1769. 1765 is the year he completed his first working model. Cf. also Marsden (2002).

⁴⁵⁷ Cf. Hylton (2007), chapter 1. While Stephenson's name is most commonly associated with the invention of the steam locomotive, the first working locomotive was actually built by a man called Richard Trevithick in 1804.

⁴⁵⁸ Cf. Holterman (2009).

⁴⁵⁹ We might even take a broader perspective still, and look at network industries in general. After all, they have many governance problems in common. Cf. Künneke, Groenewegen & Auger (2009).

⁴⁶⁰ Eg. Caron (1997).

those roads was carried out by private parties under close supervision of the *censores*, based on careful standards set out in detail in the *Leges Censoriae*, followed by a public bidding procedure.⁴⁶¹ These roads were then maintained by the local authorities using a similar procedure.

Despite the obvious academic value of this first point of observation, it is given here only for context. Not only is there no discernable continuity from the Roman system to the present day, there is also a debilitating lack of information about their actual practice, as opposed to the theory given above. One might consider, for example, that it is difficult to imagine that the *censores* would have been able to supervise effectively except for those public works that took place in the vicinity of Rome. Nevertheless, a number of aspects are worth bearing in mind, since we will encounter them again in later centuries:

1) Roman *viae* were constructed primarily for military purpose. The first one, the *Via Appia*, was built in order to facilitate the war against the Samnites, who lived in the north of the present day Italian region of Campania⁴⁶², and later roads followed this pattern. This is an important factor to consider in judging not only the efficiency of the route chosen, but also the efficiency of the overall governance structure. Compared with the ordinary civilian economy, military decision making is much more focused on making sure that the system works when it really matters. Often times, seemingly significant inefficiency is warranted by the assurance that there will be no failure at the critical moment.

2) Given what we know about the Roman republic, it is reasonable to hypothesise that the choice for private contractors was motivated less by concern about efficiency, and more by an inability or unwillingness to develop the institutional capacity necessary for the republic to perform such functions itself. Only during the period of the Roman Empire did the central government develop a civil service that exceeded the bare minimum.

3) It is important not to overlook the difference between theory and practice. To quote Yogi Berra: “In theory, theory and practice are the same. In practice, they are not...” Both the institutional design as it is laid down in positive law and the actual practice of institutions are the subject of the present thesis.

⁴⁶¹ Cf. Gibbon (1776), p. 58-59, Van Tilburg (2005) and Ramsay (1875). The wikipedia page http://en.wikipedia.org/wiki/Roman_censor (last visited on 9 May 2009) cites a number of classical Roman sources that discuss this topic, including works by Cicero and Livy.

⁴⁶² Cf. Livy, *Ab urbe condita*, book IX.

The Roman *viae* continued to be the only means of long distance travel on land throughout the Middle Ages and the Renaissance. Obviously, the institutions created by the Romans to look after the maintenance of these roads died with the Empire itself, so they generally fell into disrepair. Only in the 17th century was there any widespread effort to construct new infrastructure. In England, local acts in 1663 and 1707 pioneered the use of so-called turnpike trusts, which constructed toll roads.⁴⁶³ Similarly, the construction of canals, a practice that evolved out of the improvements carried out on various natural waterways starting in the Tudor era⁴⁶⁴, was also initially entrusted to public officials.⁴⁶⁵

Other European countries followed suit, most famously with the construction of the Canal du Midi between 1666 and 1681, connecting the Atlantic and the Mediterranean. The governance structure created for that canal was a bit unusual, due to the influence of the original backer of the plan, Pierre-Paul Riquet. His heirs continued to own and operate the canal until 1897, and to this day the legal regime that covers various aspects of that canal is different on certain points from the rest of France's inland waterways.⁴⁶⁶ Compared with England, France shares a long history of central government, but has always had more pressing military worries, which explains why the French did not leave the development of key infrastructure to local initiatives, but entrusted it to the *Corps des ingénieurs des ponts et chaussées*, the corps of engineers of bridges and roads, instead.

As for the other countries that are the main focus of this chapter, the Netherlands had little need for either long distance roads or canals in the pre-Napoleonic era, while Belgium and the different German states were either too impoverished or too internally divided to undertake any large scale construction works.

⁴⁶³ The 1663 act, 15 Car. II, c. 1, also numbered as 15 Car. II, c. 14, assigned the administration of the trust to local justices, while the 1707 Act, 6 Anne, c. 4, for the first time appointed private citizens as trustees. Cf. Albert (2007), p. 22. N.B. the notion of a “local act” will be explored at length below, in the section on Company Law.

⁴⁶⁴ The city of Canterbury obtained Parliamentary approval for the “canalisation” of the River Great Stour in 1515.

⁴⁶⁵ The oldest one is the Exeter Canal, constructed by the city of Exeter in 1564-1566. Cf. Bond (2007), p. 153.

⁴⁶⁶ Cf. Code du domaine fluvial et de la navigation intérieure, art. 224 and the Code général de la propriété des personnes publiques, art. 2124-20 – 2124-25.

3.1.1 Company Law

Before discussing Europe's first railway companies, it is necessary to consider the problem of company law, or the absence thereof, in the countries under consideration. As we will see, the absence of a convenient system of company law is extremely disadvantageous to the development of industry, the more so for enterprises that require an extremely large capital.⁴⁶⁷ Since the railway companies that emerged in the second half of the 19th century were easily the largest organisations ever created, their initial development was severely hampered by the complicated procedures that existed for the creation of a joint stock limited liability company, meaning a company who is legally treated as a person distinct from the human beings running it, which is capable of acquiring capital through the emission of shares, and whose shareholders are limited in their liability to the amount of their share.⁴⁶⁸

The first such company is generally considered to be the Dutch East India Company, which was established on March 20, 1602. Our interest here is not in the organisation of that company, but in the procedure by which it was created. It was preceded by a number of smaller companies, which are better described in contemporary legal language as partnerships, each formed by a number of merchants so as to pool their resources and reduce the risk.⁴⁶⁹ In 1602, however, the Dutch Estates-General enacted an "Octrooi", which established the existence of the company and the rules by which it would be governed.⁴⁷⁰ In the preamble, the Estates-General state that they are taking this action "uit souveraine magt ende autoriteit", i.e. from sovereign power and authority, meaning that considered that they derived the legal power to do so from custom and natural law, in a manner similar to the sovereign rulers who had come before. The other European sovereigns quickly followed suit, forming their own East India Companies.⁴⁷¹

⁴⁶⁷ Cf. Commons (1924), chapter IX.

⁴⁶⁸ Cf. Commons (1924), chapter V and Forde (1999), par. 1-03 – 1-08. Throughout this chapter, a corporation as such is defined only as having legal personality, and incorporation means the creation of a distinct legal person. Whether the other elements are present as well, depends on the applicable legislation.

⁴⁶⁹ There was some government involvement in these companies as well, since the state (either Holland or Zealand or the Estates-General of the United Netherlands) usually provided guns for the ships as well as letters of marque and reprisal. Cf. Den Tex (1962), p. 385 and articles XXV and XXXVII of the *Octrooi* of the Dutch East India Company of 1602.

⁴⁷⁰ This document is available on-line at http://nl.wikisource.org/wiki/Octrooi_van_de_VOC (last accessed on May 13, 2009). The political history of this event is recounted in Den Tex (1962), p. 384-420.

⁴⁷¹ The English East India Company already received a Royal Charter in 1600, meaning that it had legal personality, but at that time it did not yet have the ability to raise further capital by issuing shares. For an overview of the different European trading companies, cf. Gepken-Jager, Van Solinge & Timmerman (2005), Part I.

The different aspects of a joint-stock limited liability company require different theoretical rationales, which are generally outside the scope of this thesis.⁴⁷² It is interesting to point out, though, that once this legal innovation had occurred, it quickly spread due to competitive pressures. Once the route to the East Indies, which had originally been known only to the Portuguese, had been discovered by the other seafaring nations, competition started to work in the way imagined in section 2.6.2, above, driving both the underlying technology and the governance structures used to ever greater efficiency.⁴⁷³

The next innovation in company law is of more immediate interest here, since it is closely connected to the early history of the railway industry. As long as the various trading companies were the largest companies in existence, the original legal framework for the establishment of a company was more than satisfactory. Generally, such companies were established in order to give them the monopoly of trade with a certain region, which was by nature a rare occurrence.⁴⁷⁴ Because the governments viewed such juridical persons with suspicion, they tended to limit the enactments of company charters as much as possible, subjecting proposals to close scrutiny.⁴⁷⁵ That was already a problem for the turnpikes and canals of the 18th century, but only after start of the industrial revolution generally and the invention of rail transport in particular did this inefficiency reach a sufficient level for change to become inevitable.⁴⁷⁶ As Alfred Chandler put it, the railways were “the nation’s first big business”.⁴⁷⁷

⁴⁷² For limited liability for example, cf. Carney (1999).

⁴⁷³ To some extent, one could also argue for the existence of evolution here, since the greater efficiency achieved by the English and the Dutch allowed them to all but bring an end to any Portuguese activities in the region.

⁴⁷⁴ For example, the Hudson’s Bay Company, which still exists today, was founded by Royal Charter in 1670, and given the monopoly of trade with the Hudson Bay area. Even in Prussia, which had no colonies or trading empire prior to the advent of the railways, the largest corporations were the Rheinisch-Westindische Kompanie and the Deutsch-Amerikanischer Bergwerksverein, companies charged with the trade between Prussia and the Americas.

The legal profession continued to struggle with the exact nature of legal personality well into the 19th century, cf. The American case of *The Bank of the United States v Deveaux*, Supreme Court of the United States, February Term 1809, 5 Cranch 61, 9 US 61, 3 L.Ed. 38, which ruled that only the citizenship of the “members” of the corporation counted for jurisdiction purposes. This ruling was only overruled by *Louisville, Cincinnati & Charleston Railroad Company v Letson*, Supreme Court of the United States, January Term 1844, 43 US 497, 2 How. 497, 11 L.Ed. 353, and *Marshall v Baltimore and Ohio Railroad Company*, Supreme Court of the United States, December Term 1853, 57 US 314, 16 How. 314, 14 L.Ed. 953, not coincidentally two railroad cases.

⁴⁷⁵ Not to mention scrutiny subsequent to incorporation. As late as 1819 the Supreme Court of the United States affirmed the general power of “visitation” enjoyed by the sovereign, meaning “the right to oversee corporate affairs, quite separate from the power to enforce the law.” (*Cuomo v Clearing House Association*, Supreme Court of the United States, June 29, 2009, 557 US ___, on p. 4.) The sovereign was, and quite likely still is, entitled to exercise its right of visitation similarly to the church’s right to supervise its institutions. (*ibid*, p. 3, citing Blackstone (1765), p. 469.) The 1819 case is *Trustees of Dartmouth College v Woodward*, 4 Wheat. 518.

⁴⁷⁶ The first modern joint-stock company in Germany, for example, was the steel producer Dillinger Hütte, founded in 1685 and incorporated under the new regime of the Code de Commerce in 1809.

⁴⁷⁷ Chandler (1965).

In 1807, imperial France codified previous practice and gave the *Société Anonyme* a place in the new *Code de Commerce*, which also became law in Belgium, the Netherlands and parts of Germany.⁴⁷⁸ Throughout the 19th century, countries continued to make access to corporations easier, as well as to standardise their nature. For historic reasons, the UK had the furthest to go in this respect, given the restrictive regime set up by the Bubble Act of 1720.⁴⁷⁹ At the same time, the UK also had the most advanced industry, making the need for more convenient company law all the more pressing. After the Bubble Act was finally repealed in 1825⁴⁸⁰, a series of acts starting with the Joint Stock Companies Act of 1844 quickly established the legal regime that still exists today.⁴⁸¹

In the continent pre-1807, a number of legal institutions existed that dated back to Roman law, whose exact status at law was often uncertain, but which served to make the introduction of a comprehensive body of company law less urgent.⁴⁸² Title II-6 of the Prussian ALR, the *Allgemeines Landrecht* or General state laws, which was promulgated in 1794, for example, introduces a set of rules for *Gesellschaften* that are based on the law of partnerships as it goes back to Roman law⁴⁸³, but with several elements mixed in that we would consider more typical of corporations.⁴⁸⁴ The deficiencies of this body of law were clearly exposed by the early history of the *Köln-Mindener Eisenbahn*, the railroad meant to connect the Rhine and Weser rivers.⁴⁸⁵

The need for such a railroad arose out of a particularly egregious case of creative legal interpretation advanced by the Dutch regarding the Final Act of the Congress of Vienna. Annex 16 to that Treaty laid down certain Regulations for the free navigation of rivers, and par-

⁴⁷⁸ Cf. Fyffe (1895), chapters V, XI.

⁴⁷⁹ An Act for better securing certain Powers and Privileges, intended to be granted by His Majesty by Two Charters, for Assurance of Ships and Merchandize at Sea, and for lending Money upon Bottomry; and for restraining several extravagant and unwarrantable Practices therein mentioned, 6 Geo I, c. 18. Cf. Harris (1994) for a discussion of the nature of the act and its connection to the 1720 South Sea Bubble and Schmitthoff & Thompson (1968), p. 6-7 for a discussion of the deed of settlement company, a type of partnership agreement used to overcome the Bubble Act.

⁴⁸⁰ Bubble Companies, etc. Act of 1825, 6 Geo. IV, c. 91. Cf. Schmitthoff & Thompson (1968), p. 5 and Freeman, Pearson & Taylor (2007).

⁴⁸¹ The main acts were the Joint Stock Companies Act of 1844, 7&8 Vict., c. 110, the Limited Liability Act of 1855, 18&19 Vict., c. 133 and the Joint Stock Companies Act of 1856, 19&20 Vict., c. 47. Cf. Schmitthoff & Thompson (1968), p. 7-9.

⁴⁸² Cf. Raaijmakers (2000), par. 1.16 and 1.30-1.42.

⁴⁸³ Cf. for instance § 13, which states that *Gesellschaften* do not have legal personality.

⁴⁸⁴ Explicitly in § 14, for example, which declares the law of corporations, i.e. corporations established by Charter or Statute, applicable to *Gesellschaften* as well, insofar as it concerns the internal governance of the organisation.

⁴⁸⁵ Cf. Holterman (2009). Most of the facts of both that article and the present discussion of early German railway history are taken from Steitz (1974).

ticularly for the Rhine. Chief among these was the general principle of freedom of the river: “The navigation of the rivers, along their whole course, referred to in the preceding Article, from the point where each of them becomes navigable, to its mouth, shall be entirely free, and shall not, in respect to commerce, be prohibited to any one”.⁴⁸⁶ The difficulty here was that the Dutch and the Prussians could not agree on the meaning of the phrase “jusqu’à la mer”, with the Dutch insisting that this did not forbid their practice of regulating the traffic between the river and the sea. As a result, they continued to stop certain goods from being transported directly to Germany until 1831, when a compromise was reached that discontinued the blockade in return for tolls that were allowed to discriminate against the Germans.⁴⁸⁷

Starting in 1828, a number of leading entrepreneurs in the Prussian Western Provinces⁴⁸⁸ started working to resolve this problem permanently by building a rail connection between Köln on the Rhine, via the Ruhr area where most of the mining and other industry was concentrated, to Minden on the Weser, from where transport could take place by ship through Hanoverian territory to Bremen and the seven seas.⁴⁸⁹

⁴⁸⁶ Art. 109 of the Final Act.

⁴⁸⁷ Cf. Clapp (1911), p. 13-14.

⁴⁸⁸ The Prussian Provinces of Westfalen and Rheinpreussen had been annexed to Prussia by the Final Act of the Congress of Vienna (art. 23-25), the latter having belonged to France since the peace of Basel of 1795, while the former had been created by the French as an independent state following the Treaties of Tilsit of 1807. This difference in origin is important because it means that French law, including the Code Civil and the Code du Commerce, continued to apply in Rheinpreussen, but not in Westfalen, which had only the traditional German legal system and the ALR. Cf. Dunlavy (1994), p. 22. Until 1866, the Western Provinces were not connected to the rest of Prussia.

⁴⁸⁹ The Kingdom of Hannover was an independent state in personal union with Great Britain between 1815 and 1837 and a completely independent state from 1837 until its annexation by Prussia following the Austro-Prussian War of 1866. The city of Bremen was an independent republic until it became part of the North German Confederation in 1867. Cf. Holterman (2009), p. 108 for a map showing the different sovereign states in the region before 1848.



Figure 3: The trajectory of the Köln-Minden railway.⁴⁹⁰

The total length of this trajectory is about 250 km, or about five times as long as the two pioneering railways in England that had only just started operation at the time, the 40 km long Stockton & Darlington, which was opened in 1825 and the 56 km long Liverpool & Manchester, which started running in 1830.⁴⁹¹ Even a more modest connection between the heart of the Ruhr area and Minden would still have meant 150 kilometres of rail. While it was certainly not beyond the contemporary capacity to complete such a project, it was clear from the start that the costs would exceed anything previously attempted in Germany.

In 1831, the German railway pioneer Friedrich Harkort defended the idea in the *Landtag* of Westphalia. In that debate, he suggested a total budget of 660.000 Thaler.⁴⁹² When the time came to do more detailed calculations, that amount proved much too optimistic. In 1832, the estimate was 2,5 million Thaler⁴⁹³, an amount that rose to 3 million in 1834 and 3,8 mil-

⁴⁹⁰ A more detailed map is given in Steitz (1974), p. 283.

⁴⁹¹ Cf. section 3.1.2 and 3.1.3, below.

⁴⁹² Steitz (1974), p. 127.

⁴⁹³ Steitz (1974), p. 133.

lion in 1835.⁴⁹⁴ This last amount was the sum given in the prospectus of the Rhein-Weser-Committee, which undertook to secure permission and incorporation from the Prussian government in Berlin. Throughout this period, the backers of the plan, supported by the Westphalian government, kept trying to get the government in Berlin to get involved in the project, or at least to provide some of the funds, but the government was unwilling to do either.

Starting in the spring of 1835, the various railway committees set up in the cities that were to be connected by the railway started raising money for the company.⁴⁹⁵ People willing to invest would sign up for shares worth 100 Thaler each, for which they were asked to pay only 5 Thaler upon signing. The rest of the sum could then be called up by the general meeting of shareholders as the company would need the funds.⁴⁹⁶ Incorporation was finally achieved in January 1836, when the King of Prussia signed the Charter, leaving only the details of the railway concession to be worked out.⁴⁹⁷ Some of those details, however, turned out to be difficult, with particularly the role of the mails as a problem.⁴⁹⁸ Finally, the concession was approved on August 21, 1837.⁴⁹⁹

Unfortunately, it was all downhill from there. With cost estimates steadily rising, and revenue projections steadily falling, investors started to get uneasy. Already soon after the end of the negotiations in Berlin did the board of the company encounter difficulties in getting the shareholders to furnish the additional sums they owed. In October 1837, the it requested that the first 10% should fall due immediately, with additional tranches of 5% in November and 10 % in January 1838, which would combine with the sum already paid upon signing to add up to a total of 30%.⁵⁰⁰ In June 1838, the company's third General Meeting was convened, allowing the shareholders for the first time to discuss the agreement made with the government and the planned order of construction of the different sections of the railway.⁵⁰¹ To say the least, the shareholders were not happy, and started to consider whether it might not be

⁴⁹⁴ Steitz (1974), p. 142.

⁴⁹⁵ Cf. Na (2002) and Zeumer (2003), who discuss the Rheinland capitalist society in the first half of the 19th century.

⁴⁹⁶ Steitz (1974), p. 156-157.

⁴⁹⁷ Steitz (1974), p. 160. The Charter or statutes of the company are reproduced in Steitz (1974), p. 291-298.

⁴⁹⁸ Cf. Holterman (2009), p. 113.

⁴⁹⁹ Steitz (1974), p. 181.

⁵⁰⁰ Steitz (1974), p. 185. Note that the board ("Verwaltungsrat") did not have the authority to call up these sums itself, with the exception of the first 10% which was due upon enactment of the concession under § 11 of the company statutes. Cf. Steitz (1974), p. 186. For obvious reasons, the current German *Aktiengesetz* states that the board ("Vorstand") is entitled to claim such funds as it pleases, cf. § 63 (1) *Aktiengesetz*.

⁵⁰¹ Steitz (1974), p. 191. Before the meeting had even begun, it had already made the problems worse. The board had refused entry to the meeting to all shareholders who had not paid up the funds that had been called, as they were entitled to do under the statutes. Those that had not done so, in their turn, refused to pay if they did not have a vote on how the funds were to be used. Steitz (1974), p. 193.

better to dissolve the company. In subsequent months, the solidity of the company became increasingly threatened. Questions were raised about the board, which had continued to work after incorporation without allowing the shareholders of the new company to vote on their position, and about whether the board had carried out its duty to guard the financial reliability of the shareholders' registry. By the end of 1838, the board started to consider a partial or entire dissolution, which, of course, would require approval from Berlin. The fourth General Meeting, in February 1839, led to nothing amid uncertainty about whether or not sufficiently many shareholders were present to allow the meeting to decide to dissolve.⁵⁰² Finally, at the General Meeting in May 1839, it was decided to liquidate the company.⁵⁰³

In response to these events, the Prussian government decided to turn the General Conditions of their agreement with the company⁵⁰⁴ into the Railway Act of 1838 (“*Eisenbahngesetz*”), taking into account their subsequent experience.⁵⁰⁵ A few years later, they enacted the country's first comprehensive statute governing incorporation, the *Aktiengesetz* of 1843.

The story of the demise of the *Rhein-Weser Eisenbahngesellschaft* neatly illustrates, in a nutshell, a number of the topics that were discussed in the previous chapter. Section 2.1 argued that the only really important distinction was between laws of general application and other law, a fact that is illustrated by the observation that, while the legal instruments used to codify the legal relationship between the state and the company were neither fully private nor fully public law, they were all negotiated and never of general application. The same section also considered the nature of incompleteness, which turns out to be problematic. While the contracts at issue here were undoubtedly highly incomplete, the bigger problem turned out to be that they were based on incorrect assumptions – most importantly the assumption as to the likely costs of construction.

⁵⁰² Those shareholders who had not paid up were again refused entry, but they were counted for the purposes of deciding how many shareholders would constitute a quorum. This, of course, at the behest of the board, who wanted to avoid total dissolution. Steitz (1974), p. 195.

⁵⁰³ Steitz (1974), p. 195. After the collapse of the Rhein-Weser Company, the process started over with a new company and a new concession, applying a new legal regime set up by the statutes of 1838 and 1843, with greater government involvement. Construction started in August 1844, and the first train departed from Minden on October 15, 1847. Cf. below.

⁵⁰⁴ The *Allgemeine Bedingungen* were agreed on July 24, 1836, and were intended to apply to all railway concessions. They are reproduced in Steitz (1974), p. 288-290.

⁵⁰⁵ *Gesetz über die Eisenbahn-Unternehmungen* of November 3, 1838. The statute is reproduced in Steitz (1974), p. 309-320. Cf. also Dunlavy (1994), p. 70-71, Lehmann (1998), p. 17-18 and Ziegler (1996), chapter 2.2.

Another thing that becomes clear from this case is the importance of very ordinary transaction costs. The absence of a standardised system for incorporation requires actors to spend vast amounts of money petitioning parliament or the king, negotiating over the terms of the company statutes, and resolving legal disputes caused by the unsettled nature of the rules of the game. All of these are transaction costs of the type discussed by Commons and Coase⁵⁰⁶, which can be understood without any need for advanced modelling. At the same time, the government's concerns, that made it reluctant to liberalise access to incorporation, can be examined further from a generalised NIE perspective.

The problem was that the government found itself unable to write the rules of the game with sufficient completeness to stop private actors from abusing the institution of a limited liability company. With some creativity, this situation could be modelled as a vertical integration problem, the paradigmatic subject of both Williamson's TCE and the GHM models, but that is where the fit runs out. For TCE to apply, there needs to be an identifiable and significant investment in specific assets⁵⁰⁷, which is missing here. The same goes for GHM, which works best in a situation where there is an investment in specific human capital.⁵⁰⁸ It is clear, though, that the specific regime set up in the statutes of the Rhein-Weser company did not survive the evolutionary pressures of a tough competitive environment, where speculators had their choice of investments. It is equally clear that a mixture of competitive pressure from other nations and a general paternalistic desire to do well forced the Prussian government to abandon its cautious high-transaction cost approach to incorporation. Once the industrial revolution had forced other European nations to go down that path, the Prussians could not avoid following suit.

It is interesting to consider the order in which this reform was adopted by the various countries. The French were first, because they decided – for exogenous reasons – to start over from scratch. Writing on a blank slate, they considered the optimal legal rules and brought them together in a series of legal codes. Subsequently, however, this innovation spread roughly as follows:

- a) First to large counties. This makes sense because the cost of considering a legal reform is independent from the size of the country, while the potential benefit increases with size.

⁵⁰⁶ Cf. section 2.3, above.

⁵⁰⁷ Cf. section 2.4, above.

⁵⁰⁸ Cf. section 2.5, above.

- b) First to countries that were more industrialised, and therefore in greater need of this reform. In countries whose economies relied more on agriculture or trade the impetus to innovate was smaller.
- c) First to countries that lacked a suitable alternative in the private law of partnerships. Particularly in the UK, the existence of the institution of a deed of settlement company reduced the pressure to innovate.⁵⁰⁹

This is what an economic explanation of institutional change looks like. Changes in the realm of law are explained by pointing to changes in the real economy that have the effect of reducing the efficiency of the old law relative to possible alternatives. The difference between theories lies mostly in which costs and benefits are included in the notion of efficiency. Public choice theory, for example, focuses more on the cost and benefits of decision makers, rather than society in general.⁵¹⁰

Having discussed some of the history of company law in Western Europe, it is now time to consider the history of the world's first real⁵¹¹ railway: The Stockton & Darlington.

3.1.2 Stockton & Darlington

The world's first modern railroad was – of course – the result of private initiative. As was already noted above, Britain, like most countries in the world at the time, had no central planning ministry for infrastructure projects.⁵¹² Instead, the initiative was left to private citizens, or at least to local government. In the case of the Stockton & Darlington railway, the initiative came from the town of Darlington, faced with the prospect of being bypassed by a canal that was proposed across County Durham to the river Teese near Stockton. The leading promoter was a wool merchant named Edward Pease.⁵¹³

⁵⁰⁹ Cf. Schmitthoff & Thompson (1968), p. 6-7 for a discussion of the deed of settlement company, a type of partnership agreement used to overcome the Bubble Act.

⁵¹⁰ Cf. section 4.1, below.

⁵¹¹ In the sense of offering a service to the public over a significant distance using at least some steam traction. Cf. Hylton (2007), p. 18-20 and Wolmar (2007), p. 1-9 for a discussion of some of the earlier enterprises in the field.

⁵¹² The one exception to this rule was France, mostly due to greater military necessity. Cf. Caron (1997), p. 12-23.

⁵¹³ For a more detailed account of the history of the Stockton & Darlington, cf. Gordon (1910), p. 81-93, Hylton (2007), p. 24-30, Tomlinson (1915), ch. III and IV, and Wolmar (2007), p. 9-20.

Because of the technological uncertainties of the project, the proposals made to Parliament consistently avoided any commitment to a particular means of traction. The first Bill, which was introduced in 1819, did initially contain a reference to something called a “locomotive”, but the word got struck in the drafting process as being nonsensical.⁵¹⁴ When this Bill was narrowly voted down, a similar proposal was introduced again in 1821. In April 1821 the Bill was passed.⁵¹⁵

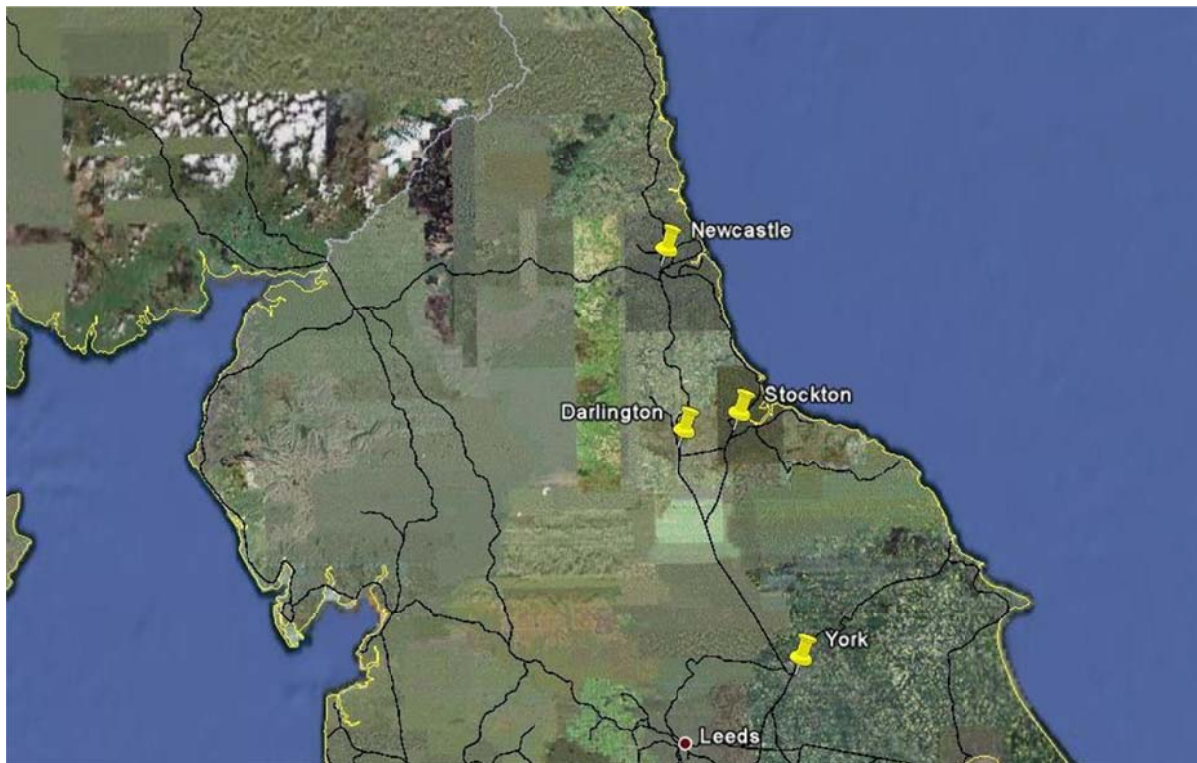


Figure 4: *The Stockton & Darlington*

Only at that time was Pease contacted by railway pioneer George Stephenson, who convinced him of the benefits of steam traction. As a result, Stephenson was hired as the project’s chief engineer. The changes he made to the route, as well as the explicit possibility of using steam, were ratified by Parliament in 1823.⁵¹⁶ The result was a 26 miles (40 km) long railway connecting the coal district around Darlington and the river Teese at Stockton, officially open for business as of September 27th, 1825.

⁵¹⁴ Hylton (2007), p. 26.

⁵¹⁵ Stockton & Darlington Railway Act (1821), 1&2 Geo. 4, c. xliv.

⁵¹⁶ The key section of the Act is quoted in full in Hylton (2007), p. 27. The Act is the Stockton & Darlington Railway Act (1823), 4 Geo. 4, c. xxxiii, which was, in turn, later amended or replaced by the Stockton & Darlington Railway Act (1824), 5 Geo. 4, c. xlvi the Stockton & Darlington Railway (Consolidation of Acts, Increase of Capital and Purchase of Middlesbrough Dock) Act (1849), 12&13 Vict., c. liv, and a number of other acts thereafter, until the company was taken over by the North Eastern Railway in 1863.

The most important thing to remember about the Stockton & Darlington is that it was never really conceived as a transport company. The promoters of the S&DR thought they were creating an infrastructure company, creating the means by which others would transport coal, people and other goods from the inland to the shore, much the same way the promoters of turnpikes and canals had already done for more than a century.⁵¹⁷ To be sure, the company took full advantage of its close relationship with father and son Stephenson, who in 1823 had set up a locomotive building company in Newcastle, together with Edward Pease and steel magnate Michael Longridge.⁵¹⁸ After all, the reason why they built the railway is that they thought that using it to transport coal to Stockton would be a lucrative proposition, one in which they were eager to share. But the company never intended to have the monopoly of use of the railway. As a result, the early life of the company was characterised by a number of unusual rules and practices.

First of all, until 1833 traction was by horse power as well as steam power. In and of itself, this was less problematic than one might think, since the difference in speed between them was not yet so great as to cause great difficulty.⁵¹⁹ There was no timetable, or, in fact, any central management of any kind. Whoever wanted to use the railway simply turned up when it suited them, with their own wagons and means of traction. Predictably, this led to daily arguments about right of way, which were only exacerbated by the limited number of passing loops.⁵²⁰ Where the railway crossed a regular road, the road had right of way, a fact enforced by the existence of fences across the rails. The train driver would have to stop the train, open the fence, cross the road, stop again and close the fences, before continuing.⁵²¹ Only in the late 1830s and 1840s did Parliament give the railway companies the authority to regulate the manner in which outsiders would use their track⁵²², although by that point open access had already been effectively abolished.

⁵¹⁷ Earlier horse-powered railways, going back to the Surrey Iron Railway which was opened in 1803, had also worked on this basis, at least those railways that were not specifically tied to a specific mine. Cf. Hylton (2007), p. 18-20 and Wolmar (2007), p. 7.

⁵¹⁸ Longridge was the owner of the Bedlington Ironworks in Northumberland, where in 1820 the first modern rails had been invented.

⁵¹⁹ During its inaugural run in September 1825, the steam train took two hours to complete the first 12 miles (19 km) of the trip. Cf. Hylton (2007), p. 28.

⁵²⁰ There were four in every mile, although there is some uncertainty about the exact method used. Cf. Wolmar (2007), p. 18. Of course, no amount of planning can avoid the trouble caused by drivers leaving their vehicles on the rails while they go for a drink. Cf. Hylton (2007), p. 73.

⁵²¹ Cf. Hylton (2007), p. 71 and s. 9 of the Railway Regulation Act (1842), 5&6 Vict., c. 55. This rule was only unequivocally abolished in by the Railway Regulation Act (1844), 7&8 Vict., c. 85.

⁵²² Cf. Railway Clauses Consolidation Act (1845), 8&9 Vict., c. 20, s. 1115 and 117-121.

The question of the feasibility of open access is an extremely important one.⁵²³ If open access is feasible, this would greatly enhance the benefit that can be achieved by forbidding the owner of the railway from operating trains on their track, a policy that is in place at the moment in many European countries. If open access is not possible, this approach only substitutes two smaller monopolists for one bigger one.

Again carrying out the analysis one step at a time, we see that the contract providing for each carrier's access to the road was incomplete as a result of having too few parties. A real Coasian bargain would involve a contract specifying the obligations of every carrier as well as the railway company, negotiated freely by all. Instead, the company used a standardised and rather short contract between it and every individual carrier, which left many problems unsolved, especially those concerning the relationship between carriers. Of course, the carriers could have negotiated additional contracts between each other, but that would have been prohibitively expensive. The same goes, incidentally, for the theoretically ideal Coasian bargain.

A second best solution in such circumstances is to have the railway company act as an agent for all other carriers, and in that capacity set the rules for use of its railway. That is what was done by the turnpikes and the canals, and it is still done by the government in enacting the Highway Code. After all, in a Coasian world with no transaction costs, it would be inefficient to have one standard rule dictating who should give way when two vehicles meet. Instead, the parties would simply bargain so that the driver who desires the right of way the most, for instance because he is in a hurry, compensates the other for giving way. The road owner only steps in because the alternative would be prohibitively costly.⁵²⁴

This raises a question of agency theory: Can it be assumed that the railway company will act as a good agent for the other carriers? Considering the analogy with the turnpikes and canals⁵²⁵, it appears the key difference is the directness with which the actions of one carrier affect the others. Any delay immediately causes all others to be late as well.⁵²⁶ The solution

⁵²³ Which is why it is important not to forget to return to it when evaluating policy alternatives today. Cf. Becker & Hoekstra (1999), Annex I.

⁵²⁴ An even more fascinating example is the way in which the US Department of Commerce negotiated with ICANN in the late 1990s to set up a system of governance for the internet. In so doing, it acted as an agent for the world.

⁵²⁵ N.B. Until 1845, British canal companies were forbidden from acting as carriers on their own canals absent explicit permission from Parliament. This rule was abolished by section 1 of the Canal Carriers Act (1845), 8&9 Vict., c. 42.

⁵²⁶ Cf. section 2.5.5, above, which discusses the model of Hart, Shleifer & Vishny (1997). This model contains a parameter $b(e)$, which describes the influence of cost-saving investments by the seller on the buyer's benefit

would be a rule against delay, or at least against parking one's train on the main railway line. Unfortunately, enforcing such a rule with a fine is inefficient, because such a fine discourages delays even when they are welfare enhancing, i.e. when they do less damage than the benefit they generate for the carrier.⁵²⁷ Instead, the optimal enforcement mechanism would involve compensation being paid by the carrier causing the delay to those who suffer from it. However, determining the size of the compensation requires a lot of information, making this an unattractive option from a transaction cost point of view.

In short, because the carriers are heterogenous in their ability to reduce the risk of delay and the extent to which they suffer as a result of delays caused by others, and because a delay by one carrier affects the others so much more directly than in the case of a turnpike, the railway company is placed in a position similar to the central planner discussed in Hayek (1945): dependent for its planning function on information possessed by others, which they have every incentive to lie about, and which may be of such a nature that it cannot effectively be communicated at all, the central planner inevitably acts highly inefficiently. In Hayek's case, the recommendation is to decentralise decision making power, which, applied to the railroad, would involve setting up an efficient arbitrage mechanism where the carriers can file their claims against each other, was the only alternative.⁵²⁸ For the railways, however, it is unlikely that this would bring a great reduction in transaction costs. Instead, the railway company did something that an entire society cannot do: it integrated vertically. Whether this necessarily involves integration between the owner of the rails and the carrier is a question we will consider later in this chapter, but at the very least our analysis of the experience of the Stockton & Darlington seems to indicate that open access is not a feasible structure for a railway, no matter how appealing this would seem when viewed from the point of neo-classical micro-economics.

from the product or service. Under a perfect Coasian bargain, all carriers would be each other's buyer and seller, making this problem of delay a good example of how this part of the HSV model should be understood. After all, cost cutting by a rail carrier could very well mean a greater risk of delay, even if the delay is not intentional. These delays, in turn, would reduce the benefit of all other carriers. The analysis in the main text here is essentially a plain english rendering of what the HSV model would have to say about this problem.

⁵²⁷ This is called Efficient Breach Theory, cf. for example the recent paper by Adler (2008). The main proposition of Efficient Breach Theory is that damages rules in contract law should be such as not to discourage breach of contract when such breach produces a net social benefit.

⁵²⁸ Cf. Murray (2001), who discusses how railway companies in modern Britain resolve the question of who owes who how much in damages for any given delay. The structure of the current British railway industry will be explained in section 3.4, below.

3.1.3 *The Liverpool & Manchester*⁵²⁹

In many ways, the Liverpool & Manchester is the first modern railway.⁵³⁰ Gone are the horses, gone is the open access, gone is the predominant focus on the transport of bulk goods, specifically coal, gone is the single track, and gone is the local perspective of the backers.⁵³¹ Instead, the Liverpool & Manchester was conceived from the start as an ambitious way to break the monopoly of the canals – there were three⁵³² – on the transport between the great industrial heart of Lancashire and the port that gives it access to the rest of the world.

As with the Stockton & Darlington, the technical work was initially entrusted to George Stephenson. It was his job was to work out the technical details, as well as to survey the intended route. However, when the bill was brought before parliament in February 1825, he ran into difficulty. It is important to remember that the Great Reform Act was only enacted in 1832⁵³³, meaning that in a still profoundly rural Britain, both Houses of Parliament were even more dominated by the landed gentry and the nobility than the country at large. These landed interests had two related reasons to oppose the railway. Firstly, they tended to have significant financial stakes in the canals, either because they owned shares, or because they received financial compensation in the form of so-called wayleaves, which acted as consideration for the easement the landowner had granted to the canal. No such compensation could be obtained from the railways, whose enabling acts tended to allow compulsory purchase of the land needed.⁵³⁴ This compulsory purchase itself being, of course, the second reason why land owners did not look fondly on the railways.

The most important spokesperson for the other side was the 43 year old self-taught railway pioneer George Stephenson, a man of modest background, who taught himself how to read when he was 18, and who had gotten interested in the railways from a background in the coal and steel industries.⁵³⁵ Always more the visionary than the PR prodigy, Stephenson had the added disadvantage of having to divide his time between the preparation of Liverpool &

⁵²⁹ The classic source on the early history of the Liverpool & Manchester is Carlson (1969).

⁵³⁰ Cf. Wolmar (2007), p. 19-20.

⁵³¹ Another novelty: the use of a timetable. Cf. Hylton (2007), p. 37.

⁵³² The fascinating story of the competition, or lack thereof, between the different canal connections between Liverpool and Manchester is recounted in Hylton (2007), p. 14-16.

⁵³³ An Act to amend the representation of the people in England and Wales, 2&3 Wm. IV, c. 45. One of the effects of this act was to give representation in the House of Commons to Manchester, which had a population of about 200,000 at the time, and was now given two MPs. However, pocket boroughs, whose seats in parliament were essentially controlled by a local nobleman, continued to exist until the Reform Act of 1867, 30&31 Vict., c. 102.

⁵³⁴ Cf. section 6 of the Railway Clauses Consolidation Act (1845), 8&9 Vict., c. 20.

⁵³⁵ Cf. Smiles (1859, 1868).

Manchester, the construction of the Stockton & Darlington, and several other projects. When he appeared before the Parliamentary committee to testify and to be cross-examined by the opponents of the bill, he turned out to be woefully unprepared.

In widely publicised hearings lasting several days, Edward Alderson, counsel for the objectors, tore into the witness, exposing the many errors in his survey.⁵³⁶ In Stephenson's defence, it is of course difficult to produce an accurate survey if the work has to be done under cover of night while trespassing in order to avoid being shot at or having the dogs set on you.⁵³⁷ Still, Stephenson always considered this event to be the low point in his career, and it continued to affect his reputation for several years, just as the railways were becoming a reality. But for the vigorous opposition by Alderson, most of the flaws in the plans would not have come to light until after incorporation had been granted, during the actual work. That would have either required the company to go through the costly process of asking Parliament for an amendment, or it would have deprived Parliament of the ability to adequately supervise the work being done, in case the fixed had been implemented without an amending act. Despite the preposterous arguments that were occasionally advanced by the opponents of the bill, the adversarial nature of the process allowed Parliament to acquire information that it did not itself have the institutional capacity to acquire. Using this information, it voted down the bill in May 1825.

With Stephenson replaced and the survey redone, the bill did get passed at the second attempt in May 1826.⁵³⁸ This time, Parliament received the correct information, and it took the decision that was in the best interest of the country. The question remains, though, whether this was really the most efficient way in which the budding railway sector could have been governed by the state. What are the transaction costs involved? First of all, of course, there are the costs of lobbying and other Parliamentary procedure. As we have already seen, these could occasionally be extremely high, in the order of magnitude of 10% of the total costs of the project.⁵³⁹ For this sum, society obtained an instrument that – at least in theory – relied entirely on a one-time, *ex ante* review. The enabling act was expected to contain all the legal

⁵³⁶ Cf. Hylton (2007), p. 38-39.

⁵³⁷ Cf. Hylton (2007), p. 38 and Wolmar (2007), p. 24.

⁵³⁸ Liverpool and Manchester Railway Act (1826), 7 Geo. 4, c. xlix, as amended by the Liverpool and Manchester Railway Act (1827), 8 Geo. 4, c. xxi, the Liverpool and Manchester Railway Act (1828), 9 Geo. 4, c. vii, the Liverpool and Manchester Railway Act (1829), 10 Geo. 4, c. xxxv, the Liverpool and Manchester Railway Act (1831), 1 Will. 4, c. li, the Liverpool and Manchester Railway Act (1839), 2&3 Vict., c. xli, and finally the Grand Junction Railway Act (1845), 8&9 Vict., c. cxviii, which authorised its merger with the Grand Junction Railway. The early amendments are discussed in Hylton (2007), p. 41.

⁵³⁹ Cf. Wolmar (2007), p. 10 and p. 26-27

rules imposed on the company in return for the state's blessing. Subsequent amendments were of course possible, at the request of, and expense of the railway company, in order to make small changes to the route or to make other small changes to the agreement⁵⁴⁰, but that does not take away from the fact that the enabling act was intended to fix the regulatory regime of the company for all eternity, despite the dizzying uncertainty involved. In later decades, Parliament standardised these enabling acts, by enacting the Railway Clauses Consolidation Act (1845)⁵⁴¹, but under article 1 of the Act it only applied to future acts, presumably on the theory that general application would be improper as a violation of the companies' property rights and legitimate expectations.⁵⁴² However, as we have seen, it soon became clear that more general changes needed to be made, causing Parliament to enact successive acts creating a more comprehensive set of rules for the railways. These acts frequently overruled the earlier enabling acts. The Liverpool and Manchester Railway Acts, for example, followed the old rule which required the barriers at level crossings to cover the railway, rather than the road. When Parliament decided to change this rule in the early 1840s, they did not confer with the railway companies with an eye towards amending their enabling acts, but rather enacted the Railway Regulation Acts.⁵⁴³ While this particular change favoured the railways, in general such an approach does not encourage efficiency.

Essentially, Parliament's willingness to return to agreements made exposed it to two types of problems. On the one hand, this meant that the railway companies could behave like the cable companies discussed by Williamson.⁵⁴⁴ They could initially make whatever promises Parliament wanted to hear, confident in their ability to later secure an amendment or a regulatory Act fixing their problem. The backers of the Stockton & Darlington were initially noncommittal about their preferred method of traction in part because they were genuinely unsure, but mostly because downplaying the aspect of traction was the easiest way to get Parliamentary approval. Similarly, the Liverpool & Manchester started out as just an infrastruc-

⁵⁴⁰ Cf. Hylton (2007), p. 41, discussing the early amendments to the Liverpool and Manchester Railway Act (1826), which dealt with matters such as the route, the gauge and the method of traction. More importantly, the 1829 amending act authorised the company to act as a carrier.

⁵⁴¹ 8&9 Vict., c. 20.

⁵⁴² Of course, the latter doctrine was only fully established by the 1984 House of Lords case of *Council of Civil Service Unions v Minister of the Civil Service*, [1984] 3 All ER 935, [1983] UKHL 6, [1984] 3 WLR 1174, [1985] ICR 14, [1985] AC 374, [1985] IRLR 28. The sanctity of property rights was already laid down in the Magna Charta, to be precise in clause XXIX of the 1297 version, but its application to these facts would have been impossible in 19th century English law, cf. Commons (1924), p. 11-28 (discussing the pioneering line of American Supreme Court cases in the late 19th century extending the notion of property). Nevertheless, it would have certainly been considered improper for Parliament to unilaterally and retroactively change all enabling acts.

⁵⁴³ *Op cit.*

⁵⁴⁴ Cf. section 2.4.3, above.

ture company, but by the time the line actually opened it had not only persuaded Parliament to let it act as a carrier⁵⁴⁵, but also managed to get the authorities to look the other way as the company essentially discouraged other carriers as much as possible.⁵⁴⁶ This way, the company took advantage from the lessons it had drawn from the Stockton & Darlington while at the same time avoiding the political difficulties that would have ensued if they had asked for these things in the first instance.

Apart from inviting later rent seeking and initial reticence, Parliament's willingness to revisit its initial regulatory regimes also hurt efficiency by increasing uncertainty. As we have seen in section 2.4.1, above, Williamson argued that uncertainty increases transaction costs by increasing the range of states of the world for which an action set is to be defined if a contract is to be viable. In this way, an increase in uncertainty means an increased likelihood of vertical integration or total failure. Since vertical integration, i.e. government operation of the railways, was not politically possible in 19th century England, the hypothesis would be that an increase in uncertainty means that some enterprises were not pursued even though they might otherwise have been viable, because the would-be backers worried about being "held up" by the government.

It is important to be clear about this: a straightforward application of the Williamson model to the railway sector would lead one to conclude that it is the railway companies who are the potential victims of opportunistic behaviour by the government, not the other way around.⁵⁴⁷ After all, the theory is that an investment in specific assets creates an exposure, and the most obvious specific investment is made by the railway company. Whatever investment was made by the government in specific human capital is negligible in comparison. Moreover, from a strictly legal point of view, the incompleteness of the contract between the state and the company is extremely one sided. Whatever might be said about the possibility of writing a contract that cannot be rewritten⁵⁴⁸, there is no question that the doctrine of parliamentary sovereignty implies that the state's obligation to obey the terms of the contract depends on its own continued willingness to do so. In the words of Dicey (1885): "Parliament (...) has, under the English constitution, the right to make or unmake any law whatever; and further, that no person or body is recognised by the law of England as having a right to override or set

⁵⁴⁵ Liverpool and Manchester Railway Act (1829), 10 Geo. 4, c. xxxv.

⁵⁴⁶ Cf. Hylton (2007), p. 42-43.

⁵⁴⁷ The same goes for the Oakland cable company discussed in Williamson (1976), of course. Cf. section 2.4.3, above.

⁵⁴⁸ Cf. Hart & Moore (1999), p. 118-124.

aside the legislation of Parliament. *Parliament is not bound by its predecessor.*⁵⁴⁹ One example of such a change of mind was already discussed: Parliament unilaterally changed the rule for level crossings in 1844. Analysing this situation through the lens of normal contract law, we would have to conclude that all of the government's undertakings are illusory⁵⁵⁰, meaning that no contract actually exists.⁵⁵¹

It follows that the Williamson framework suggests that no private company should be willing to enter the railway industry in Britain, because they should fear opportunistic behaviour by the state. For example, the company might fear that the government might change the rules to benefit important constituents after the construction of the rails has finished, for example by imposing a stricter speed limit than before, or increasing or decreasing the prices that are to be charged.⁵⁵² Such fear would have been justified, since the government certainly did intrude in the running of the railways more and more as time went on, starting with Gladstone's attempt at widespread nationalisation in 1844. The Act that emerged from that – otherwise abortive – attempt, the Railway Regulation Act (1844)⁵⁵³, not only changed the rule for level crossings, but also required the railways to run special trains that became known as Parliamentary trains, which were required to run at an average speed of at least 12 mph and charge a price of no more than 1 pence per mile.⁵⁵⁴ Despite the best efforts of the railway companies, these trains turned out to be a great success.⁵⁵⁵ Of course, that does not take away from the fact that this was a case of opportunistic behaviour by the government, rewriting the terms of the contract at will.

Even though the railways sector would seem to be the archetypal case of a sector where Williamson's Transaction Cost Economics would be a useful explanatory tool, it appears to fail miserably at explaining why a transaction that is characterised by enormous investments in specific assets, enormous uncertainty and enormous incompleteness would nonetheless be governed by a contract instead of a hierarchy. The only way to save the model is to

⁵⁴⁹ Dicey (1885), p. 38. Emphasis added.

⁵⁵⁰ “[The doctrine of illusory promise] is used to invalidate contracts where one party has not actually bound itself to do anything. The traditional example of an illusory promise is “I will give you ten dollars if I feel like it.”” Lowe (2007), p. 710. Cf. section 2.1.5, above.

⁵⁵¹ N.B. this conclusion is dependent on the doctrine of parliamentary sovereignty, meaning that it does not apply with the same severity in countries other than the United Kingdom.

⁵⁵² A higher price for rail transport could result from rent seeking by competing industries, such as the canals. A lower price could be the result of a desire to appease the general voting public, similar to the attempts in the 21st century by the European Commission to win over voters by making phone calls cheaper.

⁵⁵³ 7&8 Vict., c. 85.

⁵⁵⁴ Cf. Wolmar (2007), p. 78. Note that price is expressed in pre-decimalisation currency, meaning that the maximum price can also be expressed as 1/240 of a pound.

⁵⁵⁵ Cf. Wolmar (2007), p. 78.

invoke additional elements, either from other branches of economic science, or from one of its sister social sciences. One could, for example, posit the existence of trust between the government and the private sector, although that would come dangerously close to positing the absence of opportunism, which is one of the fundamental assumptions of the Williamson model.⁵⁵⁶ While trust has occasionally been modelled as an economic concept, such attempts usually rely on reputation mechanism, meaning that the thing they model is the very opposite of trust.⁵⁵⁷ Another way to save the model is to incorporate aspects of path dependency or public choice, for example by pointing out the high number of MPs who had some kind of connection to the railway industry. This possibility is discussed further in chapter 4.

*3.1.4 The Netherlands*⁵⁵⁸

Unlike the UK, the Dutch government did explicitly consider what the role of the government should be. In 1836, the King ordered an official inquiry, under the chairmanship of Anton Reinhard Falck, in order to examine whether the construction of railways would be a good idea, and if so where. A third question put to the committee concerned the appropriate role of the state in any railroad construction.⁵⁵⁹ Unfortunately, the answer to this last question focused mostly on the various ways in which the state could encourage private endeavours, limiting its evaluation of the possibility of total state construction to the observation that leaving the initiative with private enterprise would tend to guarantee that no money would be wasted⁵⁶⁰, an observation that could reasonably have been doubted by anyone familiar with the periodic railway bubbles that had already occurred in Britain since the late 1820s.⁵⁶¹

⁵⁵⁶ An interesting attempt to do this is Lulfesmann (2007), who modifies the GHM framework to work with a “benevolent government”.

⁵⁵⁷ Cf. Williamson (1993a) for a discussion of the problems involved in incorporating trust in economic modelling.

⁵⁵⁸ Apart from the general books on Dutch railway history, such as Jonckers Nieboer (1907, 1938) and Veenendaal (2004), an important source for this section is Fritschy (1983).

⁵⁵⁹ Cf. Fritschy (1983), p. 196-197, Jonckers Nieboer (1907), p. 14-20, Veenendaal (2004), p. 27. The committee’s rapport is available in Van Citters & Van Roosendaal (1877-1944), Vol. 1, p. 7-26.

⁵⁶⁰ “In die vrijlating [van private initiatieven] ligt in den regel de zekerheid opgesloten, dat in het wijzigen en uitbreiden der middelen van gemeenschap het algemeen belang met het particuliere vereenigd en geenszins om hersenschimmige voordelen of ten behoeve van eene onberedeneerde voorkeur, eenige gedeelte van het kapitaal nutteloos verspild worden zal.” Van Citters & Van Roosendaal (1877-1944), Vol. 1, p. 22. This passage is also quoted in Jonckers Nieboer (1907), p. 19-20.

⁵⁶¹ Cf. Hylton (2007), p. 118-120. Wolmar (2007), p. 21: “Lines worth a total of £ 22m. (about £ 1.32 bn today), an unprecedented amount of capital at the time, including an ambitious scheme for a London-Edinburgh railway, were put forward in 1824-5, though most never got further than a prospectus and a vague scrawl on a map.”

In response to the committee's report, the government got to work creating a rail connection between Amsterdam and Cologne, the one route specifically prioritised by the Falck committee to the point of suggesting that government intervention might be defensible for that project. Unfortunately, the Dutch parliament did not necessarily agree, and the bill proposing the use of 15 million guilders of East-Indies money⁵⁶² for a railway between Amsterdam and Arnhem, with a branch line between Rotterdam and Utrecht, ran into severe difficulties as soon as it reached the relevant parliamentary committee. The government's reply gives us the opportunity, for the first time in this chapter, to talk about transaction costs as they are perceived by the relevant actors, instead of merely using a transaction costs framework to explain why the choices made were right or wrong.⁵⁶³

Quoting the government's reply at length:

*Het is natuurlijk in overweging moeten komen, of deze geheele onderneming niet aan particulieren zoude kunnen worden overgelaten tegen den afstand van alle opbrengsten, welke zoodanige weg zal kunnen opleveren en inderdaad zou daartoe mede op zich zelve beschouwd, afgescheiden van alle andere omstandigheden en wanneer men aan het denkbeeld van het vestigen van een monopolium ten behoeve van particulieren niet wilde hechten, redden genoeg kunnen bestaan; doch, al wil men ten regel aannemen, dat al wat van dezen aard aan particulieren overgelaten kan worden, door het Rijk niet behoeft of behoort te worden ondernomen, zoo zijn er nogtans gevallen, waarin deze regel uitzondering lijdt en men vermeent, dat zoodanig geval alhier bestaat, vermits, wanneer de kapitalen van particuliere associaties op zoodanigen weg mochten gevestigd zijn, de tarieven der opbrengsten noodwendig onveranderd als de hypotheek daarvan zouden moeten blijven bestaan, zonder dat men meester zou zijn om naar gelang van de omstandigheden daarin verandering te brengen, welke omstandigheden nogtans zich zouden kunnen voordoen, indien men in alle gevallen de onderneming haar doel zoude willen zien bereiken en zich kunnen wapenen tegen eene mededinging, die dit noodzakelijk maken mogt.*⁵⁶⁴

⁵⁶² In a statute of 24 April 1836, the state had decided that the colonies in the East Indies owed it a debt of f 140 million in compensation for "the burdens previously undertaken by the state for the colonies". Cf. Jonckers Nieboer (1907), p. 21.

⁵⁶³ Cf. section 1.2, above.

⁵⁶⁴ Jonckers Nieboer (1907), p. 22-23. Present author's translation: "It was of course necessary to consider whether the entire enterprise could not be left to private parties in return for all revenues, which such a road shall produce. And indeed, there could be sufficient grounds for that, viewing the problem in itself, separate from all other circumstances and assuming one does not mind establishing a monopoly for the benefit of private parties. Nevertheless, even if one is willing to consider the general rule that everything that can be left to private parties need not and should not be undertaken by the State, there are nonetheless cases where that general rule requires an exception, and one considers that such an exception occurs here. Since, when capital of private associations should be invested in a road of this kind, the future revenues necessarily need to continue to exist as a mortgage thereon, without allowing the possibility to make changes therein, should the circumstances so require. Such circumstances could nonetheless occur, if one would like to see the enterprise reach its goals in all cases, and to enable itself to defend against any competition, which might make [a change] necessary."

In other words, the government objected that privatisation of such an important route would require an overly rigid regulatory structure, which they would not be able to adapt to changing circumstances. This is, of course, the beginning of a transaction costs argument. What is missing is an explanation why renegotiation should be foreclosed.

A good place to start looking for an answer to that question is in the Royal Decree authorising the Netherlands' first railway: the HSM⁵⁶⁵ railway between Amsterdam and Haarlem.⁵⁶⁶ The Amsterdam and Haarlem line had been authorised two years before the debate about the Rhine Railway, in 1836, and was a strictly private initiative similar to the early English railways discussed above.⁵⁶⁷ When the government argued that a strictly private railway would not do for the route between Amsterdam and Germany, they must have been thinking of a private railway of the form chosen for the HSM. The curious thing is that the enabling act for the Amsterdam & Haarlem railway already gave the Dutch state a great deal more flexibility than equivalent British Acts:

- Article 2(a) required the company to negotiate a more detailed contract with the Ministry of the Interior, which might contain additional regulations;
- Article 4 gave the state the authority to nationalise the railway, though no earlier than 33 years after its opening. The price the state would have to pay was fixed at the total of investments for initial construction and later improvements.
- Article 5 fixed the “toll” rate at 2½ cents per km for passengers and 5 cents per tonne per km for freight. However, the passenger rate was only an average, to be divided over the different classes in a scheme yet to be agreed, and open to later amendment.
- Article 6 allowed the crown to later reduce the maximum toll if the profit of the company should exceed 10%.
- Article 7 banned the levying of any other tolls, for the benefit of either the state, or the province, or any third party.
- Article 9 created some flexibility for charging other costs in relation to the transport of goods.

⁵⁶⁵ The name of the company was *Hollandsche Ijzeren Spoorweg Maatschappij*. At different times and at different occasions, this was either abbreviated as HSM or HIJSM. Cf. Veenendaal (2004), p. 554, fn. 4. In the present thesis, we will use the acronym HSM throughout.

⁵⁶⁶ Koninklijk Besluit van 1 June 1836, no. 75: *Concessie voor den aanleg van eenen spoorweg van Amsterdam tot Haarlem*, reproduced in full in Jonckers Nieboer (1907), p. 236-239.

⁵⁶⁷ Typically for the murky constitutional system of the 1814-1848 period, there was some confusion about whether or not it was lawful for this authorisation to be done in the form of a Royal Decree, rather than a statute, especially given that it authorised the use of eminent domain power (art. 3). Cf. Van der Pot & Donner (2001), p. 118-121 and p. 563-565 for a discussion of the distribution of powers between the king and parliament during that period.

- Article 10 mandated that the railway should be based on the principle of open access, just like the Stockton & Darlington was.
- Article 11 established that the company should have the same obligations in relation to the mails as other carriers.
- Article 12 required the company to obey and enforce any regulations that would be enacted for the railways.⁵⁶⁸
- And Article 13 put the Ministry of the Interior in charge of enforcing the rules of the concession, giving it the authority to demand information of various nature.

Clearly flexibility as such is not the problem. Instead, the more plausible reading of the government's demurrer is that privatisation would add an additional veto player, in the manner of the GHM models.⁵⁶⁹ Just like in the case discussed by Hart & Moore (1990)⁵⁷⁰, having the skipper own the yacht means that there are three veto players instead of two, the government's argument expresses concern about having two veto players instead of one.

Applying the Hart & Moore (1990) framework to the railway sector shows that there are always going to be at least two veto players. Even if the state owns all the physical assets involved, it will still need a number of people to invest in transaction specific human capital, i.e. the knowledge of how to operate the railways. At the same time, it is exactly that actor, the state's counterparty, who is expected to apply effort to reduce costs and increase the value generated by the railways, in other words, to innovate. It follows that the better question is how to stop the would-be innovator from having the value generated by his innovation appropriated by the state. In fact, but for the fact that we cannot have the state taken over by the railway company, proposition 2 of Hart & Moore (1990) would suggest exactly that.⁵⁷¹ Unless we can identify an "investment" – in the sense of that proposition – that is expected of the state, the model's application to the present problem can only be saved by imposing an exogenous constraint⁵⁷² saying that the state may not be taken over by the railway company. In

⁵⁶⁸ "Zij zijn gehouden de reglementen, zoo van policie als andere, die op het gebruik der spoorwegen in het algemeen en van den hinnen in het bijzonder gemaakt zullen worden, naauwkeuring in acht te nemen en door hunne bedienden in acht te doen nemen."

⁵⁶⁹ Cf. section 2.5, above.

⁵⁷⁰ Cf. section 2.5.3, above.

⁵⁷¹ "If only one agent *i* has an investment, then he should own all the assets." Hart & Moore (1990), p. 1131.

⁵⁷² The term is chosen deliberately to invoke a type of linear programming approach to solving the governance problem.

itself, this constraint would be eminently reasonable, but it would be preferable if the model could be made to work without the need to resort to exogenous constraints.⁵⁷³

While this solution comes dangerously close to playing with semantics, one might argue that the state's investment consists of restraint. Exceptions aside, it continually refrains from interfering with the railway company's business. This restraint cannot be contracted, and if the railway company were to vertically integrate and acquire the state, it still would not be able to compel this restraint through fiat power.⁵⁷⁴ The only possibility is to make the state a residual claimant.⁵⁷⁵

Alternatively, the Dutch government's argument can be cast in terms of classical Coasian transaction costs: adding external shareholders as potential veto players, and increasing the bargaining power of the railway managers by assigning additional residual control rights to them means that negotiating will inevitably be more costly. This aspect cannot be captured by the GHM models, because they assume costless and perfect *ex post* bargaining. In reality, though, negotiations between the state and private actors in a privatised sector can be difficult and costly, and it is quite rational for the government to prefer to assign as much bargaining power to itself so as to increase its future room for manoeuvre.

In 1838, however, the discussion was a good deal simpler. For a majority of the Dutch parliament, the idea of direct government involvement in the Rhine railroad was a bridge too far. In April 2nd of that year, the vote in the Second Chamber was 46 to 2 against, with one of the two yea votes coming from the minister for Finance himself, who had defended the bill in parliament. That was not the end of the story, though. Instead of accepting his defeat, Dutch King Willem I decided to use his significant constitutional powers to bypass parliament's decision. On April 30th, 1838, he signed a Royal Decree mandating the Ministry for the Interior to organise the construction of a Rhine Railroad from Amsterdam to the German border near Arnhem, with a branch line from Rotterdam to Utrecht.⁵⁷⁶ The result was an organisation whose legal nature was so unclear, that academics have been unable to ascertain its exact na-

⁵⁷³ Cf. section 1.2, above.

⁵⁷⁴ Cf. art. 83 of the Dutch Constitution of 1815: "De leden dezer [Tweede] Kamer stemmen voor zich zelve, en zonder last van, of ruggespraak met de vergadering, door welke zij benoemd zijn." A similar rule is currently established by art. 67 (3) of the Dutch Constitution.

⁵⁷⁵ Of course, there are limits to the state's freedom, such as the protection of property mandated by art. 4 and 164 of the 1815 constitution, but those have little independent force given that Dutch courts, like their British counterparts, are generally denied the right to declare statutes unconstitutional.

⁵⁷⁶ The Decree is reproduced in Van Citters & Van Rosendaal (1877-1944), Vol. 1, p. 78-79.

ture to this day⁵⁷⁷ – it was the ultimate hybrid. On the one hand, it was not really private. Unlike the HSM, the Rhine Railroad was not an NV, a public limited company. In fact, it was not incorporated in any way know to private law. Since it apparently did have legal personality, this legal personality must have been of a public law origin. Moreover, the construction was carried out under the responsibility of the ministries of Internal Affairs and Public Works.



Figure 5: The Rhine Railroad

At the same time, the Rhine Railroad was not fully public either. The money to build it was raised from private investors, with the King only contributing a guarantee as to the minimum return on their investment.⁵⁷⁸ In the end, this mess was only sorted out shortly after the railroad was opened in 1845, when the state transferred whatever ownership and control rights it might have had to a consortium of six private investors. As of 1 September 1845, the company would be fully private, and known as the Nederlandsche Rhijnspoorweg Maatschappij (NRS).⁵⁷⁹ Together with the HSM, which completed its connection between Amsterdam and

⁵⁷⁷ Cf. Gravendaal (1990) for a valliant effort. In fact, apart from its legal nature, there were a number of other dubious aspects to this story, which are, however, outside the scope of this thesis. Cf. Van den Broeke & Van Nederveen Meerkerk (2001), p. 5-6 and Fritschy (1983), p. 197-198.

⁵⁷⁸ Under art. 6 of the Royal Decree, the King would guarantee a minimum return of 4,5%. Art. 7 established that any sums paid by the King under art. 6 would have to be repaid by the Railroad as soon as it was able.

⁵⁷⁹ Cf. Veenendaal (2004), p. 63. This privatisation was, again, a curious affair. Given the relative success of the enterprise, there was a significant faction in parliament who were willing to consider the possibility of voting money for full nationalisation. This course of action had even been recommended by the commission of inquiry

Rotterdam via The Hague in 1847⁵⁸⁰, the NRS would be a prominent player in the Dutch railway industry for the next 50 years.⁵⁸¹

3.1.5 France

Compared with their British and Dutch counterparts, the French authorities displayed a much greater degree of purposiveness. In fact, they did exactly what the armchair scholar would like to see: they examined the situation, discussed the matter, and designed a governance structure to deal with this new phenomenon.⁵⁸² Unfortunately, it did take them a rather long time to get there.

As far as the basic technology goes, France had about as much experience with railroads as England did, having had rudimentary railroads at least since the time of Louis XIV, and not only for the King's amusement⁵⁸³, but also for use in the mining industry in the Massif Central.⁵⁸⁴ As a result, the first efforts at railway building in France followed quickly after the English example. On February 26, 1823 King Louis XVIII granted a concession for a railway between Saint-Étienne and the river Loire at Andrézieux, over a distance of 23 km. This railway was entirely the result of private initiative. It was pioneered by Louis Antoine Beaumier, director of the mining college in Saint-Étienne, who wrote a letter to the King in 1821 asking for the concession, and in 1824 became the first director of the *Compagnie du chemin de fer de Saint-Étienne à la Loire*.⁵⁸⁵ While being the first railway on the European continent, opening a little over two months before the first section of the Budweis–Linz–Gmunden railway in Austrian Bohemia⁵⁸⁶, the Saint-Étienne line still showed many signs of its origins: Its primary purpose was to transport coal – in fact it did not open to passengers until 1832. Until 1830, traction was exclusively by horse power, with gravity being used for downward travel, although the gauge of the track was chosen with Stephenson's locomotives

of 1844, which had investigated possible improprieties surrounding the founding of the company. However, no such proposal was ever put to parliament, presumably because the combination of the King's guarantee and the previous King's financial and other creativity regarding the company's accounts meant that privatisation was the only way to extricate the King from this mess without causing a scandal. Cf. Fritschy (1983), p. 198-200.

⁵⁸⁰ The NRS connection between Rotterdam and Utrecht was completed in 1855

⁵⁸¹ The third of the "Big Three", the Staatsspoorweg, was only founded in 1860. Cf. section 3.2, below.

⁵⁸² Cf. Caron (2005b), p. 15.

⁵⁸³ Cf. Mayor (1998) and Wolmar (2007), p. 1.

⁵⁸⁴ Cf. Caron (1997), p. 27-30.

⁵⁸⁵ Ordonnance of July 21, 1824.

⁵⁸⁶ The former opened on June 30, 1827, the latter on September 7th.

in mind, causing his 4ft 8½in “Standard Gauge” to become standard in the European continent as well.

As further ventures followed, the debate continued. Even more than the *ancien regime*, the Restoration France was profoundly conservative.⁵⁸⁷ In a political climate focused on the position of the King, the dominant political factions logically were the loyalist conservatives, who were naturally inclined to resist any kind of innovation, and pro-revolution liberals, who were equally naturally favoured English-style *laissez-faire*. Only after the position of the King had been resolved by the July revolution of 1830⁵⁸⁸, and after the first railways had been in operation for a while, was there room for a pro-government intervention party to emerge in the debate. In so doing, they pointed not only to the experience France already had with government-constructed roads and canals, but also to the characteristics that set France apart from its peers, particularly the United Kingdom. The most important among these was the difference in terrain. In England, even when the early railroads were constructed in areas more difficult than the low lands of Lancashire⁵⁸⁹, they generally covered only short distances. In France, on the other hand, the most obvious routes to start with were either very long, such as the connections between Paris and the southern port cities of Marseille and Bordeaux, or through difficult terrain, such as the early railroads in the coal mining region of the Massif Central near Saint-Étienne.

Of course, high costs of construction alone are not enough reason for government intervention. However, there was a second problem. As we have seen, the early railway companies were among the largest enterprises ever undertaken, especially if we ignore the colonial trading companies as being essentially state guaranteed. In Britain, where the industrial revolution had already been going on for several decades, this simply meant added pressure for the repeal of the Bubble Act and the creation of a comprehensive legal framework for joint stock companies. In France, on the other hand, there were grave doubts as to whether there was any private party capable of organising such a thing. The only candidates seemed to be *la haute banque parisienne*, consisting of such banking pioneers as James de Rothschild and the *saint-simoniste*⁵⁹⁰ former Prime Minister⁵⁹¹ Jacques Laffitte.⁵⁹² While they were certainly consid-

⁵⁸⁷ Cf. Fyffe (1895), chapter XIII.

⁵⁸⁸ Cf. Fyffe (1895), chapter XVI.

⁵⁸⁹ Which is where the Liverpool and Manchester was built.

⁵⁹⁰ Claude Henry de Rouvroy, count of Saint-Simon (1760-1825) was an important precursor of the socialist thinkers of the later 19th century. His writing speculated about a post-industrial age where everyone would be equal. He inspired not only early socialists such as Proudhon and Marx, and early French feminists, but also industrial pioneers such as the builder of the Suez Canal, Ferdinand de Lesseps and the banker Jacques Laffitte

ered capable of financing the railways, it was less clear whether they were also capable of organising a national railway system.

In addition, there was one final critical difference between France and the United Kingdom which tilted the balance towards greater government interference in the railways. Even though in the 1830s Western Europe was in the middle of an unprecedented period of peace, commonly known as the Concert of Europe⁵⁹³, in the 1830s the Napoleonic Wars were still fresh in people's minds, most of all in France itself. The Emperor had found himself fighting simultaneously or in quick succession on all sides of France, from the 1807-1814 Painsular war in and against Spain and Portugal⁵⁹⁴ to the constantly recurring Coalition wars that were fought mostly in Germany⁵⁹⁵, and from the recurrent fighting in Italy⁵⁹⁶ to the sea battles against the British. Clearly France, like the united Germany of the 20th century, faced a very real risk of having to fight a war on many different fronts at the same time. And even if it should turn out to be possible to avoid this fate, there was still the simple prospect of having to fight a land war somewhere on the border without being able to predict with any certainty which one. In the end, it turned out to be a war with Prussia in 1870.⁵⁹⁷

Concern for war is not necessarily an argument for government involvement. When Wolmar (2007) discusses the performance of the British railways during World War I⁵⁹⁸, he concludes that their overcapacity was an important contributing factor in their outstanding performance, an overcapacity caused by 90 years of almost unfettered capitalism. By that time, competition between the major railway companies had caused a great amount of redundancy: almost any connection between major cities could be achieved in more than one way. For example, both the London & South Western and the Great Western connected London

mentioned in the text. It also earned him a spot as one of Isaiah Berlin's "Six Enemies of Human Liberty", Berlin (2002), p. 105-130.

⁵⁹¹ 2 November 1830 – 13 March 1831, i.e. shortly after the "*Trois Glorieuses*" of 1830. Earlier, Laffitte had been governor of the *Banque de France* (1814-1819).

⁵⁹² Cf. Cameron (1991), p. 7.

⁵⁹³ There was no large power war between the battle of Waterloo in 1815 and the Austro-Prussian War of 1866, except the Crimean War of 1853-1856, which occurred far away from Western Europe. The Concert System is discussed in chapter 4 of Kennedy (1988).

⁵⁹⁴ Cf. Fyffe (1895), chapters VIII and IX.

⁵⁹⁵ These included the War of the Third Coalition (1803-1806), the War of the Fourth Coalition (1806-1807), the War of the Fifth Coalition (1809) and finally the War of the Sixth Coalition (1812-1814) that resulted in Napoleon's downfall. Cf. Fyffe (1895), chapters V-XI.

⁵⁹⁶ Cf. Fyffe (1895), chapter III, in particular.

⁵⁹⁷ Cf. Fyffe (1895), chapter XXIV. One of the proposed reasons for France's swift defeat in this war was the inefficiency of its railway system, with many connections between Paris and the borders, but not enough cross-connections between lines, which was contrasted with the Prussian railway system which was structured more as a grid. Cf. Caron (1997), p. 419-426.

⁵⁹⁸ Wolmar (2007), chapter 11.

with the key port cities of Southampton and Portsmouth, the former running a direct line from Waterloo station, while the latter made the connection from Paddington station via Reading. That way, whenever there was a problem on one of the lines, trains could be rerouted using the other one.

While the idea that free competition might lead to a useful amount of overcapacity makes sense from a 20th century point of view, this would have been an unrealistic suggestion in the 1830s. At a time when railways were still an extremely risky and extremely costly venture, the more realistic concern was over too few railways, particularly in areas with little commercial importance, difficult terrain and great strategic value, such as the relatively impoverished region of Alsace-Lorraine, whose modest economic value did not stop it from being highly contested during the more than 300 years between the start of the Thirty Years' War and the end of World War II. If there was going to be a railway between Alsace-Lorraine and the rest of France, the argument ran, the government would have to build it itself.

A related question was whether the private sector could be trusted to build the “correct” routes at all. As we have seen, the Dutch government argued the more traditional point that private sector primacy was a guarantee against wasteful construction. In France, however, the question arose whether for certain routes the private benefits might not end up being so much lower than the social benefits that socially beneficial railroads might nonetheless produce insufficient private benefit for any private party to undertake the project. Consider this: all benefits enjoyed by the railway's customers – either directly or indirectly – can be captured by the company, meaning that the price of the ticket can be made to depend on them.⁵⁹⁹ This gives the company a socially suboptimal incentive, given that it can only ever capture part of the marginal benefit generated by its choices, but at least there is some link between the company's revenue and the social benefits it creates. A greater problem is posed by the industry's positive externalities. For example, in a time without telegraph⁶⁰⁰, the opening of a railway line allowed for a massive improvement in communication. The benefit to both government and business is difficult to overstate. To the extent that these positive externalities are significant, they are a reason to fear for the private sector's ability to construct a socially optimal railroad network. That problem concerns not only the choice of which cities to connect, but also the manner in which they should be connected.

⁵⁹⁹ Ignoring, for now, the normal custom of fixing the ticket price in the concession.

⁶⁰⁰ While semaphores were already used in France as early as the 1790s, Morse code was only invented in 1837 and the first commercial telegraph was operated by the Great Western in England in 1839.

As described above, the early railroads in Prussia were often organised by local railroad committees. One consequence of that system is a tendency to try to connect too many cities. The early English railroads, on the other hand, tended to suffer from “l’esprit géométrique”⁶⁰¹, i.e. from a preference for straight lines over the possibility of linking more towns with the railroad. The Grand Junction railway between Birmingham and Manchester, for example, was opened in 1838 without any connection to the towns of Walsall, Wolverhampton and Northwich, even though all three were located less than five miles from the line and could easily have been accommodated.⁶⁰² In France, though, it was difficult to see whether it was the private sector or the government who had the greater tendency to draw straight lines.⁶⁰³

The prolonged French discussion also produced an entirely uneconomic argument. The Director-General of the *corps des ingénieurs des ponts et chaussées*, Alexis Legrand, and many others of a similarly conservative bent, seemed to argue for government involvement on principle. This is not the Saint-Simon utopia of a future where society is run as one big business⁶⁰⁴, but the idea that something as important as the railroads should not be left to “mere” private industry, that it should not be run for profit⁶⁰⁵, and least of all by a “*nouvelle féodalité*” or some *pouvoir concurrent*⁶⁰⁶, some private monopoly whose power might rival that of the state. While this last concern was easily remedied by creating a different company for every line or region, the general mistrust of private enterprise in the area of public utilities continues to this day.

As before, some of these arguments fit better with our theoretical framework than others. Considerations of national defence, for example, are undeniably a good reason for government involvement, given that defence is the archetypical public good, since it is both non-rival and non-excludable.⁶⁰⁷ However, there is a reason why the theory of public goods was not discussed in any detail in the previous chapter: the question of whether the government should be involved in a certain industry is distinct from the question as to the nature of that involvement. If the government should decide that a certain railroad ought to be built, it can either

⁶⁰¹ Caron (1997), p. 130.

⁶⁰² Wolmar (2007), p. 65.

⁶⁰³ Cf. the different private and public proposals discussed in Caron (1997), p. 122-134.

⁶⁰⁴ Cf. Berlin (2002), p. 105-130.

⁶⁰⁵ Caron (1997) quotes a government minister “L’État a toujours [l’avantage] lorsqu’il exécute lui-même de faire mieux, puisqu’il opère dans des vues d’intérêt public, vues toujours plus saines et plus hautes que celles des compagnies.” (p. 114)

⁶⁰⁶ Caron (1997), p. 115.

⁶⁰⁷ Cf. Batina & Ihori (2005), or any textbook on law & economics, eg. Eijgelshoven, *et al.* (2000), p. 142-145.

construct and operate the railroad itself, or it can encourage the creation of a certain railway line by offering to contract with a private party on favourable terms, i.e. by proposing a concession that involves a generous subsidy. This latter approach is quite popular for a variety of public or semi-public goods in the early 21st century, but in the 19th century it was not a possibility that was on anyone's mind.

This is generally a problem with the early phases of the history of the railways: a lack of familiarity with the problem did not simply cause the actors to make inefficient choices, but it also caused them to consider only a subset of all the choices they might make.⁶⁰⁸ What was needed was a process of innovation of governance, whereby new modes of governance would be invented to match the technological innovation that gave us the railways themselves. This process did eventually occur, as governments drew lessons from their early efforts, but it took a significant amount of time.⁶⁰⁹

In 1830s France, the discussion tended slowly but gradually towards some measure of government involvement. In 1842, finally, came the enactment of the *Loi relative à l'établissement des grandes lignes de chemin de fer*. Article 1 of the Act specified a list of railroads that would have to be built, from Paris to Lille, to the Canal coast, to Strasbourg, to Lyon and Marseille, to the Spanish border via Bordeaux, to Nantes, and one to “the centre of France” to Bourges. In addition, two other routes were mandated: one connecting the Rhine and the Rhône, from Lyon to Mulhouse, and one between the Atlantic and the Mediterranean, from Bordeaux via Toulouse to Marseille. In other words, the law aimed to create a proper “réseau”, a network, not trusting the private sector to “spontaneously” build the optimal network of railroads.⁶¹⁰

Having fixed, in general terms, the routes that were to be constructed, the act continued by establishing – in equally general terms – the relevant governance structures. Construction was to be carried out in collaboration between the state, local authorities and private industry, although the possibility of granting a concession along the lines of the previous system was expressly reserved.⁶¹¹ However, such a concession would have to be granted by law, not

⁶⁰⁸ Whether later governments suffered from the same problem, or whether they considered but quickly rejected whole realms of possibility, is an open question. Cf. section 3.3, below.

⁶⁰⁹ Cf. sections 3.2 and 3.3, below.

⁶¹⁰ Note that the law gave only the most general of instructions as to the precise route, leaving the details to later statutes, depending on the recommendations of the *Corps des ingénieurs des ponts et chaussées*.

⁶¹¹ Art. 2.

by Ordonnance.⁶¹² Articles 3 – 6 settled the financing of the project, putting quite a bit of the burden on local authorities.⁶¹³ The general idea was to have the state pay for the land, the necessary earthworks and the train stations (art. 5), while the railroad companies that were to be created would pay for the rails, the trains and their upkeep (art. 6). Also, the companies would pay rent to the state (art. 6) and the state would have the right to buy the property of the company upon expiration of the rent contract, at a price to be determined by independent appraisers (art. 7).⁶¹⁴ The total cost of the project was estimated at 126 million francs (art. 10-16).



Figure 6: The system of 1842.

⁶¹² Art. 2, final sentence. The concession for the Saint-Étienne – Andrézieux line, as well as for subsequent railroads before 1830, had been granted by ordonnance, i.e. by royal decree. Under the July Monarchy, founded in 1830, such use of executive power became illegitimate. Cf. *Charte constitutionnelle du 14 Août 1830*, art. 8 and 14.

⁶¹³ This element was removed in 1845. Cf. Caron (1997), p. 149.

⁶¹⁴ Cf. art. 4 of the concession given by the Dutch crown to the HSM, which was discussed in section 3.1.4, above. The difference between nationalisation at cost, and nationalisation at an appraised value should be obvious.

Strangely enough, the 1842 law had an effect that was entirely different from what the drafters intended. When the government finally decided what to do, after more than a decade of uncertainty, the private sector once again became willing to get involved in the railway industry. With the main source of uncertainty removed, *la haute banque parisienne* were willing to lend the money for companies to take on leases or even concessions for parts of the network⁶¹⁵, running them as entirely private enterprises, just like all earlier French railways.⁶¹⁶ As a result, the majority of the lines mandated by law ended up being operated entirely by private enterprises. The July Monarchy, after much debate and starting from vastly different assumptions, ended up producing a system not that much different from the British system, and one that would stay in place until the early days of the second republic.

3.1.6 Belgium⁶¹⁷

The last country we will consider here is Belgium, a country that only became independent in 1830. Having been united with its northern neighbour by the Final Act of the Congress of Vienna⁶¹⁸, 1830 saw Belgium rise up against its Dutch government, and declare independence.⁶¹⁹ The new country was understandably full of energy to take on new projects, and one of the most notable among them was the construction of a proper railway system. As early as December 12, 1830, the government orders a study into a canal or railway between Antwerp and Maastricht.⁶²⁰ Because the establishment of the provisional border between the Netherlands and Belgium in November interfered with the original proposals, the final rapport was only finished in February 1832.⁶²¹ The authors of the rapport, the engineers Simons and De Ridder, proposed a railway between Antwerp and Cologne via Aachen, i.e. just south of the Dutch part of Limburg. More notably, they also proposed a significant degree of government

⁶¹⁵ The last sentence of art. 2 of the 1842 law provided: “Néanmoins, ces lignes pourront être concédées en totalité ou en partie à l'industrie privée, en vertu de lois spéciales et aux conditions qui seront alors déterminées.” Cf. Caron (1997), p. 149 and Caron (2005b), p. 15-18.

⁶¹⁶ A total of 804 km of railway had been opened by 1842. Cf. Caron (2005b), p. 15.

⁶¹⁷ For an early discussion of the history of the Belgian railways, cf. De Economist (1862).

⁶¹⁸ Art. 65. Cf. Fyffe (1895), chapter XII and Gemelli & Royer (1864), p. 27-29.

⁶¹⁹ The rebellion started on August 25, 1830, less than a month after the “*Trois glorieuses*” in France, and the provisional government declared independence on October 4. On November 3, the new country had its first elections and on July 21, 1831, Leopold I became the first king of Belgians. Cf. Fyffe (1895), chapter XVI and Gemelli & Royer (1864).

⁶²⁰ Cf. Devys (1910), p. 15, Laveleye (1862), p. 6, Linters (1985), p. 17 and Simons & De Ridder (1839), p. III-IV.

⁶²¹ Cf. Devys (1910), p. 15, Laveleye (1862), p. 6 and Linters (1985), p. 18.

involvement in the construction and operation of this railway.⁶²² This proposal, however, was not supported by the Ministry for the Interior itself, which is why interested parties were invited to bid for a concession for a railway between Antwerp and Liège.⁶²³ While the proposed concession was in certain respect quite advantageous for the counterparty, for example because the concession would be in perpetuity, there were some stringent requirements as well, formulated by Simons and Deridder.⁶²⁴ On the whole, the private sector judged the proposal insufficiently advantageous, so the debate continued.⁶²⁵

In the end, the resolution of the railway question came much sooner than in France. A railway was to be built quickly, with a minimum of fuss.⁶²⁶ The law of 1 May 1834 is simple enough that it can be quoted in its entirety:

Art. 1. Il sera établi, dans le royaume, un système de chemins de fer ayant pour point central Malines, et se dirigeant à l'est vers la frontière de Prusse par Louvain, Liège et Verviers; au nord par Anvers; à l'ouest sur Ostende par Termonde, Gand et Bruges, et au midi sur Bruxelles et vers les frontières de France par la Hainaut.

Art. 2. L'exécution sera faite à charge du trésor public et par les soins du gouvernement.

Art. 3. Les dépenses de cette exécution seront couvertes au moyen d'un emprunt que sera ultérieurement réglé par une loi.

Art. 4. En attendant la négociation de l'emprunt, il est ouvert au gouvernement un crédit de dix millions, qui sera couvert en tout ou en partie par l'émission de bons du trésor, aux conditions de la loi du 16 février 1833.

Les avances ou les bons du trésor seront remboursés sur les premiers fonds de l'emprunt.

Art. 5. Les produits de la route provenant des péages, qui devront être annuellement réglés par une loi, serviront à couvrir les intérêts et l'amortissement de l'emprunt, ainsi que les dépenses annuelles d'entretien et d'administration de la nouvelle voie.

Art. 6. Avant le 1^{er} juillet 1835, et d'année en année jusqu'au parfait achèvement des travaux, il sera rendu un compte détaillé aux Chambres de toutes les opérations autorisées par la présente loi.

⁶²² Simons & De Ridder (1839), p. IV: "les ingénieurs insistaient dans tous leurs rapports, (...) de faire exécuter le chemin de fer aux frais et par les soins de l'État."

⁶²³ Royal Decision dated 21 March 1832. Cf. Laveleye (1862), p. 6 and Simons & De Ridder (1839), p. V.

⁶²⁴ Laveleye (1862), p. 6-12.

⁶²⁵ For a detailed discussion, including many original documents, cf. Simons & De Ridder (1839).

⁶²⁶ I.e. the lowest possible transaction costs in terms of writing and negotiating a detailed contract that would be necessary to specify the rights and obligations of both the state and the would-be private railway company.

*Art. 7. A dater de l'ouverture du chemin de fer entre Liège et Anvers, le péage sur les canaux du Hainaut sera réduit au taux du péage à établir sur ce chemin de fer, par tonneau et par kilomètre.*⁶²⁷

In other words, a railway network was to be constructed by the government, paid for with public funds obtained from a bonds emission that was yet to be arranged, in the form of a cross with its centre in Mechelen, a city between Antwerp and Brussels. A number of additional lines were added in 1837⁶²⁸, including a line from Gent to the French border near Lille, and a line connecting the city of Namur and the provinces of Limburg and Luxemburg to the rest of the network. Particularly the former is interesting, since it was apparently motivated by the desire to help the *Generale Maatschappij*, the influential bank and investment company that had been set up by the Dutch King Willem I in 1822, and that developed in the post-independence period into a kind of parallel government.⁶²⁹ In the mid-1830s, they were trying to persuade the French parliament to give them a concession for a railroad between Paris and Lille/Valenciennes, a cause certainly advanced by the construction of a connection between Lille and the Belgian port cities of Antwerp and Oostende.

⁶²⁷ Cf. Laveleye (1862), p. 14-15 and Simons & De Ridder (1839), p. CXLI-CXLII.

⁶²⁸ Law of 26 May 1837, cf. Buelens, Van den Broeck & Willems (2008), Devys (1910), p. 19, and Linters (1985), p. 25.

⁶²⁹ Until 1850, the *Generale Maatschappij* also acted as the Belgian central bank.



Figure 7: Belgium under the Laws of 1834 and 1837.⁶³⁰

This makes Belgium the only country in our sample that truly chose for the government option. In the end, the decisive argument seems to have been that the new country needed the railways quickly, as an impulse for the economy and to replace the waterways that previously connected the Walloon mining and industry regions to the world market, waterways that now ran through foreign territory.⁶³¹ While the lawfulness of giving private parties a concession was briefly the subject of debate⁶³², the proponents of state construction seem to have focused almost entirely on the unique situation the country was in, rather than the general propriety of privatising railways.

⁶³⁰ N.B. In 1834, the border between Belgium and the Netherlands had not been fixed yet, while the border between Belgium and Prussia was further west than shown in the figure.

⁶³¹ Remember, the connection between Antwerp and the North Sea had been blocked by the Dutch continuously from the fall of the city to the Spanish in 1585 until the conquest of the Netherlands by the French in 1795. During the “cold war” of 1830-1839, it was not inconceivable that the Dutch would again make things difficult for the new state. When a peace treaty was finally agreed in 1839, the river Scheldt was a central part of that treaty, with ample safeguards for a continued free passage between Antwerp and the North Sea. Cf. Wels (2001).

⁶³² Cf. Linters (1985), p. 18, citing Laveleye (1862), p. 12.

Bearing in mind the budget escalation of the Rhein-Weser railway that we already discussed in section 3.1.1 above, it is interesting to look at the costs of the Belgian railways. Using the data produced under art. 6 of the Law of 1834, Laveleye (1862) came up with the following numbers for Belgium:

Line:	Projected cost per km in 1833:	Real cost per km in 1859:	Increase:
Mechelen – Antwerp	Bfr. 66.000	Bfr. 318.400	382%
Mechelen – Brussels	Bfr. 61.430	Bfr. 323.730	427%
Mechelen – Verviers	Bfr. 97.500	Bfr. 403.100	313%
Mechelen – Oostende	Bfr. 65.300	Bfr. 184.000	182%
Total:	Bfr. 78.680	Bfr. 296.000	276%

Table 1: Construction costs of the Belgian Railways⁶³³

In other words, while the cost increase was less severe than in Prussia, it was still quite significant. It is impossible to say whether the Belgian government would have made the same decision if they had had more accurate cost data, though perhaps we can say that it is fortunate for Belgium that they were never forced to make that decision.⁶³⁴

3.1.7 Conclusion

We have now reached a point where it makes sense to look back to the list of section 2.7, and compare it to the first part of our story:

Transactions: Clearly there was no shortage of transactions taking place. The only question is what it was, exactly, that was being transferred. In a nutshell, the answer is that the state transferred to the companies a series of entitlements, whereas the companies undertook to construct and operate the railways according to certain specifications, as well as in many cases to

⁶³³ Compiled from Laveleye (1862), p. 16-17. Remember that inflation was virtually unheard of in the 19th century, due to the widespread use of the gold standard.

⁶³⁴ Cf. Damart & Roy (2009) on cost-benefit analysis for public transportation decisions.

pay sums of money in dividends or taxes. These transactions were vigorously negotiated at arm's length, without any evidence of inequality in bargaining power.

Overall efficiency: We have seen many instances where the use of the market was defended by pointing to its purported superior production efficiency, as well as at least one case where the transaction costs of using the market were taken as a reason to have the government take the first step itself. The problem is that in many cases this argument seemed a bit like a pretext, given that the efficacy of alternatives had never really been considered. To the extent that the government has preconceived notions of what kind of governance structure it would like to see, quite separate from any social welfare considerations (even when the latter are broadly defined) it is difficult to see how an economic model could explain its behaviour, at least not an economic model that takes the government as a single actor.

Costly information: There is one critical piece of information that always poses particular difficulties for railway builders, and even more so at a time when no one has ever built a railway before: the cost of the project. The budget escalation of the Rhine-Weser railway and the Belgian railways were already discussed. Similar stories can be told about the other early railroads.⁶³⁵ An estimate that is reliable enough to ease the mind of the average decision maker is simply impossible to obtain, except by actually building the thing.

Putting to one side the category of information that is essentially impossible to obtain, we have seen ample examples of protracted negotiating between the railway companies and the state. Sometimes, the former were entirely rejected, as in the early stages of the Rhine-Weser railway and the Stockton & Darlington, while mostly it just took a while to agree on an appropriate division of burdens and rights. Compared to the costs of negotiating, the actors seem to have been much less preoccupied with searching and enforcement. Concessions were often given to whoever asked for them first, under the assumption that it could be transferred – i.e. sold – later if a more suitable party came along.⁶³⁶ As for enforcement, the contracts tended to make few specific demands of the railway companies, not only because the government did not need a great many things from them, but also because it realised that such

⁶³⁵ Cf. Caron (1997), p. 171-200, Hylton (2007), p. 34, 49, 51, and Veenendaal (2004), p. 52. For Switzerland, cf. Van den Wall-Bake (1907a), and for Italy, cf. De Bordes (1886) and Van den Wall-Bake (1907b). For an early discussion of the situation in many countries, cf. Grothe (1859). Note, however, the exception of the Grand Junction railway, where the engineer Joseph Locke and the contractor Thomas Brassey managed to construct the line ahead of schedule and on budget. Cf. Hylton (2007), p. 48-49.

⁶³⁶ As a purely legal matter, there is no free trade in concessions. The government always has to approve any transfer. In practice, however, this should not be a problem if the transfer is to a financially stronger party, or if suitable guarantees are provided.

demands would be impossible to observe. Those demands that the government did make were easy to observe and enforce, such as the transport of mails, access of government rolling stock to the track, etc.

Moral Hazard: Given the amounts of money involved, both sides seem surprisingly unconcerned with the possibility of moral hazard. At times, the railway companies seem quite comfortable granting significant *ex post* flexibility to the state. Similarly, the state does not seem to mind very much that it cannot really observe what the railway company is doing.⁶³⁷ Perhaps the answer to this problem lies in Williamson's probity hazard.

As discussed in section 1.1.3, above, the railway sector is not inherently public, in the sense that it would be unimaginable for it to be entirely private. Railways do not involve "sovereign transactions", or transactions that are substantially similar to sovereign transactions. There is no core element of sovereignty in play here. It follows that our case is almost certainly not one of those cases where this additional transaction characteristic of probity is likely to "surface".

None of this means, of course, that a certain Victorian tendency towards gentlemanly behaviour and rectitude cannot explain the lack of concern for moral hazard, although that attitude certainly seems to have had its limits. Rather, it means that it would be inappropriate to categorise these explanations as a separate variable in an economic analysis. Instead, they should be left where they belong: in the realm of social psychology and sociology.

Fundamental Transformation: Even though a general evaluation of the railway industry might lead one to conclude that the state and the railway companies are forever stuck together, forced to negotiate without either side having any alternative, except nationalisation, this was not true in the early period of the railways. There seemed to be plenty of takers for any one concession, although in practice they tended to work together. Where efforts failed, new companies tended to step in fairly quickly. In other words, whenever the market was used, the transformation from a situation of one-versus-many to one-versus-one did in fact occur.

Asset Specificity: Staggering amounts of money were invested in rail infrastructure alone, investment that had no other value except as a prerequisite for providing rail transport services. In addition, there were enormous investments in the human capital needed to run the railways.

⁶³⁷ Even today, it seems impossible to agree on how many trains are delayed or why. Cf. OECD (2005), *passim* and Vromans (2005).

Uncontractible investments: The notion that something could be uncontractible seems profoundly un-Victorian. As noted: gentlemen keep their promises. That said, the governments of various countries certainly did not hesitate to enact fairly aggressive laws, where theoretically less drastic measures would have sufficed: why would the Act of 1844⁶³⁸, in creating the parliamentary trains, specify in such detail how fast they were required to go and how much a ticket should cost? Because this was the only way to get it done at all; any more flexible rule would have been circumvented by the companies. Even now, the parliamentary trains tended to run very early in the morning, when few travellers were likely to take advantage of them.

Of course, that is merely an example of uncontractibility per se, not of uncontractible investment. The two major categories of investment, infrastructure and human capital, are both fairly contractible, to say the least. Concessions and enabling acts tend to specify in great detail where and how the railway in question is to be built. Investments in human capital are somewhat less easily monitored, but then they were largely considered to be no legitimate concern of the government. The exception to that rule is of course railway accidents, but even there the governments did not seem very keen on mandating safety training directly. That leaves only the uncontractible investment in innovation.

Another possibility is to stretch the definition of “investment” somewhat, to include certain things the government might do. As already discussed, many investments by the government are uncontractible because the government is incapable of promising to refrain from many actions indefinitely. The most important of these, from the perspective of the railway company, is its investment in the project per se. The company needs the government to refrain from doing anything that undermines the railway, such as granting a concession for a competing infrastructure, casting doubt on the viability and safety of the technology, etc. With some imagination, this could be seen as the government’s uncontractible investment within the meaning of the GHM models.

Alignment mechanisms: The first stage of the development of the railway industry contained some notable examples of evolution at work. In several countries under considerations, companies that were not set up adequately were removed – through natural selection – from the population of all companies, before re-emerging, phoenix-like, from their ashes with the shortcomings rectified. In Germany, the entire system of company law was in urgent need of updating, something made clear by the initial failure of the Rhine-Weser railroad. In France

⁶³⁸ Railway Regulation Act, 7&8 Vict., c. 85.

and Britain, not all of the initial endeavours were based on sound financial underpinning, causing many of them to fail.

Competition between countries was observed, for example, in the case of Belgium, where the government was eager to establish an independent connection between Prussia and the world markets, in competition with the Dutch waterways. Otherwise, though, it was not yet much of a factor.⁶³⁹ Competition between modes of transport was wonderfully illustrated by the Liverpool & Manchester as well, a railway which was conceived in order to break the cartel of the three existing canal connections between the two cities. Competition between railways, on the other hand, was not yet possible in this early stage, in the absence of competing railway lines between the same destinations.

Design, finally, appeared in France more than in any of the other countries. Lengthy debate, lasting more than a decade, preceded any actual action. Whether that is a good thing, is a different matter, as is the question of whether they reached the correct conclusion, but at least it points towards a system of design at work, i.e. a system where the theoretically optimal alignment between transaction characteristics and governance structure is more likely to occur than any other possible alignment. Design seems to have played much less of a role in those countries – like Britain and Germany – that opted for a private model: market was the default, the governance structure chosen when no deliberate choice was made.

Concluding, we can say that institutional economics, as such, is an appropriate framework for understanding the problem at hand. The work of Commons and Coase seems to fit well. The variables identified by later authors, though, are less easily recognised. Another problem is that the predictions of the models often do not add up. In the period we have discussed so far, uncertainty was at its zenith; never is information more costly than when one is trying to do something unprecedented. NIE models tend to incorporate uncertainty by pointing out that it will make contracts more incomplete. This, in turn, means that contracts will be relatively disfavoured as a governance structure, meaning that hierarchy will be more sensible. And yet, the countries under consideration overwhelmingly favoured private enterprise. Only some of

⁶³⁹ The railway pioneer Friedrich List may have been a little ahead of his time when he wrote, in 1834: “The French force the Belgians, the Belgians force the Dutch, all of them together force the Germans to build railroads; [and] the Germans (...) force one another to build railroads.” Quoted in Dunlavy (1994), p. 109-110. There is little evidence that decision makers at that time felt any real economic pressure to build railroads, although many did consider it a matter of national pride.

the incarnations of NIE point to a solution: when the government has only few preferences that are costly to contract, using contracts instead of hierarchy can nonetheless be optimal.⁶⁴⁰ On balance, it appears that that is in fact a pretty good way to sum up the reasoning of most of the actors we have seen.

⁶⁴⁰ In Coase (1937), this conclusion follows in a fairly straightforward manner. In Hart, Shleifer & Vishny (1997), the (uncontractible) preferences of the state are captured by the variables $b(e)$ and $\beta(i)$.

3.2 Consolidation

When the railways were first invented, every line was perceived as a separate problem in need of a governance structure. Separate companies were founded to deal with them based on separate enabling acts. However, as we have seen, by the middle of the 1830s, the French and Belgian authorities already started to think about the railway industry as a *réseau*, i.e. a network in need of at least some general organising principles. Once that step had been made, it became less evident that each railway line should be organised as a separate company.⁶⁴¹ Particularly in the wake of railway crises and general economic malaise, different solutions were being explored. The present section follows the early maturity of the railway industry in the remainder of the 19th century, again considering the different countries one at a time.

3.2.1 Great-Britain

The Grand Junction railway between Birmingham and the Liverpool & Manchester was already mentioned earlier in the passing. Authorised in 1833⁶⁴² and opened in 1837, it was the world's first long-distance railroad, running at a distance of 132 km. Around the same time, another railway was built connecting Birmingham to London.⁶⁴³ As of 17 September 1838, it was possible to travel from London to Liverpool, a distance of almost 350 km., entirely by train. However, doing so did require three separate tickets, and transfers at Birmingham's Curzon Street station and at Newton Junction. As noted before⁶⁴⁴, these railways were constructed almost entirely with the intention that they should connect the end points, with little regard for the stations along the way. Even the middle point, the booming city of Birmingham, was only of moderate importance, as the two railways were always conceived as two connected enterprises, that were merely split up in order to reduce the risk of the project and improve its manageability. However, in so doing the promoters created significant inconveniences.

As already mentioned, the two companies did not offer any kind of direct connection between London and Liverpool. While they shared a station building in Birmingham's Curzon

⁶⁴¹ N.B. the French Law of 1842 still envisaged that approach.

⁶⁴² Grand Junction Railway Act (1833), 3&4 Will. 4, c. xxxiv.

⁶⁴³ London & Birmingham Railway Act (1833), 3&4 Will. 4, c. xxxvi.

⁶⁴⁴ Section 3.1.5, above.

Street, they used different platforms, meaning that passengers had to transfer and all goods had to be carried over. As for tickets, the standardised “Edmondson” ticket was not invented until 1837, and did not come into general use until the middle of the next decade.⁶⁴⁵ Before that time, the billets were partly or wholly written by hand, and needed to be reserved well in advance. Of course, such tickets were still only available for the network of the railway company selling the tickets. After all, it would be unthinkable for one company to sell the services of another without a specific agreement to that effect. And given the number of different railway companies that had been incorporated by the early 1840s, the transaction costs of making and enforcing an agreement between every two of them would be enormous.

The solution for this problem was the first real act of consolidation in the British railway industry: The Railway Clearing House.⁶⁴⁶ Founded in 1842 at the initiative of the London & Birmingham⁶⁴⁷, the idea was to apply the methods used by the London banks for cheques cashed at another bank than the bank of the drawer⁶⁴⁸ to railway tickets and material. Every month, the money owed by every company in the system to every other company in the system for use of railroads, rolling stock, or other facilities was calculated, and the net amounts paid off. Initially, only a few companies participated, with the most notable absentee being the Grand Junction railway. After all, one would expect them to appreciate this opportunity for improving the convenience of its connection with the London & Birmingham. As it happened, they joined the Clearing House soon after its success became apparent in the mid-1840s. By 1850, almost all English railway companies had joined, except for the Great Western and the companies operating to the south of the Thames valley. In that year, the initiative was placed on more sound legal footing with the enactment of the Railway Clearing Act⁶⁴⁹, which also improved the ability of the Clearing House to compel payment of the unwilling.⁶⁵⁰

While the benefits of establishing the Railway Clearing House are clear, there are also significant transaction costs. To begin with, there are the costs of this new organisation itself:

⁶⁴⁵ Cf. Hylton (2007), p. 96.

⁶⁴⁶ Cf. Francis (1851), p. 49-59, Head (1861), ch. XIV, Hylton (2007), p. 125, Lardner (1850), ch. IX and Wolmar (2007), p. 103-105.

⁶⁴⁷ The Clearing House was established next door to the headquarters of the London & Birmingham in London’s Euston street, and in its early days the chairman of the London & Birmingham was also the chairman of the Board of the Railway Clearing House.

⁶⁴⁸ A cheque is an order by the drawer of the cheque to the drawer’s bank, instructing the bank to pay a certain amount to the payee, or possibly to bearer. In order to facilitate the cashing of cheques at banks other than the bank of the drawer, the London banks created a clearing house as early as the late 18th century. Cf. Lardner (1850), p. 150-153.

⁶⁴⁹ Railway Clearing Act (1850), 13&14 Vict., c. 33.

⁶⁵⁰ Sections XIV-XXV stating, for example, that the books of the clearing house were *prima facie* evidence of the existence of the debts therein recorded (section XIX).

“In the six months ending 31st December, 1850, the amount balanced was £1,320,000, and the expense of management only £8,700.”⁶⁵¹ However, that amount, while significant, is only the start of the story. The Clearing House cannot work unless it receives adequate information, that is to say, information about all passengers, all goods, and all rolling stock that crosses any boundary between one railway company and another. What is more, that information has to be standardised as much as possible. That is why the creation of the Railway Clearing House created a powerful impetus for the adoption of the Edmondson ticket⁶⁵², which allowed for standardised registration immediately at the moment when the contract with the passenger was concluded. Similarly, the rolling stock had to be as interchangeable as possible, implying more standardisation. That is why the Great Western was reluctant to join the system, despite its obvious advantages: at the behest of its legendary lead engineer, Isambard Kingdom Brunel⁶⁵³, the Great Western used a much wider gauge than most other railway companies: 7 ft. instead of Stephenson’s 4 ft. 8½ inches. As a result, their rolling stock could not use the railroads of other companies, and other companies’ rolling stock could not use the railroads of the Great Western.⁶⁵⁴

While standardisation is useful for any number of forms of cooperation beyond the Clearing House, it always comes at a cost: different circumstances may well require different solutions in different parts of the country. One can easily imagine that the optimal gauge depends on the type of terrain that is to be crossed, and on whether the railroad will be straight or winding. Forcing a single solution on the entire national network means that the fit in certain areas may be highly imperfect.

Another problem created by the Clearing House is a reduction in the level of competition. If such a system is to work, it has to be highly transparent. That means giving each company a lot of information about its competitors. As we know from game theory, the easy detection of cheating is essential for the maintenance of a cartel, be it a formal cartel or an informal concerted practice.⁶⁵⁵ Especially as the number of parallel connections – for example between London and Lancashire – started to increase, the effect of the Railway Clearing House on competition was probably significant. As we will see again later in this chapter, that

⁶⁵¹ Francis (1851), p. 58.

⁶⁵² Cf. Hylton (2007), p. 96.

⁶⁵³ Cf. Hylton (2007), p. 63-69 for a short portrait.

⁶⁵⁴ Cf. Hylton (2007), p. 166-170, about the “gauge wars”. Despite the Gauge Act (1846), 9&10 Vict., c. 57, the standardisation of the gauge in England was only completed in 1892, when all railroads were 4 ft. 8½ inches wide, measured between the rails.

⁶⁵⁵ Cf. Jones & Suffrin (2008), ch. 11, Martin (1993), p. 162-169, Myerson (1991), ch. 7.

is often the case: measures designed to improve the efficiency of the railway sector by making it easier for railway companies to cooperate, also make it easier for them to cooperate anti-competitively.

Another important development that has its origins in the 1840s is the rise of amalgamation. Instead of improving the interaction between railways through contractual means, which is how the Railway Clearing House can be described, it is of course also possible to use different governance structures. This is a problem that can be understood quite intuitively using the tools of NIE. Consider the Liverpool & Birmingham and Grand Junction railways, for example. While they were always envisaged as being part of a single railway connection, they started out as separate companies, connected through a myriad of contractual ties. Those contracts govern transactions regarding the transfer of passengers, goods and rolling stock from one railroad to the other.⁶⁵⁶ Conceptually, the truly interesting thing about these transactions is that they are reciprocal: both companies are each other's "supplier" and "buyer".

Applying the standard Williamson framework to this problem, the question immediately arises: is it possible for a holdup to occur in such a symmetrical situation? After all, anything one company can do to the other, the other can do in return. By their nature, holdups are observable but not verifiable⁶⁵⁷; there is no such thing as a holdup that nobody knows about. It follows that unless the situation is not truly symmetrical, unless one company has some additional exposure that the other can exploit⁶⁵⁸, the two will hold each other by the

⁶⁵⁶ Clearly, this is not about transfer of ownership, but rather transfer of physical control, or – in the case of passengers – not even that. Cf. section 2.3.1, above. Regarding physical control over passengers, it should be noted that while passengers were habitually locked into their carriages in the early years of the railways – in order to stop them leaving the train at unexpected moments and getting run over – this practice was banned after a horrendous train accident near Versailles in May 1842 killed 52 people, many burned to death as a result of being locked in. Cf. Hylton (2007), p. 82, Wolmar (2007), p. 147. However, this rule only required that one door in every carriage be unlocked. Head (1861), p. 45, still describes how "the Company's porters at a brisk walk (...) unfasten one after another the doors of all the carriages" upon the arrival of the train.

⁶⁵⁷ Cf. section 2.1.5, above.

⁶⁵⁸ It is a classic rule of negotiation theory that whichever side "needs" the deal more will be in a difficult position. In terms of Harvard Negotiating Project, this is described as having a poorer "Best Alternative to a Negotiated Agreement" and in game theory, it is called having a poorer "outside option". Cf. Fisher, Ury & Patton (1991), ch. 6 and Schelling (1956), respectively. Despite having an ostensibly symmetrical position, one company might nonetheless be in a weaker position, for example because it has higher leverage and can therefore ill afford a breakdown in service, or because it has a poorer connection to the rest of the network.

throat without either being able to dictate terms. Williamson's TCE does not seem very applicable here.⁶⁵⁹

As for the various GHM models, some of them – particularly Hart & Moore (1990) – assume the presence of significant human capital specificity.⁶⁶⁰ While there is undoubtedly some skill involved in many of the tasks carried out by the various employees of the railway, the human capital involved is undoubtedly insignificant in comparison to the value of the physical capital deployed. Moreover, this human capital – unlike the rails for example – is not specific to the company. Human capital acquired while working for one railway company can be deployed for another without too much difficulty. It follows that the employees of the railways cannot derive bargaining power from their human capital in the manner described by Hart & Moore (1990).

Many of the other GHM models, on the other hand, do ring true when applied to this tale of two railway companies. For example, from the perspective of Hart, Shleifer & Vishny (1997), it is clear that each company will cut costs to the extent that the reduction in expenditure is not outweighed by the resulting reduction in its own revenue. Each company equates the marginal costs and benefits of cost cutting for its own accounts only. However, this clearly ignores the externality that a reduction in quality creates for the other company: overcrowded carriages, carriages with overly Spartan facilities, trains with a tendency to miss their connection in Birmingham, etc., can all be justified – from the perspective of an individual company – by calculating that the expenditure saved in this way is greater than the resulting loss in revenue. But at the same time, they tend to reduce the demand for rail travel for the entire route between London and Lancashire, even if only one of the companies engages in these practices. And since many of these practices cannot be contracted, at least not at a reasonable cost⁶⁶¹, quality will tend to be lower than would be optimal from the perspective of the entire London – Lancashire route.

Finally, the most important benefit of a merger between the London & Birmingham and the Grand Junction, and for all other railway companies similarly situated, is of course located at what we would consider the production side of the equation, not the transaction

⁶⁵⁹ N.B. there is a difference between saying that the Williamson model is inapplicable, and saying that it is applicable but that it shows no reason for a merger between the companies. In this case, the model indicates significant exposures, but does not account for the possibility that these exposures might cancel each other out.

⁶⁶⁰ Cf. section 2.5.3, above.

⁶⁶¹ For comparison, consider how easy it would be for a railway company to convince one of its “suppliers” to use the same gauge, in order to improve customer service. The gauge of the railway is a decision which only has to be taken once, and which can be verified easily.

side.⁶⁶² While there are significant difficulties in supervising so large an organisation as a railway, especially given the means of communication available in the middle of the 19th century⁶⁶³, these difficulties do not increase very much if the company takes on more routes. As one can imagine, the management of a small railroad like the Stockton & Darlington is quite a bit easier than the management of something like the London & Birmingham. Not only is the latter about 4½ times as long, but its management is further hampered by having both termini located in major urban areas. However, adding to the 112 miles of the London & Birmingham the further 78 miles of the Grand Junction railway does not add a great deal to the burden of management. As a result, a merger between the two allows for a significant reduction in overhead.

While this example of economies of scale requires vertical integration by definition, there are other possible economies that could theoretically be generated even between independent companies. An example that we already discussed is the Railway Clearing House. However, many such economies are impossible for the reasons listed by Coase in his original article: the costs of writing a contract to govern such transactions, and the costs of monitoring their execution are simply greater than the amount of money that would be saved. Examples include possible joint buffer capacity for various inputs, such as coal or rolling stock, or joint purchasing of these inputs.

For some of these transactions, a joint venture along the lines of the Railway Clearing House might have been possible, although that type of governance structure was still quite unusual in the mid 19th century, meaning that there would have been a great deal of uncertainty about the best way to arrange such a thing. This uncertainty is reflected in the history of these clearing houses, which started out with the relatively simple problem of cheques, followed several decades later by the railways, with the latter following extremely closely in the footsteps of its predecessor. Absent any greater familiarity with various hybrids, the only realistic options were independence or amalgamation.

⁶⁶² N.B. Unlike many other authors, the present thesis works from a definition of transaction which is limited to transfers between legally distinct entities. Cf. section 2.3, above. All other costs, including the inefficiencies created by the inability of management to adequately supervise the entire organisation, are considered “production costs”.

⁶⁶³ The history of the rise of the Electric Telegraph is closely connected to the rise of the railways, particularly the Great Western and the London & North Western. These organisations, more than any other at the time, benefited from this new opportunity for careful coordination over large distances. Cf. Head (1861), ch. XIII.

For all of these reasons and more the Grand Junction merged with the Liverpool & Manchester in 1845⁶⁶⁴, and in July 1846 with the London & Birmingham and Manchester & Birmingham railways.⁶⁶⁵ Henceforth, the largest corporation in the world was the London & North Western railroad.⁶⁶⁶ Around the same time, further north and east, George Hudson was in the process of setting up his railway empire.⁶⁶⁷ While the details of the story of the rise and fall of George Hudson are outside the scope of this thesis – being more appropriately dealt with by a forensic accountant – the result was the amalgamation of scores of small railway companies into the York and North Midland. After the collapse of Hudson’s empire in 1849, this company became the foundation of the North Eastern Railway in 1854⁶⁶⁸, while other parts of his empire were amalgamated into the Midland Railway.

By the 1860s, the British railway industry had consolidated around a number of giants, supported by myriad smaller companies. While the three companies already mentioned had been formed as a conscious effort to consolidate independent companies so as to improve efficiency, many of the others reached their position by starting from an expensive and important trunk line and expanding from there. The Great Western started from a railway between London and the West Country, initially as far as the important port city of Bristol, and later extending all the way to Plymouth and Penzance. From this basis, the company extended to Southampton and Weymouth on the south coast of England, and north as far as Birmingham and Wolverhampton. Similarly, the Great Northern⁶⁶⁹ started out as a railway between London and York, which continued to be its most important route throughout its existence, being the first part of the East Coast main line.⁶⁷⁰ To the south of the Thames valley, there was less

⁶⁶⁴ Grand Junction Railway Act (1845), 8&9 Vict., c. cxcviii. Cf. Hylton (2007), p. 47.

⁶⁶⁵ London & North Western Railway Act (1846), 9&10 Vict., c. cciv. Cf. Hylton (2007), p. 50.

⁶⁶⁶ Cf. Gordon (1910), ch. 1.

⁶⁶⁷ Hylton (2007) has a short biography of Hudson on p. 57-63. Cf. also Francis (1851), ch. VIII, Tomlinson (1915), ch. XIII and XIV and Wolmar (2007), p. 97-103.

⁶⁶⁸ After 1854, the North Eastern Railway continued to expand in Yorkshire, County Durham and Northumberland, acquiring, for example, the Stockton & Darlington in 1863. Cf. Gordon (1910), ch. 3 and Tomlinson (1915), ch. XVII.

⁶⁶⁹ Cf. Gordon (1910), ch. 2.

⁶⁷⁰ Throughout the second half of the 19th century, there were two main routes that connected London to Scotland: The East Coast Main line from London to York using the Great Northern, and from there to Berwick and Edinburgh with the North Eastern, and the West Coast Main line from London to Lancashire using the London & North Western, and from there with the Caledonian Railway to Glasgow and Edinburgh. The rivalry between these routes was at its peak during the famous races to Edinburgh and Aberdeen in 1888 and 1895. Cf. Gordon (1910), p. 60-66 and Wolmar (2007), ch. 9. After 1876, there was a third connection, using the Midland from St. Pancras station, and connecting with its subsidiary, the Settle & Carlisle. From Carlisle, the rest of the way could be covered using the Caledonian or the Glasgow & South Western. However, due to the rough terrain covered by the Settle & Carlisle, this third route was always significantly slower than the others.

space for major railways. The major players there, other than the Great Western, were the London & South Western, which competed with the Great Western on the route to Southampton as well as connecting to a number of other destinations in Hampshire and West Sussex, and the London, Brighton & South Coast (LBSCR) and South Eastern Railways, which competed in the ever increasing commuter market in Sussex.

These last two railways are in many respects more like the largest of the small companies, rather than the smallest of the giants. Both suffered significant financial difficulties, due to mismanagement and lack of scale, despite running potentially lucrative routes. The South Eastern's routes to the Channel ports of Dover and Folkestone were also reached by its smaller and equally impecunious brother the London, Chatham & Dover, until the two finally merged in 1899. A similarly vicious competition between the South Eastern and the LBSCR doomed any chance of financial solvency for the main Sussex routes.

Such stories can be told about most of the smaller railway companies that existed in Britain until the amalgamation following World War I. They constantly faced the dual threat of being outcompeted on the railways and being outcompeted in Parliament. Tomlinson's history of the North Eastern Railway, for example, is one long account of scheming and plotting of one company against another, allegiances changing constantly in a way that Machiavelli would have instantly recognised.⁶⁷¹ Many of the strategies used sound suspiciously like hold-ups the way the term was used in the previous chapter. In 1860, for example, the London & North Western plotted with a number of other companies to improve the connection between Lancashire and the North East, so that the position of the North Eastern in the coal market would be undermined. In response, the North Eastern reached out to the Stockton & Darlington, a decision leading ultimately to a merger between the two in 1863. They also persuaded the Newcastle & Carlisle Company to support them. Tomlinson (1915) explains what happened next:

After the Stockton and Darlington defection the case of the Newcastle and Derwent Valley party necessarily collapsed. They could not show, as they proposed, an independent access into Lancashire, because the Stockton and Darlington and South Durham and Lancashire Union lines were closed against them; they could not show and independent access into Newcastle, because they had to go for three miles over the rails of the Newcastle and Carlisle Company who were hostile to them, and further,

⁶⁷¹ Tomlinson (1915). The chapter that discusses the height of this manoeuvring in the period of 1857-1865, chapter XVII, is appropriately called "Schemes of Invasion that were Foiled".

they could show no access to any shipping place on the River Tyne. The preamble of their Bill was therefore declared not proved.⁶⁷²

It is impossible to say on this evidence whether this practice of strategically closing lines to rival companies constituted a hold-up, although we do know that “objections had been taken, both in a Court of Law and in Parliament, to the arrangement for closing the Newcastle and Carlisle line to traffic between Newcastle and Liverpool”.⁶⁷³ This certainly suggests that the Newcastle & Carlisle acted in less than entirely good faith in the execution of their contractual and other legal obligations, in order to acquire a larger share of the gains from trade than they were originally supposed to get. It also serves to show the benefit of having laws against abuse of market power⁶⁷⁴ and unlawful monopolization⁶⁷⁵, which, of course, did not exist during the period in question, the first modern antitrust law, US Sherman Act, not having been enacted until 1890.⁶⁷⁶

This last example illustrates perfectly how the British railway industry functioned throughout the second half of the 19th century. The government was involved, but only by the setting of safety rules and by enacting the enabling acts. As a result, a lively competition emerged, both in the running of the railways and in parliament. As we have seen, the way these companies interacted can be understood quite well using traditional micro-economic tools such as neo-classic economics, game theory, and the models discussed in chapter 2. However, for present purposes the interesting question is not primarily how to understand the governance of transactions between companies, but rather how to look at the governance of the transactions between the companies and the state. A study of the former question would probably result in the conclusion that the market was quite inefficient.⁶⁷⁷ The subsequent application of Williamson’s remedialness criterion⁶⁷⁸, however, would suggest that there is no clear way in which the companies themselves could have improved on this outcome.

Studying the latter question, about the efficiency of the transactions between the state and the companies, is much trickier. After all, one of the ingredients in such a study is the

⁶⁷² Tomlinson (1915), p. 587.

⁶⁷³ Tomlinson (1915), p. 593.

⁶⁷⁴ Art. 82 EC, cf. section 18 of the Competition Act 1998 (c. 41).

⁶⁷⁵ 15 USC 1, cf. *Aspen Skiing Co. v. Aspen Highlands Skiing Corp.*, Supreme Court of the United States, 19 June 1985, 472 U.S. 585.

⁶⁷⁶ In domestic British law, no explicit ban against abuse of market power existed until the enactment of the Competition Act 1980 (c. 21).

⁶⁷⁷ For statistical evidence, cf. Crafts, Mills & Mulatu (2007).

⁶⁷⁸ Cf. Williamson (1991).

objective function of the state: What is it that the state wants to get out of its relationship with the railways? Looking at the historical sources, it is clear that initially the state was mostly concerned with minimising the intrusion by the railways on established property rights, including the right of peaceful enjoyment of one's property. In other words, they wanted to avoid the use of eminent domain power as much as possible, and they wanted to protect the property owners in the vicinity of the railways against all the disasters the railways were alleged to cause.

As early as Gladstone's tenure as President of the Board of Trade (1843-1844), this started to change. His efforts during this period, some of which ended up in the Railway Regulation Act of 1844⁶⁷⁹, clearly showed a concern for the overall efficiency of the industry, as well as for the safety of both passengers and staff.⁶⁸⁰ However, in this early period these concerns received little weight compared to the more traditional values of freedom to contract and free enterprise. While the governance structure initially chosen seems to have provided a good fit, this was increasingly less the case as the government's priorities changed. Nevertheless, during the rest of the 19th century the fit was never so poor that the government's policy of general regulation and encouragement of amalgamation was under any kind of serious political attack.

3.2.2 The Netherlands

In section 3.1.4, above, we discussed the history of the HSM, connecting Amsterdam and Rotterdam via Haarlem and The Hague and the NRS, connecting Amsterdam and Rotterdam with the Prussian border near Arnhem.⁶⁸¹ What was clear about this account was that the Dutch were committed to the "private option" almost as much as the British were, and more than any of the other countries in our sample. The result was an entirely private *Hollandsche IJzeren Spoorwegmaatschappij*, and a curiously hybrid but ultimately private Rhine Railroad.

During the subsequent period, the Dutch railway network expanded only slowly. As before, very few people seemed to consider it a worthwhile endeavour, given the existing network of rivers and canals. Two foreign railroad companies crossed into the Netherlands in

⁶⁷⁹ 7&8 Vict, c. 85.

⁶⁸⁰ The death of William Huskisson, the local MP and former President of the Board of Trade, during the opening of the Liverpool & Manchester in 1825 does not seem to have immediately encouraged stricter safety rules. Cf. Hylton (2007), p. 46-47.

⁶⁸¹ Cf. figure 5, above.

the 1850s: the Antwerp & Rotterdam, which extended as far as Moerdijk⁶⁸² and the Aachen & Maastricht, which connected the Flemish city of Hasselt with Aachen in the Prussian Rhine Province by way of Maastricht in the Netherlands.⁶⁸³ Including those, the Netherlands still only had 4 railroads by 1860, 35 years after the opening of the Stockton & Darlington. In kilometres, the network in that year was 5½ times shorter than that of lowly Belgium.⁶⁸⁴

Considering the arguments advanced by the government in the 1840s, this could mean one of two things: Either there were no other routes available where a railroad could be built profitably, or the private sector was not working as it should. In the 1850s and 60s, public opinion tended more and more towards the latter. However, the earlier concerns about state involvement in the private economy had by no means disappeared.⁶⁸⁵ So the government looked for a way that they might have their cake and eat it, too. The result was a hybrid almost as curious as the NRS in its early days⁶⁸⁶, a hybrid in many ways similar to the system France would have had if the Act of 1842 had been carried out as contemplated.⁶⁸⁷

Under the system created by the Railway Acts of 1860 and 1863⁶⁸⁸, the state would go ahead and build the railroads itself. However, this was meant strictly as a temporary fix: as soon as they were completed, the railroads were to be spun off into a private company, the confusingly named *Maatschappij ter Exploitatie der Staatsspoorwegen*, or Company for the Operation of the State Railroads. In addition to the HSM and the NRS, the Netherlands now had a third major railway company: the SS, responsible for more than 800 new kilometres of railroad, including several new bridges. The design of that organisation of course bore all the signs of a hefty political struggle: if the government was willing to abandon the argument of the 1840s that the state should not be in the business of building railroads because they would

⁶⁸² Société Anonyme des chemins de fer d'Anvers à Rotterdam. Cf. Jonckers Nieboer, (1907), p. 77-79 and Veenendaal (2004), p. 94-96. At Moerdijk, there are currently two bridges across the Hollands Diep. However, given the width of the river at that point – about 2,5 km – the first bridge on that location was only opened in 1872, being the longest bridge in Europe at the time. Cf. Veenendaal (2004), p. 109-110.

⁶⁸³ This railroad was relatively small by nature. It was not, as of yet, a real “Iron Rhine”. Cf. Jonckers Nieboer, (1907), p. 74-77 and Veenendaal (2004), p. 96-97.

⁶⁸⁴ Cf. Veenendaal (2004), p. 99.

⁶⁸⁵ For example, a 1857 bill proposing that the state spend 10 million guilders to support in various ways yet to be specified various companies that were yet to be founded was rejected by Parliament. The next year, a new government made a new proposal, where particularly the idea of state participation in the equity of the railways encountered much resistance. Ultimately, the amended version of this second proposal was rejected by the Senate in February 1860. Cf. Veenendaal (2004), p. 100-101.

⁶⁸⁶ Cf. section 3.1.4, above.

⁶⁸⁷ Cf. section 3.1.5, above.

⁶⁸⁸ Wet tot aanleg van spoorwegen voor rekening van de staat, d.d. 18 August 1860, Stb. 1860, 45 and the Wet van 3 Juli 1863, Stb. 1863, 100.

only build them in the wrong places⁶⁸⁹, why not keep the company in public hands? This is what happened in the Dutch East-Indies, where a similar lack of enthusiasm resulted not only in construction by the state, but also in state operation.⁶⁹⁰

The initial act had said as little as possible on the subject. Article 4 simply noted that the manner of operation was to be determined by law.⁶⁹¹ In the mean time, however, the conservative government of Prime Minister Van Hall had been replaced by the liberal J.R. Thorbecke, who had been against the railway act from the start.⁶⁹² While he did not want to reverse the course adopted under his predecessor, he forcefully argued for privatisation for reasons of principle as well as efficiency. Reitsma (1940) quotes him as saying that he would even favour privatisation of it were certain to be more expensive.⁶⁹³ As for efficiency, he argued that privatisation would produce more innovation⁶⁹⁴, both in the direction of higher quality and lower costs, and that it would be inappropriate to have operation and supervision in the hands of the same organisation.⁶⁹⁵ Even in the absence of a party system, much less party discipline, the government was able to convince a majority in parliament, which is why the Railway Act of 1863 ordered the creation of the SS as a private railway company.⁶⁹⁶

While chaotic, the enactment of the Acts of 1860 and 1863 had the desired effect. Not only did the Netherlands end up with two new railway networks – known as the Northern and the Southern lines, depending on which side of the NRS they were on – but the railway industry in its entirety received a new impulse. The existing railway companies started expanding, albeit initially only on railroads constructed as part of the Act of 1860⁶⁹⁷ and the country was even blessed with a new railway. The *Nederlandsche Centraal Spoorweg Maatschappij* or Central Railway already received its concession in 1856, but was only able to raise the necessary funds in the early 1860s. In 1863, the NCS railroad connecting Utrecht, Amersfoort and Zwolle, in the heart of the Netherlands, was opened. Unfortunately, due to its modest size

⁶⁸⁹ Note how the law of 1860 carefully mapped out the railroads over all electoral districts. Cf. Jonckers Nieboer, (1907), p. 99 and Veenendaal (2004), p. 102.

⁶⁹⁰ Cf. Reitsma (1940), p. 763.

⁶⁹¹ Cf. Reitsma (1940), p. 766.

⁶⁹² Some of his more forceful statements are quoted by Reitsma (1940), p. 767-768.

⁶⁹³ Reitsma (1940), p. 767-768.

⁶⁹⁴ “Het eigenbelang is de sterkste hefboom tot verbetering van den dienst met de minste kosten en tot vermeerdering van de opbrengst.” Reitsma (1940), p. 767.

⁶⁹⁵ Reitsma (1940), p. 767.

⁶⁹⁶ The Company was officially founded on September 26, 1863, though the concession was already agreed on August 11, 1863. Cf. Jonckers Nieboer (1907), p. 111-113.

⁶⁹⁷ The HSM leased the line between Amsterdam and Den Helder, Wet van 10 Augustus 1865, Stb. 1865, 99. Cf. Van Kerkwijk (1878, 1886) and Veenendaal (2004), p. 125. Beyond that, the companies were reluctant – asking for concessions but not necessarily following through. Cf. Veenendaal (2004), p. 140-141.

and precarious financial situation, it continued to be a minor player throughout its history, until it was partially taken over by the NRS in the 1880s⁶⁹⁸ and completely acquired by the SS in 1890.⁶⁹⁹

That last acquisition was part of the second major intervention by the government in the railway industry. A major parliamentary inquiry in 1881-1882 showed that the existing system left much to be desired.⁷⁰⁰ Insofar as is relevant here, the committee of inquiry concluded that parliament should exercise greater control over the concession procedure, so that the interests of the network as a whole might receive more attention.⁷⁰¹ Moreover, more attention should be given to the joint use of stations, for the convenience of travellers⁷⁰², and to the running powers of one company on the railroads of the others, which should preferably be made automatic.⁷⁰³ Finally, the committee recommended that the financial health of the railway companies should be scrutinised more carefully, with the weaklings being acquired by their healthier brethren.⁷⁰⁴

After a decade of debate, during which time only little actual progress was made⁷⁰⁵, the government went beyond the 1882 recommendations on several points. Both for reasons of financial health, and for reasons of systemic efficiency, it was decided that not only the financially unhealthy companies should be redistributed, but also the relatively healthy NRS.⁷⁰⁶ That way, the two remaining railway companies, the HSM and the SS, could be placed on the healthiest of financial footings, while at the same time making sure that they would each have a connection to each of the major destinations, for maximum competition. For example, the SS connected Amsterdam and Rotterdam with Germany by way of Utrecht and Arnhem, following the old core line of the NRS. The HSM offered the same connection from Rotterdam next to the river Waal via Gorinchem and Geldermalsen to Nijmegen, and from Amsterdam close to the southern shore of the Zuiderzee, via Amersfoort and Apeldoorn

⁶⁹⁸ Cf. Veenendaal (2004), p. 207.

⁶⁹⁹ Cf. Veenendaal (2004), p. 224.

⁷⁰⁰ Cf. Veenendaal (2004), p. 201-202 and Van Wickevoort Crommelin (1883).

⁷⁰¹ De Economist (1882), p. 1107-1108.

⁷⁰² De Economist (1882), p. 1108-1109.

⁷⁰³ De Economist (1882), p. 1109-1110.

⁷⁰⁴ De Economist (1882), p. 1118. In this part of their rapport, they did not only have the NCS in mind, but also the Noordbrabantsch-Duitsche and the Nederlandsche Zuidooster-spoorweg, which ran between Boxtel and the German border, and between Tilburg, Den Bosch and Nijmegen, respectively.

⁷⁰⁵ Cf. Reitsma (1940), p. 769-770, and Veenendaal (2004), chapter 10 for a discussion of the debate during this period.

⁷⁰⁶ The NCS and the NBDS continued to exist as separate entities, albeit as subsidiaries of the SS, while the NRS was purchased by the state – compulsory purchase under the concession being impossible until 1898 – and divided between the HSM and the SS. To complete the picture, it should also be noted that a few SS lines were transferred to the HSM. Cf. Jonckers Nieboer (1907), chapter IV and Veenendaal (2004), p. 213.

to the border at Winterswijk and Enschede.⁷⁰⁷ Likewise, the HSM connected Amsterdam with Leeuwarden by going straight north from the capital, going by ship from Enkhuizen to Stavoren, and from there again by train, while the SS reached the same destination by using the eastern shore of the Sea, via Utrecht and Zwolle.⁷⁰⁸ In this way, the HSM and the SS were on roughly equal footing for all major destinations, except for the connection with the far south, where the SS had a monopoly.

As a result of all these reforms, the Netherlands ended the 19th century with a system of managed competition, or, as the government summarised its policy, a system of competition and concentration.⁷⁰⁹ Before we consider further whether we can use Coasian NIE to understand this choice, we should first discuss the situation in France, which – as we have seen – started out with completely different choices, but ended up with a system much like the Dutch one.⁷¹⁰

3.2.3 France

Earlier in this chapter, we considered the history of the railways in French up to the enactment of the Act of 1842, which ostensibly created a system of state constructed railways, but ended up giving a critical impulse to the private sector instead. Once the new republican government had constructed a proper railway network, the private sector was more than willing to operate it. Some of the major companies, that would continue to dominate the French railway industry until nationalisation, were founded shortly after the enactment of the Act of 1842. The most important ones among these were the *Compagnie d'Orléans à Bordeaux* (October 1844), the *Compagnie du Nord* (July 1845) and the *Compagnie de Paris à Strasbourg* (July 1845).⁷¹¹ By the early 1850s, the network planned by the law of 1842 had been divided up amongst all comers, in bits and pieces, organised along the lines of the SS in the Netherlands, i.e. with state construction and private exploitation.

The consequence of all this enthusiasm was, of course, a railway industry that was much less rationally organised than the lawmakers had imagined in 1842. Just like in Eng-

⁷⁰⁷ Veenendaal (2004) has maps of the situation in 1860, 1889, 1914 and 1994 on the inside of the cover.

⁷⁰⁸ Before the redistribution of 1890, the SS used to run the route from Stavoren to Leeuwarden, giving it ample opportunity to make life difficult for the HSM on its route north, thus favouring its own easterly option. Cf. Veenendaal (2004), p. 215-217.

⁷⁰⁹ Cf. Reitsma (1940), p. 770.

⁷¹⁰ Cf. De Economist (1874), p. 255-256.

⁷¹¹ Cf. Caron (2005b), p. 18.

land, there were many companies, large and small, some of them playing fair, others less so, some of them solvent, others less so. The latter problem came to light as a result of increasing speculation in the second half of the 1840s combined with an economic crisis in 1846-47⁷¹², and resulted in the bankruptcy of a number of major companies, including several that had obtained concessions for parts of the core network. The Paris & Lyon, for example, ceased operations in August 1848⁷¹³, while the Lyon & Avignon did not even last that long: it went bankrupt in October 1847.⁷¹⁴ It should be obvious what those difficulties did to the prospects of the Avignon & Marseille, which was opened in September 1847.⁷¹⁵ In this way, the young republic⁷¹⁶ was immediately forced to take important decisions concerning the future of an obviously essential industry, decisions that it did not have the institutional capacity to make any more than the monarchies that preceded it. Given the chaos of the times, it is surprising how well these decisions turned out. Where essential railway companies failed, the state invoked its authority under the concession to buy back the railroads already constructed. These would then be constructed further under the direct supervision of the government, which simultaneously tried to organise a new private enterprise to operate the railway. The Paris & Lyon, for example, was re-privatised in 1852, even though the line was still not completely finished by that time.⁷¹⁷

Under the second empire⁷¹⁸, this policy of privatisation and consolidation continued. This was not yet consolidation for the sake of reducing transaction costs between smaller companies, as we have seen in the UK and in the Netherlands, but mostly consolidation so that financially weak companies might be held upright by their healthier brothers. However, as the 1850s progressed, and the financial health of the country and the sector improved, attention shifted back to the goal of creating a rational railway system, as intended by the law of 1842.⁷¹⁹ This process culminated in the grand agreements of 1859.⁷²⁰

⁷¹² Cf. Caron (1997), p. 191-197.

⁷¹³ Cf. Caron (1997), p. 201 and Caron (2005b), p. 47.

⁷¹⁴ Cf. Caron (1997), p. 180.

⁷¹⁵ Cf. Caron (1997), p. 195.

⁷¹⁶ The second republic was proclaimed on February 24, 1848. Cf. Fyffe (1895), chapter XIX.

⁷¹⁷ Decree of June 5, 1852. Around the same time, the Lyon & Avignon was merged with the still-private Avignon & Marseille and a few other railway companies in the Provence to form the Lyon & Méditerranée. Cf. Caron (2005b), p. 21 and 48.

⁷¹⁸ On December 2, 1851, Louis-Napoléon Bonaparte, the president of the Republic, carried out a coup d'état, and exactly one year later the second empire was officially established under the emperor Napoléon III. Cf. Fyffe (1895), chapter XX.

⁷¹⁹ Cf. Caron (2005b), p. 20-22.

⁷²⁰ Technically, they were negotiated in 1857-1859, but they were only ratified by parliament in July of 1859. Cf. Caron (1997), p. 235-238 and Caron (2005b), p. 22-31.

The purpose of those agreements⁷²¹ between the state and the railway companies was to re-distribute the leases so as to create six major railway companies, each dominant in a distinct geographical area. These companies would be financially strong, capable of offering transport services without unnecessary transaction costs, and would be entrusted with the responsibility for the rapid development of France's railway industry, which was still expanding more slowly than the emperor liked to see.⁷²² This way of distributing the market is similar to what each of the major English companies was trying to do around its own main line, something that failed due to the hostility of the government against such an outcome, and because the companies were never able to arrange a stable system of multi-market collusion.⁷²³ It is notably different from the Dutch reform of 1890, which attempted to maintain redundancy whenever possible: they wanted travellers to be able to choose which company to use for as many destinations as possible.

What the French did not do, however, is divide the entire market between these giants; there were still many smaller railway companies, just not on the core network. Financial support was given not in the form of subsidies or equity participation but in the form of government guarantees as to return on investment.⁷²⁴ Using those guarantees, the companies were expected to build a *nouveau reseau* to complement the network already built under the law of 1842. These new lines were to be built particularly in more sparsely populated regions such as Brittany or the *Midi-Pyrénées*, or in areas that were important from a military point of view, such as Alsace-Lorraine. The Cartesian spirit of the French approach to regulation was, at least for a while, restrained by the liberal instincts of the time, and particularly of *la haute banque parisienne*. Still, it was a very French version of private enterprise, with volumes of careful regulation. The financial guarantees, for example, were felt to require ample govern-

⁷²¹ Note that the French government literally had to engage in protracted negotiation to achieve agreement on its plans. Not even the imperial government of Napoléon III could push these reforms through without the approval of the major railway companies. This was testimony not only to the difficulty of his intended reform, but also to the bargaining power of the oligopolistic companies.

⁷²² Cf. Caron (2005b), p. 21.

⁷²³ Cf. Chadwick (1859), who compares France favourably to the English railways of his time.

Multi-market collusion refers to a result in micro-economics whereby a collusive outcome can be maintained in an oligopolistic market where each of the players has a "home market", which he can defend by threatening reciprocal entry against any of the other players who might otherwise "invade". Cf. Van Witteloostuijn & Van Wegberg (1992) and Van Wegberg & Van Witteloostuijn (1992). It has been suggested that this may be a particular problem in the Common Market, where the former national monopolists may not take up the invitation to invade each other's home markets. Cf. Van Wegberg, Van Witteloostuijn & Roscam Abbing (1994).

⁷²⁴ Cf. Caron (1997), p. 236 and De Economist (1874). The latter gives the detailed financial outlays made by the government under this agreement until 1872, putting the total at 279 million francs. (p. 251)

ment access to the accounts of the railway companies, which, in turn, required a degree of standardisation in their preparation.⁷²⁵

The six giants that emerged were the following⁷²⁶:

- The *Compagnie de l'Est*⁷²⁷, i.e. the Company of the East, connecting Paris with the German border at Strasbourg. Created in 1853 through a merger between the Paris & Strasbourg and the Montereau & Troyes. Until 1871, this company also operated the railways in Alsace-Lorraine. Total length: 1836 km.
- The PLM, i.e. the *Paris-Lyon-Méditerranée*.⁷²⁸ As already noted, there were some financial difficulties in this part of the network, but in 1857 the Paris & Lyon and the Lyon & Méditerranée merged to form the PLM, which operated the route running from Paris to Dijon, and from there via Lyon to Marseilles. Total length: 4080 km.
- The PO, i.e. the *Paris-Orléans*⁷²⁹, which, however, extended far further than Orléans. Not only did it successively buy the railways connecting Paris with Bordeaux and the rest of the South-West, merging with the Orléans & Bordeaux in 1852, for example, but in 1857 it merged with the *Grand Central*, allowing it to extend further inland into the *Massif Central*. Total length: 3976 km.
- The *Compagnie de l'Ouest*⁷³⁰, i.e. the Company of the West, was formed in June 1855 through the merger of six companies connecting Paris with Normandy. Subsequently, and with significant government encouragement and occasional financial support, the company also extended into Brittany. Total length: 2079 km.
- The *Compagnie du Midi*⁷³¹, i.e. the Company of the Centre, was formed in 1852 and given the concession for the connection between the Atlantic and the Mediterranean. Total length: 1530 km.

⁷²⁵ Cf. Caron (1997), p. 237.

⁷²⁶ The length of railway in concession that is given for each company concerns the situation in 1859 and is taken from De Economist (1874), p. 245. The total length in concession for all other companies at that time was 1011 km, for a total of 16.071 km.

⁷²⁷ Cf. Caron (2005b), p. 43-47.

⁷²⁸ Cf. Caron (2005b), p. 47-53.

⁷²⁹ Cf. Caron (2005b), p. 54-57.

⁷³⁰ Cf. Caron (2005b), p. 57-65.

⁷³¹ Cf. Caron (2005b), p. 65-70.

- The *Compagnie du Nord*⁷³², i.e. the Company of the North, finally, was formed in 1845, as we have already seen. It operated the lines connecting Paris with the Belgian network and with Calais and Dunkerque. Total length: 1559 km.

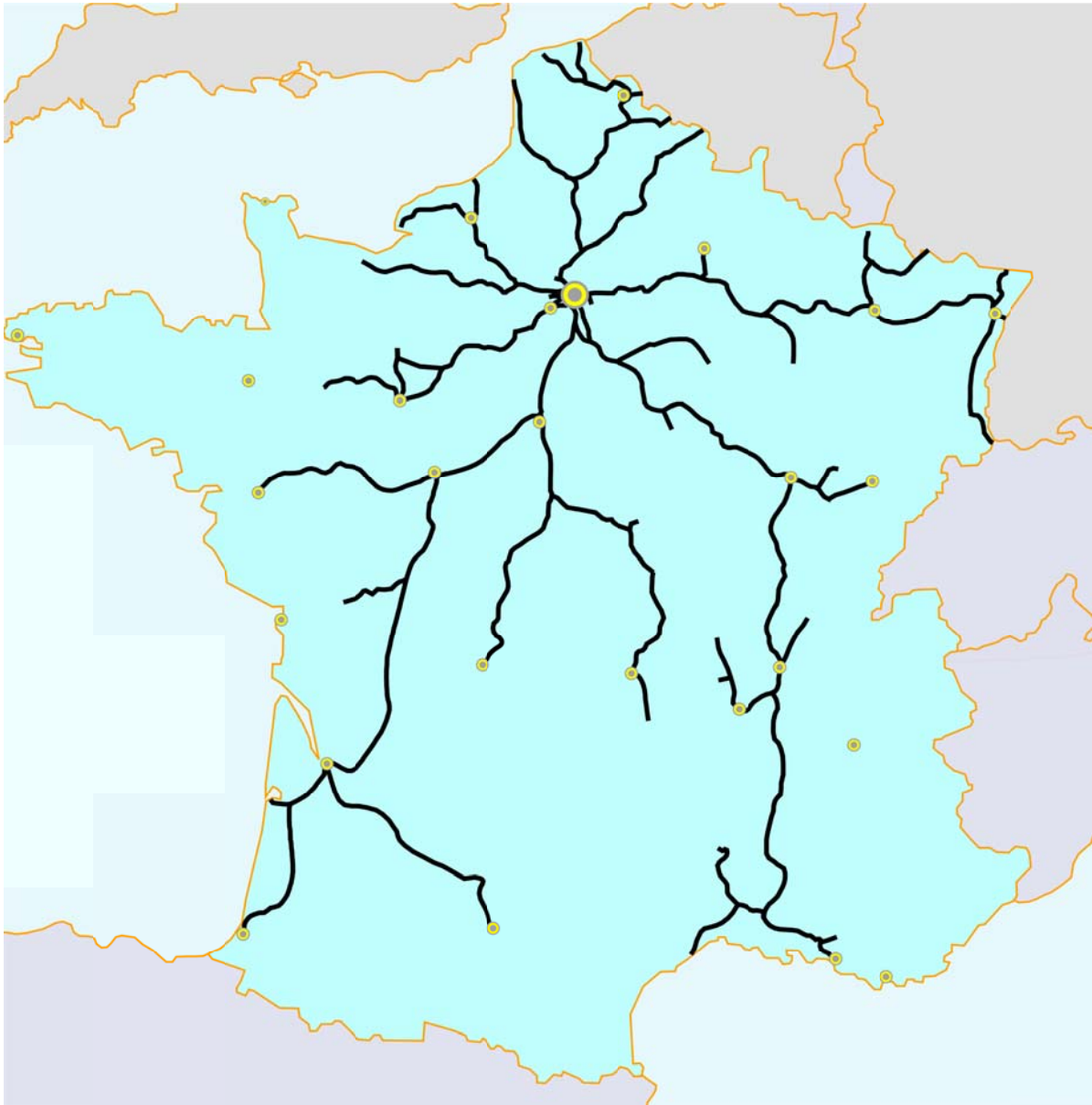


Figure 8: The French Railway system in 1856⁷³³

During the following decade, the expansion of the railways continued at a steady pace. The 13,5% average annual growth in the length of the network between 1851 and 1858⁷³⁴ was

⁷³² Cf. Caron (2005b), p. 70-88.

⁷³³ Adapted from Caron (1997), fig. 11.

never going to continue, but between 1858 and 1875 the industry still managed an impressive 5% average annual growth.⁷³⁵ Unfortunately for the emperor, by the time the war of 1870 arrived, the design of the network still showed the signs of its beginnings: Of the great six, only the Midi did not reach Paris. As figure 8 shows, as late as 1856 there was no connection between any two of the other five, except in Paris. By 1870, there were a few such connections, but not enough to enable the rest of the army to quickly come to the emperor's aid as he was suffering defeat and capture in Sedan.⁷³⁶ As we shall see, the Prussian railway network was much better suited for the rapid deployment of a large army.⁷³⁷

The most important immediate consequence of the fall of the empire was a weakening of the government in Paris. The Third Republic was organised as a parliamentary democracy, at times the very opposite of the decisiveness of the Second Empire. Add to that the difficulties caused by the loss of Alsace-Lorraine⁷³⁸, and it is understandable that railway reform was not an immediate priority. Nevertheless, the military lessons of the defeat of 1870 and the economic lessons of the operation of the railway sector under Napoléon III were slowly being drawn. For the first time since the early 1840s, parliamentarians were again debating whether the state should prefer concentration or not, a discussion that was of course difficult for the smaller companies to win, given their disadvantages in lobbying.⁷³⁹ Also, their cause was not helped by the constant financial difficulties of the poster child of a *chemin de fer d'intérêt local*, the *Compagnie des Charentes*.⁷⁴⁰ Founded in 1863 by local businessmen in the Charentes region in western France, the company was in financial trouble throughout the 1870s. A plan for it to be taken over by the PO was rejected by the *Assemblée Nationale* in 1877, resulting in its ultimate bankruptcy and nationalisation in 1878. Other small companies had similar difficulties. The debate as a whole was won effortlessly by the great companies: there would be no deconcentration, no preference for smaller companies for future concessions, and no reduction in government support for the industry.

⁷³⁴ From 3554 km to 8670 km. Cf. Kalmbacher (1998), p. 19.

⁷³⁵ From 8670 km to 19746 km. Cf. Kalmbacher (1998), p. 19.

De Economist (1874), p. 249, gives the total length per company in 1873 as follows: Nord: 1989 km, Est: 2632 km, Ouest: 2913 km, PO: 4350 km, PLM: 6243 km, Midi: 2568 km, others: 2665 km. Total: 23.360 km.

⁷³⁶ The Prussians crossed the border on July 19th, and the battle of Sedan took place on August 31 and September 1, 1870.

⁷³⁷ Cf. Caron (1997), p. 419-426.

⁷³⁸ In a law dated June 17, 1873, the *Compagnie de l'Est* was compensated for the loss of its Alsace-Lorraine network with 840 km worth of new concessions for railroads yet to be constructed in the neighbouring regions. Cf. Caron (1997), p. 462, De Economist (1874), p. 247-248 and Kalmbacher (1998), p. 19.

⁷³⁹ Cf. Caron (1997), p. 442-443.

⁷⁴⁰ Cf. Caron (1997), p. 443-445. By 1876, the company held the concession for 738 km of railway, 498 km of which was in operation.

Instead, the government and the great six entered into a renewed round of negotiation, resulting in a second major agreement⁷⁴¹: the agreement of 1883. While in 1859, the main goal was to guarantee the continued solvency of the industry, enabling it to expand from a sound financial basis, by 1883 this high level of solvency had created an industry that was satisfied with the network as it was, loathe to risk renewed financial trouble by expanding. Having rejected the possibility of deconsolidating the market, the only other way to end this complacency was to increase government control over the market. The Freycinet plan of 1879⁷⁴² committed the state to the construction and possible exploitation of thousands of kilometres of new railroad, costing an estimated total of 3,5 billion francs.⁷⁴³ Exactly how this was meant to be financed, and who would be responsible for their day to day exploitation was not considered with any great care. When both those questions caused difficulties in subsequent years, the state was forced back to the negotiating table, hoping to bring the big six on board to the plan. Needless to say, this hardly improved the state's bargaining position.

After a few months of negotiation, terms were agreed whereby the great six would operate almost 10.000 km of new railroad, to be completed in the next decade, at an appropriate price and with appropriate guarantees.⁷⁴⁴ Combined with the occasional reallocation of leases between the state and the companies⁷⁴⁵, or between companies, as well as the occasional authorisation for a merger, all of which were added to the agreement to "sweeten the deal", this meant that the great six were even more dominant after 1883 than they had been before.

This set the tone for the period until the 1930s: whenever the state wanted to give an impulse to the railways, they needed private companies to operate the new lines. The only companies that were financially sound enough to do this were the great six, who bargained hard for the privilege. The only limit to the ability of the great six to bargain with the state in this way was the increasing ability and willingness of the state to operate the railways itself.⁷⁴⁶

⁷⁴¹ The modifier "major" reflects the fact that there were a number of smaller agreements between 1859 and 1883, most notably those of 1863, 1868 and 1875. Cf. Caron (2005b), p. 31-35 and De Economist (1874), p. 246-247.

⁷⁴² Law of 17 July 1879, cf. Caron (2005b), p. 38-39, Kalmbacher (1998), p. 21 and Lamming (1998).

⁷⁴³ Initially, the plan concerned 4500 km of railroad, but the project became significantly larger once the members of parliament, elected in a district system, became involved in the plan. Cf. Kalmbacher (1998), p. 21.

⁷⁴⁴ Cf. Caron (2005a), p. 1-24 and Caron (2005b), p. 89-114.

⁷⁴⁵ Cf. the deal with the PO, Caron (2005b), p. 97.

⁷⁴⁶ Cf. section 3.3, below.

3.2.4 Belgium

It was described above how Belgium started out with an ambitious plan of state construction of the railways. Politically, this was possible because the two main power blocks – the conservative catholics and the liberal bourgeoisie – had joined hands to get rid of the *neerlandophone* and protestant North with its heavy taxation and corresponding tax burden.⁷⁴⁷ This coalition lasted long enough to enact the railway act of 1834, but ruptured around the time of the peace treaty with the Netherlands in 1839.⁷⁴⁸

The first private investor to take the plunge was Gustave Deridder, who we have already met as one of the key engineers behind the introduction of the railways, as well as their construction by the state. In 1843 he obtained the concession for a railway between Antwerp and Gent, direct via St. Niklaas and Lokeren instead of around via Mechelen and Dendermonde.⁷⁴⁹ His *Compagnie du Chemin de Fer d'Anvers à Gand par Saint-Nicolas et Lokeren* turned out to be a well run and well engineered enterprise, and was highly profitable as a result.⁷⁵⁰ Following his example were a number of similarly small companies, originally dominated by English investors.⁷⁵¹ Starting from the 1850s, domestic investors, too, started entering the market in larger numbers.⁷⁵² At the same time, the state was satisfied to operate the four main lines, running in four directions from Mechelen. It felt that it had constructed those lines that were essential for the prosperity of the new country, and that further railway construction should be left to the market.

The result was that by 1870 there was 3.348 km of railroad in operation, a density that not even England could match. Of that length, 869 km was owned and operated by the state and 605 km by the *Grand-Central Belge*, a company owned by the *Société Générale*, including the line from Hasselt to Maastricht in the Netherlands, and from there to Aachen in Prussia. In addition, another 946 km. was controlled by the *Banque de Belgique*, the *Société Générale's* main competitor. The remainder was controlled by dozens of smaller companies.⁷⁵³ In other words, while Belgium had started out with a purely state-controlled system, by 1870 the state ran only 26% of the network, albeit some of the most important lines.

⁷⁴⁷ Cf. Gemelli & Royer (1864).

⁷⁴⁸ Cf. Gemelli & Royer (1864), p. 329-330.

⁷⁴⁹ Cf. Reynebeau (2003), p. 50.

⁷⁵⁰ Cf. De Economist (1862), p. 459 and Reynebeau (2003), p. 52.

⁷⁵¹ Cf. Buelens, Van den Broeck & Willems (2008). Van den Broeke & Van Nederveen Meerkerk (2001), discuss how the Dutch NRS was similarly dominated by British shareholders.

⁷⁵² The role of the *Société Générale* in the early history of Belgium is complex, and better left for another time.

⁷⁵³ Data taken from Linters (1985), p. 36-37.

While most European countries were giving free reign to private enterprise as a means of developing the new technology, in Belgium the state took the lead. While in most European countries, the railway sector was consolidating, in Belgium the only consolidation that took place was of a financial nature: the two major investment companies of *Société Générale* and *Banque de Belgique* owned 46% of the network by 1870, but they did so through large numbers of subsidiaries, who remained fairly free to run their own business. As a result, this “light” consolidation only did little to reduce transaction costs between companies.

After 1870, the Belgian railway sector was slowly consolidated through nationalisation. Since nationalisation is the subject of section 3.3, below, the history of Belgium after 1870 is better discussed there.

3.2.5 Germany

In the chapter so far, only a small part of the German railway history has been discussed. The obvious reason for this is that, given the division of 1830s Germany into 37 sovereign states and 4 free cities, a comprehensive discussion of the history of this period would take up a volume of its own.⁷⁵⁴ Instead, only an example was given of the difficulties encountered by German railway pioneers, both due to the reluctance to see the state involved in private enterprise that the German lands shared with the rest of Europe at the time, and due to the extraordinary conservatism of the government in Berlin. At the same time, the account of the history of the Köln-Minden railway illustrated the importance of a comprehensive system of company law, a topic that is somewhat outside the scope of this thesis, but that is nonetheless important to bear in mind.

By the time of the revolution of 1848, only the small duchy of Braunschweig had followed the Belgian example and constructed a state owned railway.⁷⁵⁵ In other parts of the country, though, public opinion was shifting. In Bavaria, where the German railways had their

⁷⁵⁴ Or at least a chapter, cf. Le Chatelier (1845), chapter 1. Cf. also Dunlavy (1994).

⁷⁵⁵ Cf. Gall & Pohl (1999), p. 14. The first section of the railway, from Braunschweig to nearby Wolfenbüttel, was opened on 1 December 1838.

There were, however, other German states that had moved at least partway towards publicly owned railways, including Baden, where the railways were mostly public starting with the railway act of 1838, Württemberg, where a programme of public construction was created in 1843 and Bavaria, where the München & Augsburg railway company was nationalised in 1844, starting a process of ever greater state involvement in the industry. Cf. Le Chatelier (1845), p. 2-13 and Ziegler (1996), p. 31-32.

origin⁷⁵⁶, the government had written a carefully designed set of laws to govern many aspects of the railway industry, including the use of eminent domain power, and their laws were adopted in many other German states as well.⁷⁵⁷ In Prussia, the military was starting to come around to the value of a railway system, particularly for East-Prussia, where the sparse population meant that private parties were unlikely to be interested, while the threat of war with Russia or rebellion in Poland was constantly present.⁷⁵⁸ As early as 1842, the state had started to give guarantees similar to the ones used in France, albeit with more stringent conditions.⁷⁵⁹ However, when in April 1847 the King proposed that the state might construct an “*Ostbahn*” railroad between Berlin and Königsberg, the capital of East Prussia⁷⁶⁰, which could then be transferred to a private company for operation, the United Landtag refused to approve the loan.⁷⁶¹ After the revolution of 1848, both the constitutional situation and the attitude towards state construction had changed somewhat. Instead of the somewhat irregular United Landtag, Prussia now had a normal parliament, as required by the constitution of December 1848.⁷⁶² This new bicameral Prussian Landtag quickly consented in the construction of the *Ostbahn* by the state.⁷⁶³

In this way, the different German states each moved in their own direction, at their own pace, depending on their own constitutional situation and the changes caused in each state by the revolution of 1848. Sometimes, this caused friction, since the often erratic manner in which the borders had been drawn created ample scope for both collaboration and free riding.⁷⁶⁴ The purpose of the railway in Braunschweig was, for example, was to form a connec-

⁷⁵⁶ On 7 December 1835, Germany's first railway was opened between Nürnberg and Fürth. Cf. Le Chatelier (1845), p. 6-7.

⁷⁵⁷ Cf. Gall & Pohl (1999), p. 14.

⁷⁵⁸ The two most important Polish uprisings against Prussian rule were in 1846 and 1848. Cf. Pease (1999).

⁷⁵⁹ Cf. Dunlavy (1994), p. 54-55 and Gall & Pohl (1999), p. 21. Notably, the Prussian conditions not only required the railways in question to pay the state a share of their profits, but also gave the state control rights over the company should this be necessary.

⁷⁶⁰ This city is currently known as Kaliningrad, the capital of the small Russian exclave of the same name, between Lithuania and Poland by the East Sea.

⁷⁶¹ The approval of this body, consisting of representatives of the provincial parliaments we already encountered above, was necessary under a promise made by King Friedrich Wilhelm II in 1820. The state of Prussia would not increase its debt without approval of the United Landtag. The fact that the representatives from the Rhine provinces were joined in this vote by the Prussian Junkers, i.e. country nobility, even though it was a project that had clear benefits for them was connected to wider constitutional conflicts in that period. Cf. Taylor (1945), p. 70 and Ziegler (1996), chapter 3.2.

⁷⁶² Art. 60-84. Art. 13 of the German Act of Confederacy of 1815 required that all German states should have a written constitution. In Prussia, however, the kings had always blocked this, referring to their divine right to rule. In 1848, however, the chaotic situation following the collapse of the national parliament in Frankfurt was finally resolved by the enactment of a constitution, which was subsequently amended somewhat in a reactionary direction in early 1850. Cf. Taylor (1945), chapter 5.

⁷⁶³ Cf. Lehmann (1998), p. 31.

⁷⁶⁴ Cf. Gall & Pohl (1999), p. 15.

tion between the river Elbe at Magdeburg in Prussia and the river Aller in Hannover, which would allow for two competing routes between the mining districts in the Harz mountains and the world markets: Via the Elbe to Hamburg and via the Aller and the Weser to Bremen. Entrepreneurs in the Kingdom of Hannover were planning to build such a connection, but intended to pass to the south of Braunschweig, so that the line would only run through two states instead of three. To avoid being sidelined like that, the Duchy of Braunschweig had no alternative but to construct the railway with public funds. Similarly, there was frequent conflict between the Kingdoms of Prussia and Saxony about alleged attempts of each to outmanoeuvre the other.⁷⁶⁵

Like in England, this kind of conflict meant a significant loss of efficiency in the short term, but increased redundancy and competition in the long run.⁷⁶⁶ However, this system of competition between the sovereign states as well as between companies was predicated on the continued division of Germany into small states. Within the enormous reaches of Prussia, Austria and Bavaria, the authorities could be as inefficient as they liked: they lacked even the relatively weak alignment incentives created by a democratic system of government to push them in the direction of an optimal alignment between transaction characteristics and governance structure. Similarly, the incentives to create and maintain an efficient governance structure for the railway industry were constantly reduced as, between 1848 and 1870, Prussia came to control a larger and larger part of the German territory.

It is worth pausing a moment to consider what this development means for our analysis, as it was outlined in the previous chapter. On the one hand, clearly the unification of Germany meant less competition between jurisdictions, and therefore a reduced likelihood that the optimal alignment will be arrived at, regardless of what it turns out to be.⁷⁶⁷ On the other hand, there is the question of how German unification would change the locus of the

⁷⁶⁵ Cf. Gall & Pohl (1999), p. 15-16. The main problem was the connection between the Prussian heartland and Bavaria. At the moment, the normal railway connection between Berlin and Bavaria runs via Leipzig in Saxony. (cf. bahn.de) However, when this route was first constructed, there would have been significant benefits for Prussia to pass Saxony to the west, through the present day German Land of Thüringen, which, at the time, was divided into many small states, such as Goethe's Weimar. Cf. Ziegler (1996), p. 31. This was all the more attractive since the Saxon-Bavarian railway company had been nationalised by the Saxon government in April 1847, because it was about to go bankrupt.

⁷⁶⁶ As noted above, the German railway system proved far superior to its French counterpart in the war of 1870, especially because it had such redundancy built in. Cf. Gall & Pohl (1999), p. 26-27.

⁷⁶⁷ The reduction in competition and evolution is compensated somewhat by a higher incentive to design an efficient system. After all, the costs of design are fixed, while the benefits increase with the scale of the jurisdiction. Cf. section 2.6, above.

optimum, if at all. Applying the Williamson framework⁷⁶⁸, we see that unification would reduce the government's uncertainty by giving it more expertise in dealing with railway issues. At the same time, the more centralised governance structures become more inefficient, since administrative controls are less effective if they have to be used over large distances in a time of difficult communication. Also, centralised governance is likely to be insufficiently adaptive, the more so if it is applied to a larger area.

Similarly, considering this problem from the GHM point of view⁷⁶⁹, we would conclude, for example, that applying a centralised governance structure over a larger geographical area will likely require an additional layer of bureaucracy, who will likely demand a share in the gains from trade, thus reducing everyone's incentives to invest. Such residual control rights can be justified, but only if there are externalities between the different regions that require resolution at a higher level of aggregation, which is unlikely to be the case.⁷⁷⁰ It follows that the optimal level of government involvement in the railway industry will tend to be lower the larger the jurisdiction is; as Germany unifies, it should tend away from public railways, *ceteris paribus*.⁷⁷¹

On the whole, the curious result is that in Germany, nationalisation generally preceded consolidation. In the 1850s and 60s, more and more German states tended towards state influence in the railways, often for similarly pragmatic reasons as the ones that drove the Dutch government down that path. This was particularly true in the south of Germany. In Prussia – where a combination of fiscal problems and conservative fear of parliamentary influence made a state-run railway system nearly impossible – the state was involved in the railways in a variety of different ways, ranging from mandatory board seats, to the usual approval right for the route, to the acquisition of shares.⁷⁷² In the rest of northern Germany, the states tended to have similarly “mixed” governance structures.⁷⁷³

A side-effect of the widespread tendency towards at least some state influence in the railway industry was that the German states were increasingly reluctant to federalise the management of the railways, even after the establishment of the German Empire in 1870, for fear

⁷⁶⁸ Cf. section 2.4, above.

⁷⁶⁹ Cf. section 2.5, above.

⁷⁷⁰ At least not when the question concerns a change in scale from the level of Prussia to the level of Germany.

⁷⁷¹ That last qualification is, of course, important to remember.

⁷⁷² Cf. Lehmann (1998), p. 34-36 and Ramuschkat (2002), p. 8.

⁷⁷³ Cf. Ziegler (1996), chapter 4.

of losing control over their transport infrastructure to the (Prussian) government in Berlin.⁷⁷⁴ While they had worked together since the foundation – in 1847 – of the *Verein Deutscher Eisenbahnverwaltungen*, i.e. the Association of German Railroadauthorities, this body never acquired any major prominence.⁷⁷⁵ The would-be Paulskirche constitution of 1848 made the railways a mixed competence, with significant federal authority, but it never entered into force.⁷⁷⁶ In 1871, Germany finally did get a proper constitution, and it had an entire chapter devoted to the railways, reflecting the political sensitivity of the issue.⁷⁷⁷ On the whole, this constitution left the control over the railways with the states, although the empire could enact laws overruling the veto of the local state to construct a railway when this was necessary “in the interest of the defence of Germany or in the interest of the common transport”.⁷⁷⁸ Not only did the constitution abandon the idea of a unified railway system, it also partly – though implicitly – abandoned the idea of a state railway. In several places, it referred to private railway companies, something that would have been quite unlikely in the early 1850s.⁷⁷⁹ The constitution did, however, create the basis for the creation of a *Reichseisenbahnamt*, i.e. an Imperial Railway Authority in 1873.⁷⁸⁰ While the full consolidation of the Germany railways had to wait until the Weimar constitution of 1919, which required both the nationalisation and the federalisation of the railway industry⁷⁸¹, the REA would act as a powerful force towards centralisation throughout the rest of the history of the German empire.⁷⁸²

3.2.6 Conclusion

We have now once reached a point where it makes sense to look back to the list of section 2.7, and compare it to the story so far. Some of the matters discussed at the end of section 3.1 apply to the railway industry generally, and do not need to be repeated here. Instead, we will focus on the changes in the industry compared to its pioneering era, and how they affect the applicability of the new institutional economics to the problem of the governance of semi-public institutions.

⁷⁷⁴ Cf. Gall & Pohl (1999), p. 27.

⁷⁷⁵ Cf. Lehmann (1998), p. 33.

⁷⁷⁶ § 28. Cf. Taylor (1945), chapter 4, for a discussion of the events of 1848-49.

⁷⁷⁷ Chapter VII, covering art. 41-47. Cf. Ziegler (1996), p. 178-179.

⁷⁷⁸ Art. 41 (1).

⁷⁷⁹ Cf. Ziegler (1996), chapter 2 and 3.1.

⁷⁸⁰ Cf. Ziegler (1996), p. 180-181.

⁷⁸¹ Art. 89.

⁷⁸² Cf. section 3.3, below.

Transactions: While there are still plenty of transactions going on, both between the government and the railway companies, and between the companies amongst themselves, the emphasis is starting to shift. As to the former, there is evidence that transactions are not always conducted at arm's length. If the less-than-arm's-length connection between government and industry took the form of some kind of vertical integration, that would be no problem, but it is far from clear that the kinds of backroom connections described above – where in a manner of speaking railway executives smoked cigars with government ministers in the gentlemen's club they both belonged to – can be so considered. At its most favourable, perhaps we can say that this is a form of what the GHM models were said to predict earlier: a takeover of the government by its contractual counterparty.⁷⁸³ Between companies, on the other hand, the period of consolidation is defined by the “internalisation” of transactions, something completely consistent with – and in fact integral to – institutional economics.

Overall efficiency: This is an important point to pause at. Economists often have a tendency to emphasise the non-materialistic nature of their utility and social welfare functions in theory only. In practice, they often find it convenient to stick to those consequences of proposed policy choices that can be translated into monetary terms relatively easily, such as tax increases, increases in unemployment and unemployment benefits, etc. During the second half of the 19th century, we see a gradual shift away from the government as facilitator, keeping an eye mostly on the needs of the industry and of the property owners involved, towards a model where the government considers a wider range of interests. While the benefit of the railways for mail delivery had always been at the forefront of everyone's mind, governments became increasingly aware of the military implications of the new technology, as well as its social implications. Many of these new parts of the equation are much more difficult to put a number on than one would perhaps prefer.

Examples of the social implications of the development of the railways range from Wellington's famous comment at the opening of the Liverpool & Manchester that the new invention would simply “encourage the lower classes to travel about”⁷⁸⁴ – which was evidently meant as a point of criticism – to the increasing attention given to the needs of the urban poor displaced by railway construction. In the early decades of the railways, the compa-

⁷⁸³ Cf. section 3.1.4, above.

⁷⁸⁴ Wolmar (2007), p. 41-42. Wolmar does not source the quote, but it is also given in Hilton (2006), p. 15, and in Christiansen (2002), p. 91 fn. According to the latter, the phrase was actually uttered at the opening of the Liverpool & Manchester in 1830. The Duke of Wellington was Prime Minister of the UK from 1828-1830.

nies often took advantage of the fact that the inhabitants of the urban slums lacked the kind of bargaining power that country landowners possessed. In other words, they made whatever arrangements they needed to with the people who actually owned the property, and then threw the inhabitants out on the street.⁷⁸⁵ The governments of the various countries under consideration only slowly started to intervene to prevent this practice.⁷⁸⁶ Such changes in priority are not so much signs of a reduced concern for the efficiency of the governance structure, but rather a change in the objective function of the government, which may well imply a shift in the theoretical optimum alignment. For example, to the extent that the railways cannot be induced by contractual and statutory means to seek an equitable solution for the underprivileged displaced by their construction, an increase in the weight given to this problem implies a shift of the optimum in the direction of nationalisation.

In describing the government's preferences in this way, it is important to draw a clear line between the government's preferences and the preferences of the individuals who control the government. The fact that the individuals who make up the government and its bureaucracy also have preferences is the key explanatory hypothesis of public choice theory, and as such we will encounter it again in the next chapter. Institutional economics, in all its forms, treats the parties to a given transaction as units, and inquires no further into the content of that black box.⁷⁸⁷ For the time being, we will define the government's preferences as its legitimate, stated policy goals. In a democracy, that will normally mean the government's – or the governing party's – manifesto, while in less democratic countries the government's legitimate, stated policy goals will have to be observed by looking at the statements of those in power or – more problematically – inferred from the government's actions. Either way, the government's preference ordering is expected to look much like the utility function of an individual, with all the usual requirements.⁷⁸⁸

With that proviso, we can conclude that the overall efficiency requirement was met both more and less than in the initial period. On the one hand, the relevant actors more fully examined the range of possible governance structures, and in fact displayed a greater range of choices. The initial taboo in most countries against government involvement, which often made a full examination of the problem impossible in the 1830s and 40s, dissipated in the

⁷⁸⁵ Cf. Ramuschkat (2002), p. 219-235, Wolmar (2007), p. 117-118.

⁷⁸⁶ Cf. Labouring Classes Dwelling Houses Act (1866), 29&30 Vict., c. 28. Cf. also Gall & Pohl (1999), ch. 4.

⁷⁸⁷ Except, of course, by examining yet more (potential) transactions.

⁷⁸⁸ Cf. Cowell (1986), chapter 4.

wake of an increase in overall political pluralism throughout Europe, i.e. the rise of socialism, nationalistic mercantilism, and technocratic conservatism taking their place next to liberalism in the European political landscape.⁷⁸⁹ All options were finally being examined, even in the UK.

On the other hand, we see a problem with the government's preferences: not all the preferences that drive the outcomes are government practices. In Britain, we see railway executives sitting in parliament, in Belgium, we see powerful banking conglomerates forcing the government to get involved, etc. These are facts that argue against a model that works from a unitary government objective function.

Costly Information and Moral Hazard: As is to be expected, information becomes less costly as all parties become more experienced. The British railway boom of the mid 1840s was followed by several more, but none of them were quite as enthusiastic as that first one. Similarly, the French railway crisis of the 1850s showed the authorities the consequences of the burdens they were trying to place on the companies, teaching them about the real costs of certain mandates, such as the requirement to carry the mails for free, or to build railways even in sparsely populated but politically or militarily important regions. Writing such requirements into concessions instead of arranging them separately is politically convenient, since it hides their true cost. As experience grew, and as the industry lived through its early difficulties, this became somewhat less true.

That is not to say that the cost overruns were suddenly a thing of the past. If nothing else, the increasing separation of ownership and control⁷⁹⁰ meant that railway promoters had an incentive to hide the truth not only from the state, but also from their own stockholders. We have seen how George Hudson not only raked up enormous debt, but also cooked the books in order to make his railway empire possible. Then again, when he did so, at least he hurt mostly his own investors – both minority shareholders and bondholders – and not his counterparty in the transaction that we are currently interested in: the state. Given the limited involvement of the state in his business, and its limited set of requirements, there was little he could do that would hurt the state. At worst, he could cause delays in the construction of railroads that the state found beneficial to the common good, or at least not adverse thereto. Matters would have been different if the British government, like its French counterpart, had giv-

⁷⁸⁹ Cf. Fyffe (1895), ch. XVIII, Von Krockow (1990), ch. 1 and Veblen (1915).

⁷⁹⁰ Berle & Means (1932).

en him a return guarantee. Such a guarantee necessarily requires detailed oversight over the company's affairs, as indeed the French government attempted to carry out.⁷⁹¹

Asset Specificity and Fundamental Transformation: Compared to the early days of the railways, barriers to entry started to develop in subsequent decades. It was no longer the case that any gentleman with an enthusiasm for new technology could get some friends together and start a railway company. Only in Britain was that still – to some extent – possible, but even there an independent would-be railway entrepreneur was at a distinct disadvantage relative to his established competitors. This was not just a matter of money. Money could, as before, be borrowed or acquired through emissions of equity, assuming the proposed project was in fact likely to return a profit. Instead, the difficulty was the expertise necessary to compete in the market and for the market. If no one has expertise, all can compete equally. Afterwards, those who initially win the fight to be allowed to compete in the market acquire knowledge about the ins and outs of the new service, knowledge that can only be obtained by outsiders at a premium. It is, for example, much easier to judge whether a proposed concession allows for profitable operation if you already operate a railway somewhere else, meaning that outsiders are at risk of overpaying should they find themselves winning a concession anyway.

This knowledge is undoubtedly an asset, and a specific one at that, but the question is: Specific to what? Mostly, this knowledge will sooner be specific to the industry generally than to any particular transaction. The only significant assets that are involved in the transaction and are specific to it are the physical infrastructure, as before. The net effect of it all is that the critical distinction is between those who have a significant presence in the industry, and those who do not. While the railway companies were always forced to do business with the monopolist-government, the government is now forced to do business only with established companies as well, a circle that is ever shrinking through mergers and acquisitions, as well as deliberate government policy. When, as part of the restructuring of the Belgian railways system, the Dutch government acquired the Dutch part of the Antwerp & Rotterdam railway, there was no one that could sensibly take that concession except the SS.⁷⁹² The combination of the (network) economies of scale displayed by the physical assets and the scarcity of the human capital necessary meant that the state was forced to do business with the SS just like the SS was forced to do business with the state. They each had only one alternative: to walk away altogether.

⁷⁹¹ Some of the data over the first few decades are collected in De Economist (1874).

⁷⁹² Cf. Veenendaal (2004), p. 173.

Uncontractible Investments: To start with the government's "investment" in the project, the change from the earliest days of the railways is, of course, that the railway companies became used to the occasional unilateral government intervention in their business. They became used to the fact that their industry was no longer considered a strictly private matter, to be regulated through contracts and contract-like instruments negotiated between the state and the company. That does not mean that such changes in the law did not affect the companies' return on investment, but it does mean that it becomes more difficult to imagine some kind of government commitment that we can categorise as a government investment. Instead, the possibility that the law might change is simply one more source of political risk, similar to the risk of war or revolution.

At the same time, the government became ever more eager to persuade the railway companies to make certain investments other than investments in physical capital. Many of those investments were to some extent uncontractible. This is a subject that, in a different guise, we already discussed above: to the extent that the government's changed preferences required uncontractible investments from the railway companies, for example in order to prevent catastrophic railway accidents, the GHM model suggests that the railway companies should own the government.⁷⁹³ Less drastically, the Hart, Shleifer & Vishny model suggests that private ownership should be preferred when uncontractible quality is relatively unimportant⁷⁹⁴, though their model reserves judgement on the opposite statement.⁷⁹⁵ This explains why an increase in the government's interest in what might generously be called "quality", i.e. everything the government cares about other than the price of services rendered, might motivate it to nationalise the industry, or at least to move in that direction.

Alignment mechanisms: Once the really big startup problems had been solved, and the really bad ideas had been filtered out, evolution became much less of a factor. To be sure, companies still went bankrupt, but only because of their own internal mistakes. Even if the bankruptcy is the result of government overreaching, as was arguably the case in France in the 1850s, the fault still lies with the company that was too optimistic about its future prospects under the deal as it was proposed. In a free market, evolutionary forces inevitably punish such excessive optimism, but this does not necessarily imply that there was something wrong with the governance structure. In the recent financial crisis, many banks and other investors were

⁷⁹³ Cf. Hart & Moore (1990), proposition 2.

⁷⁹⁴ Hart, Shleifer & Vishny (1997), proposition 3.

⁷⁹⁵ Hart, Shleifer & Vishny (1997), proposition 5.

punished for their excessive optimism with bankruptcy, but that does not mean the government should get in the banking business.

Competition, on the other hand, became more of a factor towards the end of the 19th century, for the simple reason that the railway industry became an ever more important factor in the economic success of a country. Even the Dutch economy could no longer avoid depending on this new land-based mode of transport. As mercantilism and militarism rose again after the liberalism of the earlier 19th century, the pressure to get it right steadily increased. Countries could often no longer afford to get it approximately right, to deviate from what they knew to be optimal in order to accommodate some powerful special interest, because to do so would quite possibly doom their economies or their national security. In response to this pressure, they started to adjust the governance of the industry ever more frequently, making changes at one point or another, in order to move closer to the optimum. Between 1850 and 1900 the UK parliament enacted 45 railway-related statutes, almost one per year, including 15 in the 1860s alone. The Labouring Classes Dwelling Houses Act of 1866⁷⁹⁶ was already mentioned earlier, but this was only one of many railway-related statutes during this period. None of the other countries under consideration here quite matched this productivity, but it is clear from the increased interest in the railways displayed by each of their governments that they, too, felt the pressure of competition.

Another thing they did was establish official inquiries. Remember, in chapter two we drew the line between competition and design at incrementalism: constant competition forces legislatures to gradually, incrementally adjust the governance structures used, while design involves drastic changes taking place at the beginning of the industry, or at some other suitable moment. The most obvious category of “suitable moments” will only be discussed in the next section: after a major war. The only major war that we have encountered so far was the Franco-Prussian war, which was so short and so catastrophic for the French that it did not cause them to make any major changes: they simply continued with their existing system. However, as we have seen this war did give the Belgians pause, and caused them to start a major overhaul of their railway industry, albeit in a somewhat haphazard fashion. Inquiries in other countries did not usually produce such drastic change: forces defending the status quo would normally be able to defeat most of the more drastic reform proposals, leaving open only a path of incremental change through competition.

⁷⁹⁶ 29&30 Vict., c. 28

As far as alignment mechanisms go, we can say that most of the alignment came through competition, with one notable case of design. However, this attests to the lack of need of drastic change as much as anything else. After all, the theory suggests that decreased uncertainty would move the optimum towards contracts, where it – rather inexplicably under many theories – already was since the beginning. While the increased tendency of the government to make more complex and uncontractible demands on the railways has the opposite effect, the net effect under most of the theories we discussed in chapter 2 will still be that the ability of the state and the railway companies to write complete contracts to govern their transactions will have increased. Since that was what they were already doing anyway, the alignment mechanisms had relatively little work to do.

3.3 Nationalisation

As we have already seen, the distinction between the period of consolidation and the period of nationalisation is not always clear. In Germany, consolidation came after nationalisation, while in most other countries under examination consolidation was encouraged by – if not organised or even required by – the government. In this way, consolidation and nationalisation often went hand in hand, occurring at the same time and defended with the same arguments, with the actual transition from a consolidated railway industry to a nationalised one merely delayed because of the remaining liberal sentiments that resisted such blatant state involvement in the economy.

In section 3.2, the story ended at different points in time for different countries, depending on when the period of nationalisation could be said to begin in earnest. For Germany and Belgium, that was 1870, whereas for France, the Netherlands and Britain, the push for nationalisation did not come until the Edwardian era and the Great War that followed. In order to keep this history intelligible, it is therefore necessary to take up the story in this order, starting with the early adopters and discussing the UK last.

3.3.1 Germany

As already explained, in Germany the question of nationalisation of the railways continued to be inextricably linked with the questions of their consolidation and federalisation until the Weimar constitution of 1919, which settled the matter once and for all.⁷⁹⁷ By 1870, there were virtually no purely private railways left, for the same reason that the European Treaties have special sections restricting state aid to private industry⁷⁹⁸: in an interlinked common market like the present day European Union or the German customs union of the 19th century⁷⁹⁹, such aid tends to become competitive: if one state does it, the others have no alternative but to follow suit. For that reason, the governance structures used by the German states in 1870 tended

⁷⁹⁷ Gall & Pohl (1999), p. 27, distinguish between two blocks of supporters of nationalisation, depending on whether they wanted to nationalise at the level of the empire or at the level of the state: “eine Fraktion, die für eine Verstaatlichung mit Blick auf die bestehenden Bundesstaaten eintrat, und eine, die zugleich die “Verreichlichung” des gesamten Eisenbahnwesens, eine Reichsbahn, ins Auge faßte.”

⁷⁹⁸ Art. 107-109 TFEU.

⁷⁹⁹ The so-called *Zollverein* was founded in 1818, after the question was left out of the original confederacy treaty of 1815. Initially, it only covered Prussia and the Hohenzollern territories in the south of Germany, but throughout the 19th century, more and more German states joined, until it became part of the Constitution of 1871. (§ 33-40.)

to range from a so-called mixed system of government guarantees and stock holdings⁸⁰⁰, to a mixed system in the sense of having a range of different arrangements within the same jurisdiction⁸⁰¹, to a fully nationalised railway system.⁸⁰² The only states that could afford to be relatively independent were the giants of Prussia and Bavaria.

After 1870, the imperial government led by Chancellor Von Bismarck quickly came to be viewed as the legitimate government of Germany, making it easier for the government to formulate a comprehensive federal railway policy to the extent that this was considered appropriate, and to the extent that the states did not manage to veto this. Once the Imperial Railway Act had been passed in 1873, and with it the Imperial Railway Authority (REA)⁸⁰³, the imperial authorities came to view the railways increasingly from a strategic point of view. Instead of focusing only on how to set up the best possible railway system for the lowest amount of money, they started thinking about how the railways fit in the overall economic policy of the empire, not to mention its fiscal, military, political and socio-cultural policies.⁸⁰⁴ In order to make this possible, both consolidation and federalisation were urgently required, and particularly the Prussian railway network was rapidly taken over by the state after 1870.⁸⁰⁵⁸⁰⁶

⁸⁰⁰ Eg. Saxony. Cf. Ziegler (1996), chapter 3.5.

⁸⁰¹ Eg. Prussia, where the tendency was towards relatively low government involvement, cf. Gall & Pohl (1999), p. 29, but where they were willing to make exceptions if this was necessary for military purposes. Cf. section 3.2.5, above.

⁸⁰² Eg. Baden. Cf. Ziegler (1996), p. 476-484.

⁸⁰³ Cf. Lehmann (1998), p. 36-37.

⁸⁰⁴ Some of the highlights from the minutes of the Prussian cabinet, which have recently become available, cf. Spenkuch (1999, 2003): 8 May 1881: Concern about the ability of Social-Democratic Members of Parliament to use their complimentary tickets for "agitation". 6 March 1882: Discussion about what newspapers should be made available at train stations, particularly whether these should include "opposition newspapers". 29 November 1890: Special rate for Maize, especially from Austria-Hungary, to combat the shortage of animal foods in the Eastern Provinces. 22 April 1892: The new management of the Almelo & Salzberg railway is approved, with the proviso that care should be had that no Dutch influence should be tolerated, given the military importance of that railway. 22 April 1899: Discussion about a US request for lower railway rates for American wood. 4 December 1899: No special rates are to be approved for the Paris exhibition, "da man Frankreich als einen wirtschaftlichen Konkurrenten betrachten muß, dessen Ausstellung nicht besonders zu unterstützen ist."

⁸⁰⁵ Gall & Pohl (1999) give the following data for Germany within the borders of 1871:

- 1870: 20.000 km total length, including 11.000 km privately owned. 3.000 km. privately owned but under public management.
- 1880: 34.000 km total length, including 12.000 km privately owned. 4.000 km. privately owned but under public management.
- 1888/89: 40.000 km total length, including 4.700 km privately owned, not counting the railways that were privately owned but under public management.

In other words, the percentage of the railway network that was privately managed fell from 40% in 1870, to 23½% in 1880 and 11¾% in 1888/89.

⁸⁰⁶ Curiously, one of the reasons why the railways were rapidly being nationalised even outside Prussia was that the other German states were worried that their private railways might be bought up by the Prussian government, as had happened in Braunschweig, for example. Cf. Knowles (2006), p. 219. On 9 April 1895, for instance, the

Two key aspects of Bismarck's "big picture" approach to the railways illustrate both the advantage and the disadvantage of government management, particularly in a country that is only partly democratic, as the German Empire was before 1919.⁸⁰⁷

In 1860, France and Britain signed the Cobden-Chevalier Treaty, strongly reducing the trade levies between both countries.⁸⁰⁸ This decision, resulting as it did from a development that had already started in Britain with the repeal of the Corn Laws in 1846, started a cascade of bilateral treaties between the countries of Europe.⁸⁰⁹ While the height of this fervour was short-lived – Bismarck's Germany already enacted a new tariff in 1879⁸¹⁰ – it created a presumption in favour of free trade that lasted until World War I. Politically and diplomatically it was much more difficult to enact tariffs than it had been before. Even though much of that achievement was undone in the last few decades of the 19th century under the joint attack of socialism and nationalism⁸¹¹, the German government found it useful to use the railway rates as part of its international trade policy: instead of imposing an import levy, they would simply charge more for transport of goods from the coast in-land, than for transport in the opposite direction.⁸¹²

Leaving aside for one moment the question of whether such import levies are a good idea⁸¹³, the fact that Chancellor Bismarck and his government thought to use their power over railway rates in this way clearly shows the "big picture" general equilibrium approach at work. Given that everything in the economy is connected to everything else, a government interested in managing the railway industry from this perspective would have to write an even more complex contract in order to allow for the privatisation of the industry than a government only interested in avoiding monopoly behaviour in the partial equilibrium sense. In other

minutes of the Prussian cabinet show discussion about the conflict with Saxony about the intended purchase by the Prussian state of the Weimar & Geraer Railway. The proposed solution is to transfer the Zittau & Nikrisch railway to the state of Saxony. Cf. Spenkuch (2003), p. 186.

⁸⁰⁷ The lower house of parliament, the *Reichstag*, was elected by general male suffrage (§ 20 of the Constitution of 1871), while the upper house, the *Bundesrat*, consisted of representatives of the states (§ 6). The *Reichskanzler* and the other ministers, however, were answerable only to the Emperor (§ 15). Cf. Berman (2001), p. 437-454.

⁸⁰⁸ Cf. Iliasu (1971). For a general discussion of trade policy in the second half of the 19th century, cf. Hobson (1997) and Neff (1990), chapter 2.

⁸⁰⁹ Cf. Trebilcock & Howse (1995), p. 21.

⁸¹⁰ Cf. Lehmann (2009) and Solsten (1999), p. 45.

⁸¹¹ Cf. Neff (1990), p. 61-71.

⁸¹² Cf. Knowles (2006), p. 221.

⁸¹³ Classical economic analysis would indicate that they are not.

words, while a government that views the railway sector only from the perspective of a static partial equilibrium perspective might be able to achieve its goals best by privatising while an otherwise identical government with a general equilibrium perspective would be better served by nationalising. A policy shift from a partial to general equilibrium perspective could, therefore, explain a tendency towards nationalisation.

The other side of the coin is reflected in a short report in the Dutch economic magazine *De Economist* in 1892, discussing the budget for the Prussian state railways for the year 1892/1893. The key numbers are: Revenue: 967 million marks, ordinary expenditure: 601 million marks, surplus from operations: 366 million marks. This surplus was to be transferred to the Prussian state in the form of interest: 207 million marks, and repayment of debt: 159 million marks.⁸¹⁴ Throughout the period of the second empire, the profits of the railways were one of the most important sources of income for the Prussian state⁸¹⁵, and therefore for the empire itself.⁸¹⁶

Of course, the one thing worse than a private monopolist running the railways is a public monopolist running the railways. As discussed in chapter 1, the analysis in this thesis is usually predicated on the assumption that the government will seek to promote the general welfare in some sense. That assumption could be cast in the form of an explicit social welfare function⁸¹⁷ or in the form of a general tendency to please voters, or something in between. Assuming, instead, that the government is only interested in maximising its revenue, will affect the analysis to the extent that a) railway profits are a large portion of the state budget, b) taxes are low relative to the country's national income. As long as the railway profits are relatively unimportant for the government, and as long as taxes are significant relative to GDP, the state will still have an incentive to maximise GDP, as before. In late 19th century Germany, the situation was exactly opposite, and as a result the Prussian government often refused to allow otherwise desirable price reductions on the grounds that this was fiscally im-

⁸¹⁴ *De Economist* (1892), amounts rounded to the nearest million. The fact that the railways are made to pay for their own purchase is, of course, not very important here.

⁸¹⁵ Cf. Knowles (2006), p. 223.

⁸¹⁶ Cf. art. 70 of the constitution of 1871, which states that the expenditure of the Empire is covered with tolls, sales taxes and profits from the postal and telegraph industry. The remainder is to be contributed by the states in proportion to their population. Data from 1910 indicate that at that time, about 40 million of the Emperor's 65 million subjects were inhabitants of Prussia.

Until the start of World War I, Prussia had a higher budget than the Empire. In 1898, for example, the Prussian budget was about 2,3 billion marks, while the Empire spent around 1,6 billion.

⁸¹⁷ Cf. section 2.2.3, above.

possible.⁸¹⁸ Notably, in 1891 the responsibility over the tariff policy of the railways was taken away from the ministry of public works, and placed with the cabinet in general and the finance minister in particular.⁸¹⁹

Viewing the history of the German railways in the 19th century, we come to the curious conclusion that the system – for lack of a better word – generally functioned well. It produced a well-functioning and widespread railway network, albeit at a relatively high cost.⁸²⁰ The disorganised political structure of the country allowed produced a clear case of competition between jurisdictions, with some states acting as a laboratory of democracy, as Baden did.⁸²¹ It also provided an incentive to make the railways a priority. Time and again, essential decisions were made for reasons completely unrelated to the railways, such as the Prussian king's unwillingness to compromise with the united Landtag in 1848, or Bismarck's desire for an alternative to import levies in the 1870s. Despite all this, by the end of the century, Germany had one of the best-run railway systems in Europe.

At that time, the only remaining check on the German or Prussian government's ability to operate the railways as a monopolist was the power of the other German states. The rise of economic nationalism had caused the enthusiasm for free trade that characterised the mid-19th century to all but disappear, remaining important only in the United Kingdom. As a result, the interest of the railways for the economy became ever less important relative to the interest of the railways for the state's social policy and particularly for the military.⁸²² Given the high degree of centralisation and nationalisation, there was no effective check on this development.

⁸¹⁸ Eg. Prussian Cabinet meeting of 13 October 1892: Finance minister blocks a reform in the rates for passenger transport. Meeting of 22 January 1898: Finance minister expresses concern over the new rates system for piece goods. The minister for public works expects additional revenue of 30 million marks, and also notes the general economic benefits of the proposed reform, which is ultimately adopted. Meeting of 17 June 1898: Finance minister calls on the railway administration to be careful about their spending, "denn ihre Überschüsse braucht man, um den Etat zu balancieren." Cf. Spenkuch (2003), p. 119, 297 and 316, respectively.

⁸¹⁹ Cf. Spenkuch (2003), p. 8, as well as the minutes of the Prussian cabinet of 25 May 1891, point 3, and 30 June 1891, point 2. Cf. Spenkuch (2003), p. 85 and 90, respectively.

⁸²⁰ For an introduction to the enormous bureaucracy that was involved, cf. Lehmann (1998), p. 42-50, discussing the new organisational model introduced in 1895. Cf. Lehmann (1998), p. 54-57 for a discussion of the financial and other results of the Prussian railway just before World War I.

⁸²¹ "It is one of the happy incidents of the federal system that a single courageous state may, if its citizens choose, serve as a laboratory; and try novel social and economic experiments without risk to the rest of the country." United States Supreme Court, Justice Brandeis dissenting, 21 March 1932, *New State Ice Co. v. Liebmann*, 285 US 262, at 311.

⁸²² Cf. Gall & Pohl (1999), p. 55-70.

The one remaining step was full consolidation. Under the confederacy that was the German Empire that was impossible, as already explained. The Weimar Republic, however, removed the imbalance between Prussia and the rest of the country, as well as the imbalance between the crown and the democratically elected parliament, making it possible to operate the *Reichsbahn* as a public monopoly, for the benefit of the country as a whole.⁸²³ Given the long history of state influence over the railways, especially during its management by the military during World War I⁸²⁴, this decision was not based on a great amount of careful consideration, certainly not consideration of an economic nature. Instead, the federalisation of the railways was in many ways simply a logical part of the overall reform of the German state.

To be precise, the argument for federalisation was mostly one of power and money. Regarding the former, the framers of the Weimar constitution were eager to move a significant amount of power from the states to the federal government. Given the size of the railways, both measured in money and in number of employees⁸²⁵, the question of its control had significant power implications, as it had done throughout the Second Empire. Related to this was the concern that the railways should be a “binding force” in the new Germany, just like they had always been a source of state pride before the war.⁸²⁶ As for money, there was the joint concern of the reparations that Germany had to pay under the Versailles peace treaty⁸²⁷ and the concern that the railways might not be as profitable after the war as they had been before. In fact, in many states, the government was more than content to get rid of this “clunker”.⁸²⁸ Finally, there was the problem of post-war resource shortages.⁸²⁹

In a country where the railways are nationalised but not consolidated, how do you decide who will have access to unusually scarce resources? In a free market, that is what the price mechanism is for: private actors bid for scarce resources, and whoever bids the most can have them. But when the railways are operated by the various states, who – to make matters worse – have a significant stake in their profitability, that solution is unlikely to please many. A second possibility is that those states that have significant coal and iron ore sources within

⁸²³ Cf. Gall & Pohl (1999), p. 71-107.

⁸²⁴ Cf. Lehmann (1998), p. 57-64.

⁸²⁵ One of the members of cabinet estimated the number of railway employees at 700.000 in 1919. Cf. Gall & Pohl (1999), p. 75.

⁸²⁶ Cf. Gall & Pohl (1999), p. 76 and 71, respectively.

⁸²⁷ Under art. 248 of the Treaty, the railways had been put up as collateral for the debt. Cf. Gall & Pohl (1999), p. 76.

⁸²⁸ Cf. Gall & Pohl (1999), p. 78, quoting federal cabinet minutes as saying “daß die Eisenbahn in den nächsten Jahren und vielleicht Jahrzehnten kein werbendes, sondern ein fressendes Unternehmen sein wird”.

⁸²⁹ Cf. Gall & Pohl (1999), p. 79.

their borders, i.e. Prussia, use their sovereign power over their own territory to win themselves a better deal.⁸³⁰ But would they be generous enough to extend such a deal to the rest of Germany?⁸³¹ In such circumstances, access to scarce resources that are essential for the railway industry is again a matter of power, where the other states would prefer to place that power with the federal government, rather than with the state of Prussia.

For these reasons, the German parliament decided to write full federalisation of the railways into the Weimar constitution. In 80 years, the railways had gone from being resisted hand and foot by at least part of the state bureaucracy in Berlin⁸³² to being the proud property of the new German republic: a *Reichsbahn*. That is how it would stay for the next 80 years.

3.3.2 Belgium

As we have seen, Belgium, too, was somewhat of an outlier. Unlike the other countries under consideration, it started out with a fully public core network, but subsequently left the rest of the railway industry to the private sector, with very little regulation. In the private sector, there were two major players – the *Société Générale* and the *Banque de Belgique* – but they focused on investing in potentially profitable railways, without attempting to consolidate them. But then, in 1870, the emperor of France was captured by the Prussians during the battle of Sedan, a town less than 15 km from the Belgian border in the Ardennes. One year earlier, in 1869, the English owners of the *Grand Luxembourg*, which connected Luxembourg to Liège and Brussels, proposed to sell it to the French *Compagnie de l'Est*.⁸³³ Given the precarious political situation between the two superpowers on its borders, the Belgian government stepped in and prevented the sale, ultimately buying the company itself in 1872.⁸³⁴ In this way the war of 1870 encouraged a reluctance to allow foreign companies, particularly foreign

⁸³⁰ Cf. the minutes of the Prussian Cabinet of 18 December 1899: Discussion of a new contract between the Prussian railways and the Rheinland-Westphalian Coal Syndicat. 6 September 1900: Discussion of possible measures to combat coal shortages. An export ban is rejected, as is the abolition of the preferential rates for export of coal. Instead, they decide to reduce the import tariffs on coal. Spenkuch (2003), p. 363 and 381, respectively. After 1900, the cabinet only dealt with coal policy in general, until the war, when an Imperial Coal Commissioner (Reichskohlenkommissar) was appointed.

⁸³¹ Cf. the minutes of the Prussian Cabinet of 16 February 1899: Discussion of an exemption from the usual rates for English coal going to Saxony, at the request of the Saxon government, who were experiencing a shortage because of strikes in the Bohemian coal mines. The Prussian cabinet was reluctant, due to fears that the strikes might be encouraged by their generosity. Spenkuch (2003), p. 368-369.

⁸³² Cf. Holterman (2009).

⁸³³ Cf. Buelens, Van den Broeck & Willems (2008), p. 105-106. Apparently, this proposal from France was connected to the failure of the French state to purchase Luxembourg outright in the previous year.

⁸³⁴ Cf. Van der Hertten (2004), p. 344.

companies that were closely connected to their governments, access to important Belgian lines. This attitude was already present with the nationalisation of the Mons & Manage in 1857⁸³⁵, but strengthened due to the war.⁸³⁶ Another major nationalisation that can be explained in this way is the purchase of the Pepinster & Spa from the *Compagnie de l'Est* in 1872.

A second motive for nationalisation was the troubling relationship between the government, on the one hand, and the country's two major financial players on the other hand. Both the *Société Générale*, which functioned as Belgium's central bank until the creation of the *Banque Nationale* in 1850, and the *Banque de Belgique*, which had been founded in 1835 as a more patriotic competitor for the SG⁸³⁷, had so much influence in the young country that they often behaved as an alternative government.⁸³⁸ When these companies encountered financial difficulties, they naturally looked to the government as a source of capital. One way to arrange that was to sell a railway or two to the state for a good price.⁸³⁹ An example of this was the purchase by the state of the BB's *Chemin de fer des Bassins Houillers du Hainaut* in 1871. The largest purchase of them all, the nationalisation of the *Grande Compagnie du Luxembourg* in 1873 for 240 million Francs was the result of both motives working simultaneously: the BB could use the financial stimulus⁸⁴⁰ and the alternative was its sale to German investors.⁸⁴¹ By 1882, when the first phase of the nationalisation process came to a natural halt due to the state's own lack of funds, the state operated almost two thirds of the Belgian railways, and owned more than half.⁸⁴²

As the state's influence in the railway industry increased, a third and fourth motive came into play. The more the state came to have a financial stake in the railways, the more it came to regret the unfortunate circumstance that it had competitors, inevitably undermining

⁸³⁵ The first time the Belgian government purchased a private railway was in 1857, when the *SA des Chemins de Fer de Namur à Liège et de Mons à Manage* ran into financial difficulties. The Namur & Liège section was leased to the French *Compagnie du Nord*, while the Mons & Manage section was added to the state railways. Cf. Buelens, Van den Broeck & Willems (2008), p. 104. The reason why only half the company was allowed to be leased by the French was that the Belgian government did not want a French company connecting to Brussels

⁸³⁶ Cf. Buelens, Van den Broeck & Willems (2008), p. 106 and NMBS (2001), p. 7.

⁸³⁷ Cf. Houtman-De Smedt (1994), p. 50.

⁸³⁸ Cf. Cuyvers & Meeusen (1976) and Van Nieuwerburgh, Buelens & Cuyvers (2005), p. 12.

⁸³⁹ Cf. Linters (1985), p. 37.

⁸⁴⁰ Cf. Linters (1985), p. 37.

⁸⁴¹ Cf. Buelens, Van den Broeck & Willems (2008), p. 106.

⁸⁴² Cf. The table in Linters (1985), p. 39. For 1880, it shows 1090 km of railway that was state constructed. In addition, there was 1379 km of nationalised railway, making a total state ownership of 2469 km. Furthermore, there was 323 km of railway that was privately owned but state operated, for a total of 2792 km of state operated railway, in a total railway industry of 4112 km.

everyone's return on investment. The simple solution was to nationalise the competition.⁸⁴³ Likewise, with increased state control came increased opportunity for patronage, using the construction of new train stations or the improvement of the service in certain districts as an incentive for voting the right way.⁸⁴⁴

Of course, to the extent that they are not illegal⁸⁴⁵, the choice between a merger and a cartel is exactly the kind of problem institutional economics is best equipped to handle. A successful cartel acts exactly like a monopolist, meaning that in a world without transaction costs, there would be no reason to prefer a merge over a cartel or vice versa. In the real world, which does have transaction costs, the problem associated with organising and policing a cartel are well-documented, particularly when the courts refuse to take part in this policing on the grounds that contracts in restraint of trade are against public policy.⁸⁴⁶ While cartels tend to be studied by micro-economists⁸⁴⁷, the problem can be viewed as just another application of the wider field of institutional economics.⁸⁴⁸ The state railways had a problem with excessive competition, and they quite reasonably concluded that a merger with their competitors, i.e. their nationalisation, was the most efficient solution to the problem.

Excessive competition aside, there were some genuine problems in need of solving. More than their counterparts in the UK, the Belgian railway companies seem to have had difficulties working together in such matters as ticketing and the sharing of rolling stock.⁸⁴⁹ Particularly the collaboration between the private companies and the state railways was difficult after the latter's expansion in the 1870s.⁸⁵⁰ Nevertheless, as we have seen in our earlier discussion of the situation in England in the 1840s, these difficulties alone can be solved without full consolidation. All things considered, it appears the desire to avoid "wasteful competition" was the main reason for nationalisation.

⁸⁴³ Cf. Linters (1985), p. 38.

⁸⁴⁴ Cf. Linters (1985), p. 39, Mahaim (1906) and NMBS (2001), p. 11.

⁸⁴⁵ Which, in the present case, they would not be because there was no antitrust legislation in Belgium in the 19th century, and even if there had been, it certainly would not have covered behaviour authorised by statute.

⁸⁴⁶ In fact, Pollock (1913) discusses a case from "1299 or 1300" that fined the members of a cartel of candle makers in Norwich. Regular law on restraint of trade is of a much later date, particularly in Continental Europe.

⁸⁴⁷ Eg. Martin (1993), chapter 6.

⁸⁴⁸ Eg. Williamson (1985), p. 278.

⁸⁴⁹ Cf. NMBS (2001), p. 7: "Op economisch vlak gaf de zeer vrije wetgeving op de concessies niet het gewenste resultaat: dubbele uitbatingen, gebrek aan eenheid in de tarieven en moeilijkheden in de organisatie van het spoorverkeer."

⁸⁵⁰ Cf. Devys (1910), p. 27.

After the turn of the century, there were only a few private railway companies left. In the following years, they, too, were purchased by the state.⁸⁵¹ Afterwards, all that was left was to reorganise them into a more or less autonomous company. To some extent, that was done immediately upon acquisition, when the bureaucracy of the state railways would assume control of one more route, but only in 1912 did the government start to consider the possibility of protecting the railways against the worst forms of patronage by placing them at some distance from the direct control of parliament.⁸⁵² They also came to appreciate some of the other disadvantages of state ownership, such as the lack of flexibility in adopting new innovations, such as electrical traction or metal carriages.⁸⁵³ Moreover, the government quite simply needed the money; especially given the damage done in World War I.⁸⁵⁴ In 1926 a new company was created: NMBS/SNCB.⁸⁵⁵

In many ways, this temporarily private NMBS – the state bought back the NMBS shares that were in private hands shortly after the war – was a precursor of the private railway companies of the 21st century. It was vertically separated, with the state staying on as the owner of both the infrastructure and the rolling stock.⁸⁵⁶ It was given the task of operating the railway system without running a deficit, at least none that would have to be made up by the state.⁸⁵⁷ However, more than the contemporary railway companies, the state continued to have at least the option of getting involved. Since the manner in which semi-public railway companies are to be controlled by the state is a topic that we will return to later, it is useful to take a moment to consider the governing bodies of the NMBS in 1926⁸⁵⁸:

- The General Meeting of Shareholders, where the state had a majority of the votes, regardless of how many shares it actually held.⁸⁵⁹
- The Board of Administration (Raad van Beheer), where 18 members were appointed by the Crown based on criteria set out in the Act of 1926, and 3 members were appointed by the employees. If the Minister so desires, he may attend meetings of the Board, in which case he has a vote, and chairs the meeting.

⁸⁵¹ Cf. Linters (1985), p. 38 and p. 43.

⁸⁵² Cf. NMBS (2001), p. 9.

⁸⁵³ Cf. NMBS (2001), p. 11.

⁸⁵⁴ Cf. NMBS (2001), p. 11.

⁸⁵⁵ Nationale Maatschappij der Belgische Spoorwegen/Société Nationale des Chemins de Fer Belges.

Law of 23 July 1926, cf. Linters (1985), p. 43.

⁸⁵⁶ Art. 4 of the Law of 23 July 1926, cf. Linters (1985), p. 43 and Van Neste (1990), p. 171.

⁸⁵⁷ Cf. Linters (1985), p. 43.

⁸⁵⁸ Taken from Linters (1985), p. 44.

⁸⁵⁹ In fact, the state never held less than 50% of the shares.

- A Committee of six commissioners, appointed by Parliament.
- A Director-General, in charge of the daily management of the company, and answerable to the Board of Administration.

In addition, many matters were kept under the control of the minister or the government.

In other words, the NMBS was placed at exactly so much distance that the members of parliament and other politicians would no longer interfere in the detailed management of the company, but no more than that. To call this reform a privatisation would be overstating the case. Instead, the Belgian railways continued to be public in every way except legal, culminating in the completion of the nationalisation process in 1958, when the last private railway in Belgium was purchased by the state. From the 1950s until the 1990s, the NMBS continued almost unchanged in the organisational structure set up by the Act of 1926, albeit with the state as its only shareholder.⁸⁶⁰

3.3.3 *The Netherlands*

Around the turn of the century, the Dutch railway industry had been consolidated to the point that there were only two companies left: the HSM and the SS, whose routes were divided between them in such a way as to maximise the opportunity for competition between them.⁸⁶¹ At a time when both Germany and Belgium had all but completely nationalised their railway industries, the Netherlands continued to have a fully private railway sector, except for the state ownership of a significant part of the infrastructure. Predictably the question was posed whether it might not be better for the two companies to be merged into a single railway company, be it public or private.⁸⁶² However, no further changes were made until World War I.

In 1908, the government established an official commission of inquiry to examine “whether, and if so how the current policy concerning the railways was to be changed”.⁸⁶³ For

⁸⁶⁰ The company was given greater autonomy by art. 155-168 of the Act of 21 March 1991 reforming certain economic public corporations. (Staatsblad 1991, p. 6155.) Cf. section 3.4, below.

⁸⁶¹ Also known as the “balancing system”, cf. Jonckers Nieboer (1907), p. 208. Part of this system was also that the government tried to treat the two companies as equally as possible, even though the state owned a much greater proportion of the railroads used by the SS than of those used by the HSM. Cf. Jonckers Nieboer (1907), p. 208-214.

⁸⁶² Cf. Jonckers Nieboer (1907), p. 227-229.

⁸⁶³ Royal Decision of 8 September 1908, No. 25.

a number of reasons, this is an interesting document. First of all, the system in place at the time combined competition and state involvement, without elevating either to the level of dogma. Secondly, compared to the 1836 Falck rapport⁸⁶⁴, or even the 1881 rapport that paved the way for the 1890 reform⁸⁶⁵, significant learning had taken place. Thirdly, unlike the governments of Prussia and Belgium, the Dutch government did not depend on the railways for fiscal or political reasons; it could afford to approach the problem relatively objectively.

The resulting rapport, published in 1911, focused mostly on the effectiveness of the governance structures established by the arrangement of 1890. In many respects, that system was found to perform as intended: as one would expect, the two companies alternated somewhat between competition and cooperation, but neither mode dominated structurally to the detriment of overall efficiency. A number of the concerns that had been expressed in 1890 turned out to be justified. Most importantly, the companies inflated their profits by depreciating at too low a rate. In their contract with the state, the state had only stipulated a minimum rate of depreciation for the rolling stock, not for the infrastructure. As a result, the committee calculated that, adding together the financial accounts, properly prepared, for all three players in the industry, there had been a decline in equity of 45 million guilders during the intervening 20 years.⁸⁶⁶

Another point of concern was the practical functioning of the system of shared use, which had been created in order to allow the companies to have equal access to certain key locations, such as the port of Rotterdam. Instead of allowing one of the companies to derive significant market power from its control over such locations, the agreement of 1890 gave control to one company, while allowing the other certain pre-defined rights of access and use. In practice, this turned out to be a difficult arrangement. Problems included passenger confusion – particularly before the entry into force of the unified railway tariff system on 1 January 1911⁸⁶⁷ – as well as delayed goods trains, “particularly during the period of very sharp competition.”⁸⁶⁸ The commission recommended replacing shared use with mere running powers whenever feasible.⁸⁶⁹

⁸⁶⁴ Cf. section 3.1, above.

⁸⁶⁵ Cf. Jonckers Nieboer (1938), p. 175-184, Van den Broeke (1989), p. 30 and Veenendaal (1998), p. 75-76.

⁸⁶⁶ Rapport of the Commission of Inquiry, dated 24 May 1911, p. 48-73. Cf. Jonckers Nieboer (1938), p. 240.

⁸⁶⁷ Cf. Van den Broeke (1989), p. 37 and Veenendaal (1998), p. 79.

⁸⁶⁸ Rapport of the Commission of Inquiry, dated 24 May 1911, p. 6-8.

⁸⁶⁹ Rapport of the Commission of Inquiry, dated 24 May 1911, p. 8.

On the other hand, the fear originally expressed that the companies might “hold up” the state by threatening to use their right to cancel the contract – forcing the state to nationalise – turned out to be unwarranted.⁸⁷⁰ However, this is the only Williamsonian element in the Commission’s rapport. There seems to be no recognition of any “fundamental transformation”, not in the least because the rapport only considers reforms in the direction of more state involvement, not less. Going from four parties (three railway companies and the government) to three, as the Dutch did in 1890, is hardly a “fundamental transformation”. Neither then, nor in 1911, was there any possibility of bringing in more parties, in order to create some kind of competition for the market. Two of the most notable ingredients that were missing are sufficiently liquid capital markets, necessary to quickly raise the money in case one’s bid for a concession is accepted, and a system of remedies to protect the lucky winner from being held up by the state.⁸⁷¹ Without the possibility of awarding the concession to another player, the parties are as condemned to work together before they agree the contract as afterwards, meaning that Williamson’s framework does not apply.

The question of how the railways respond to the possibility of *ex post* government regulation of their business is a constant factor in the present day railway debate. As we have seen, this concern was already voiced in the discussions surrounding the UK Railway Act of 1844. In the 1911 rapport, however, this quintessentially GHM problem received relatively little attention. Again, the reason for this oversight is that the authors focused on the possibility of giving the government more power over the railway industry, not less. To the extent that the possibility of government regulation acted to distort the sector’s incentives to invest, this would have to be remedied with more detailed regulation.⁸⁷²

This does not mean that the commission did not consider transaction costs at all. On the contrary, they did little else. In chapter one of their rapport, they evaluated the government’s attempt to write an efficient contract with the railway companies, highlighting the places where the contract was incomplete, as well as the places where the contract created

⁸⁷⁰ Rapport of the Commission of Inquiry, dated 24 May 1911, p. 74.

⁸⁷¹ Remember, it is not the state but the railway company who has the most obvious investment in transaction-specific assets.

A system of administrative law was almost entirely absent in the Netherlands until well after World War II. Cf. Van der Pot & Donner (2001), p. 504-506. Instead, the agreement of 1890 gave the railway companies one very drastic remedy in case of state overreaching: as noted in the main text, they could cancel the contract, forcing the state to take over the railway industry. Cf. Rapport of the Commission of Inquiry, dated 24 May 1911, p. 50.

⁸⁷² Cf. for example Rapport of the Commission of Inquiry, dated 24 May 1911, p. 81.

perverse incentives.⁸⁷³ The second chapter discussed the various non-financial objectives the government would have when dealing with the railway industry, such as safety, employee rights and the interests of the military. Again, the predominant question was contractual: to what extent can the government achieve its goals through contracting or legislating? Chapter three made up the accounts for the entire sector, discussing in the process whether it would be possible for the government to acquire this kind of information under the existing contractual arrangements.

Of course, for lack of anything to compare with, the members of the commission were unable to agree on one clear recommendation.⁸⁷⁴ Then again, for the present thesis, the outcome of their analysis is less interesting than the manner in which they came to it.⁸⁷⁵ And the members of the commission came to their conclusion by considering, as Coase did twenty years later, what the limits are of what can be done using contracts on the marketplace. Their analysis does not, however, anticipate the more advanced modelling done by later authors building on Coase's work.

Like the Commission of 1908, Dutch politics was divided on the matter of further consolidation and nationalisation, and remained so until World War I made it necessary for the state to take over the railways for the national defence. Plans for such an event had been ready since the 19th century⁸⁷⁶, but during the war they were amended on several occasions.⁸⁷⁷ Most importantly, in 1916 the SS and the HSM agreed to merge their operations, forming a joint venture allowing the railway industry to be operated as a single enterprise, albeit with a number of signs of the old system still in place.⁸⁷⁸ Parliament approved the joint venture with rela-

⁸⁷³ Remember, when a contract is insufficiently detailed, it is incomplete in the sense of being insufficiently state contingent. Cf. section 2.1.4, above. Sometimes, such incompleteness leads to perverse incentives. In other cases, the perverse incentives are caused by the contract simply being wrong, meaning that parties – perhaps irrationally – left gains from trade on the table.

⁸⁷⁴ Five members voted for immediate nationalisation, two members supported gradual transition towards state control, five members voted to continue with the system of 1890 and four committee members wanted a single private railway company. Considering individual proposals, nationalisation was rejected 10-5, and operation by a single railway company was accepted 8-7. Rapport of the Commission of Inquiry, dated 24 May 1911, p. 97-98. Cf. Jonckers Nieboer (1938), p. 240-241.

⁸⁷⁵ Cf. section 1.3, above.

⁸⁷⁶ They were laid down in the Spoorwegwet (Railway Act) of 1875. Cf. Rapport of the Commission of Inquiry, dated 24 May 1911, p. 43-47.

⁸⁷⁷ The organisation of the Dutch railways during World War I is discussed in Jonckers Nieboer (1938), p. 243-254, Reitsma (1940, p. 774, Van den Broeke (1989), p. 38-40, Veenendaal (1998), p. 104-105 and Veenendaal (2004), chapter 15.

⁸⁷⁸ The new organisational model was introduced starting 1 January 1917. Cf. Veenendaal (2004), p. 323-326.

tively little debate, based on the same arguments used by the Commission of 1908, as well as the subsequent experience of three peacetime years with little actual competition and two war-time years of fruitful collaboration.⁸⁷⁹ After the war, the legal reality of the railway industry was gradually changed to match the practical reality, until on 2 August 1937 the NV Nederlandse Spoorwegen was founded.⁸⁸⁰

3.3.4 France

By decree of 25 March 1878, seemingly unimportant at the time, the government of France created a proper administrative body to administer a number of railways in the western Poitou-Charentes and Vendée regions that had come back under state control due to financial difficulties. Given that parliament had rejected the possibility of selling these companies to the Paris-Orléans, the government had little alternative.⁸⁸¹ Nevertheless, the entirety of the French state railway system, the SNCF⁸⁸², would grow out of this *Réseau de l'État*.

During the first 30 years of its existence, the *Réseau de l'État* continued to be relatively unimportant, compared to the big six, totalling just 2615 km of local railway in 1878, and growing little afterwards.⁸⁸³ At the beginning of the 20th century, however, the ever increasing financial distress of the *Ouest* in particular coincided with a steadily changing political mood towards nationalisation. The *Ouest* had always been the weakest of the big six, focusing as it did on the relatively unrewarding markets of Southern Normandy and Brittany. Politically, throughout the history of the Third Republic there was almost always a centre-left majority. The composition of that majority, however, gradually shifted from anti-Bonapartist Republicans to social-democrats as we know them today. To the extent that these social-democrats were willing to work together with the *étatist* conservatives on the right, this meant a gradually increasing support for the nationalisation of all sorts of services of general interest. This happened not as a result of any new arguments being advanced, but simply because an overall ideological shift, not dissimilar from the ideological shift in the rest of Europe around the same time.

⁸⁷⁹ Another noteworthy argument was the experience of other countries, including the USA, with consolidated and nationalised railways. Cf. Veenendaal (2004), p. 323.

⁸⁸⁰ Cf. Veenendaal (2004), p. 350.

⁸⁸¹ Cf. section 3.2, above and Neiertz (1991).

⁸⁸² *Société Nationale des Chemins de fer Français*.

⁸⁸³ Cf. Caron (1997), p. 476.

By 1902, the socialist parliamentarian Jean Bourrat had managed to persuade the *Assemblée Nationale* to adopt a resolution proposing the purchase by the state of the *Ouest* and the *Midi*, the two companies that were underperforming the most.⁸⁸⁴ This partial nationalisation was defended not only as a pragmatic compromise between those looking to nationalise the entire industry and those who resisted nationalisation, but also as a chance for the state to develop the skills necessary for nationalisation.⁸⁸⁵ While this first resolution was not followed by any government action, a second resolution that was passed in 1906 resulted in the purchase of the *Ouest* by the state in 1908.⁸⁸⁶

So what exactly was the problem that this nationalisation was meant to solve? Well, to sum it up in a single sentence: the Big Six were too big to fail, and they knew it. During bad economic times and in cases of mismanagement the government bailed them out, without the ability to attach any significant consequences to that bailout. There was no pressure forcing the companies to be particularly customer friendly, to run the trains on time, or care about quality at all except to the extent that this had a direct influence on company profit. In a period before the long-distance transportation of goods and passengers by automobile became a real possibility, the connection between quality and profit was generally weak, while in periods of government bailout, the connection was entirely non-existent. In his programme of government following the 1906 elections, Clemenceau supported his plan to nationalise the *Ouest* by complaining about the delays and about the railways' chronic shortage of rolling stock, which "severely hurt the interests of commerce and industry".⁸⁸⁷ To the extent that a financial solution for the industry generally and for the *Ouest* in particular was impossible, he said, it was nothing but a "régie désintéressée", a public company without a stake in the outcome of its activities, and should be organised accordingly.⁸⁸⁸

In the following years, the debate continued. Notably, in 1913 the newspaper *Le Matin* made a first attempt to compare the efficiency of the railways in different countries using hard data.⁸⁸⁹ While their approach still had many flaws, this was certainly a step forwards. Interrupted by the war, this debate continued only in 1919. During the war itself, the French rail-

⁸⁸⁴ Cf. Caron (2005a), p. 360 and Carron (2005b), p. 288-307.

⁸⁸⁵ Initially, the *Réseau de l'État* had been run just like any other railway company. Starting in 1882, its administration was gradually changed to the normal state way of doing things. Cf. Caron (2005a), p. 362. However, even as an integral part of the government apparatus, the *État* was still too small to make the state ready for full nationalisation.

⁸⁸⁶ Cf. Caron (2005a), p. 361 and 365.

⁸⁸⁷ Caron (2005a), p. 365.

⁸⁸⁸ Caron (2005a), p. 365.

⁸⁸⁹ Cf. Caron (2005b), p. 325.

ways were under the direct control of the War Ministry and the Army.⁸⁹⁰ Unlike in other countries, however, this did not cause a permanent change in thinking, at least not about the manner in which the industry should be organised.⁸⁹¹ After the war, each company again resumed full control over its network, while the government tried to calculate how much it owed each of them for the use of their facilities and for the damage done thereto.⁸⁹² The only difference the war made is that the state now acquired control over the railway system in Alsace-Lorraine, which the *Compagnie de l'Est* was not interested in taking back.⁸⁹³ However, the AL was kept separate from the *État*, understandably so because they were active on opposite ends of the country.

When the dust settled in 1921, it became clear that the railway companies would need the government's help in order to recapitalise.⁸⁹⁴ The sheer amount of damage done to the network would have posed problems anyway, and the French railway industry was far from financially healthy even before the war. The price they paid for the state's assistance was a further step towards nationalisation. The agreement of 1921 created a *conseil supérieur des chemins de fer*, a high council for the railways, consisting of representatives of the railway companies and of the state, which took over a lot of control from the boards of the individual companies. In addition, the agreement established a railway fund, which would act as a pool of the profits and losses of the railway companies, paying out to cover losses, receiving money in case the company made a profit. Essentially, this meant that the shareholders of the private railway companies were forever assured of their guaranteed return on investment, without the chance of receiving either less or more.⁸⁹⁵ The railways became public in all but name, without any need for the government to purchase the industry outright.⁸⁹⁶

This first bailout was enough to return the industry to some semblance of solvability and profitability, at least for the remainder of the decade. When the depression hit, however,

⁸⁹⁰ Cf. Caron (2005a), p. 533-634 for a discussion of the French railways during World War I.

⁸⁹¹ Cf. Caron (2005a), p. 606-609.

⁸⁹² There was a commission consisting of the secretaries-general of the big five that studied the possibility of nationalisation, but they could not agree on the terms the state would have to offer. Other proposals made in parliament similarly encountered financial difficulties. Cf. Caron (2005a), p. 687-690. For a discussion of the approach to calculating the war damage, cf. Caron (2005b), p. 341-355.

⁸⁹³ Decree of 19 June 1919.

⁸⁹⁴ Cf. Ben Sassi (2008), p. 65, Caron (2005a), p. 639-686 and Caron (2005b), p. 341-372.

⁸⁹⁵ Cf. the last entry in the list of expenses that the companies were allowed to deduct from their revenues for the purpose of calculating their contribution to the fund, Caron (2005b), p. 373.

⁸⁹⁶ Other terms included the closure of 11.000 km of railway, the simplification of many rates, and the transfer of billions of francs worth of war-related debt back to the state. Cf. Ben Sassi (2008), p. 65, Caron (2005a), p. 693-702, and Caron (2005b), p. 373-383.

the companies once again found themselves writing deep red. By 1937, only the Alsace-Lorraine was still profitable. After years of trying to solve the shortages in the fund by increasing the rates and putting yet more government money in the sector, by 1937 the state⁸⁹⁷ and the railway companies agreed to finally complete the process of nationalisation and consolidation: effective January 1, 1938, they created the SNCF, a public limited company which controlled and operated the entire French railway industry. The state owned 51% of the shares, and the railway companies the other 49%.⁸⁹⁸ The private companies' share was to be successively decreased over a 45 year period. The board of the new company was to consist of 33 members, 21 of which would be designated by the state.

3.3.5 Great-Britain

As we have seen, in the second half of the 19th century, the logic of the market drove British railway companies inexorably towards consolidation, only restrained by the reluctance of Parliament to approve such mergers combined with its unabated willingness to approve new schemes proposed either by new companies or by companies previously operating in a different part of the country. For this reason, Britain continued to have a highly diverse railway industry, with giants as well as dwarves, until the Great War.⁸⁹⁹

During World War I, the British railway sector was controlled by the state, as it was in the rest of Europe.⁹⁰⁰ All of a sudden, the excess capacity that had plagued the profitability of particularly the southern railway companies for so many decades proved to be fortuitous, since it allowed more goods and personnel to be transported to the southern ports, with lower vulnerability to breakdowns at any part of the network.⁹⁰¹ At the same time, the shortcomings of the old system were suddenly elevated from hypothetical to very real. Wolmar (2007) summed it up as follows:

⁸⁹⁷ In 1935, the *Front Populaire* had been formed, a collaboration between the three main left-wing parties. In 1936-37, they formed the first socialist-led government of the Third Republic, under the leadership of Léon Blum.

⁸⁹⁸ Cf. Ben Sassi (2008), p. 65.

⁸⁹⁹ Cf. Arnold & McCartney (2005). By the start of the war, there were about 120 railway companies, cf. *ibid*, p. 41. Wolmar (2007), p. 224-225 mentions the number of railway companies as being 178 in 1914.

⁹⁰⁰ Regulation of the Forces Act (1871), 34&35 Vict., c. 78. It should come as no surprise at this point that the act allowing for state takeover of the railways in time of war was passed just after the Franco-Prussian war.

⁹⁰¹ Cf. Wolmar (2007), ch. 11, for a discussion of the British Railways during World War I.

Under state control, the British railway network had been unified as never before and in many ways operated far more efficiently than before the war, carrying more traffic with fewer workers and less rolling stock. The pooling of wagons meant they no longer needed to be returned empty to their place of origin, locomotives were transferred to lines short of motive power and the duplication of services on parallel lines was reduced. (...) ‘The lesson, that substantial economies could be obtained by getting away from traditional independence and operating the railways more or less as a single system, was learnt both by railway managers and outside observers.’⁹⁰²

A second impetus for post-War change came, as in the other countries, from the financial health of the sector. Compared to the French railway industry in particular, the British railways had generally been fairly profitable, with the notable exception of the companies connecting London with Sussex and Kent. However, the state was never going to be willing or able to fully compensate the industry for its services during the War.⁹⁰³ Combined with the inevitable post-War recession, this meant that many companies were in sufficient financial difficulty to welcome state intervention, or at least not oppose it.

What the British ended up with in 1923 was essentially the French system of 1859, with some aspects of the Dutch system of 1890 thrown in. The network was divided among four companies: The Great Western, the London & North Eastern, the London, Midland & Scottish and the Southern.⁹⁰⁴ As in France, all four companies reached London, each still using the terminals their predecessors had built in the 19th century.⁹⁰⁵ There was, however, one notable exception to this territorial principle: The east coast and west coast main lines connecting England and Scotland, which had been the subject of intense competition in the 1890s⁹⁰⁶, continued to belong to different companies.⁹⁰⁷ Most likely, though, this was less the result of any lingering desire to maintain some semblance of competition than the inevitable

⁹⁰² Wolmar (2007), p. 224, quoting Michael Bonavia (1980), *The Four Great Railways*, David & Charles, p. 9. Note that the Railway Clearing House, which was discussed above, allowed for the participating companies to have certain running rights on each other’s networks, but not for the level of integration achieved during World War I.

⁹⁰³ The railway companies were not entirely powerless. The North British Railway company, the largest railway company in Scotland, claimed an initial installment of £ 616.194. When the government offered £ 186.194, the company took the government to court. In the end, the state ended up paying just short of £ 10 m. Cf. Hamilton Ellis (1955), p. 211, Hamilton Ellis (1959), p. 305 and Wolmar (2007), p. 223.

⁹⁰⁴ Railways Act (1921), 11&12 Geo. 5, c. 55.

⁹⁰⁵ St. Pancras station, where the Eurostar currently arrives, was built by the Midland in 1868, similarly, the Euston Street station was since 1837 the London home of the London & Birmingham and later the London & North Western, Paddington (1854) belonged to the Great Western, Charing Cross (1864) belonged to the South Eastern and King’s Cross to the Great Northern. In 1923, the first two of these stations came to belong to the London, Midland & Scottish, and the others, respectively to the Great Western, the Southern and the London & North Eastern.

⁹⁰⁶ Cf. Wolmar (2007), chapter 9 and the footnotes accompanying the earlier discussion of the British railways.

⁹⁰⁷ To the London & North Eastern and the London, Midland & Scottish, respectively.

consequence of having four companies all reaching London: the only way to allocate the entire Scottish railway network to a single company would have been to create a company which, like the *Midi* in France, did not reach the capital.

During World War II, the railways were once again brought under state control, and heavily used. Of course, unlike the previous war, World War II also involved heavy enemy bombardment of key infrastructure, including the railways, meaning that the railway infrastructure suffered even more damage than during World War I. Under the post-war Attlee cabinet, the government embarked on an ambitious project of nationalisation, covering such industries as coal mining, health, telecommunications and road haulage.⁹⁰⁸ Naturally, the government also set its eyes on the railways. The Transport Act of 1947⁹⁰⁹ provided the necessary legal framework, and on 1 January 1948, the British Transport Commission, through its Railway Executive, became the proud owner of the British railway industry.⁹¹⁰

3.3.6 Conclusion

Transactions: Given that we have ruled out transactions occurring within an organisation, a fully nationalised railway industry must involve no transactions whatsoever. That is fine, of course, as long as the models still work in those circumstances. So far, that does in fact seem to be the case. The key for seeing this is Williamson's insight that the goal is to compare "discrete structural alternatives".⁹¹¹ Clearly those alternatives cannot all involve no transactions whatsoever; there is only one way to fully integrate the value chain. This "full integration" option will necessarily be compared to alternatives that involve at least one transaction.

Reality, of course, offers more nuance still. In chapter 2, we attached the notion of a "transaction" to the legal realities of the situation. A transaction is a transfer of an asset from one legal person to another, normally for consideration. Looking at the legal realities created in various European countries that allegedly represent "full" nationalisation, it turns out that none of them opted for something that completely eliminated all transactions as we have defined them. In all countries under consideration, the railways were owned by some kind of separate legal entity, controlled by the state but still at arm's length. Nowhere did the state

⁹⁰⁸ Cf. Hennessy (1987), p. 37.

⁹⁰⁹ Transport Act (1947), 10&11 Geo. 6, c. 49.

⁹¹⁰ Wolmar (2007), chapter 13.

⁹¹¹ Williamson (1991).

own the railways the way it owned the parliament buildings, for example. As a result, the era of state ownership still had transactions.

The terms of the laws that governed those transactions were, of course, fixed unilaterally by the state, not negotiated at arm's length. However, to leave it at that would be to underestimate the freedom of manoeuvre of the officials tasked with managing the nationalised railway bureaucracies. This is a point that has been emphasised particularly by Grossman, Hart and Moore: ownership of physical assets is only one aspect of integration. In other words: as long as there is a significant amount of human capital necessary to operate those assets, there are still transaction costs.

Overall efficiency: Compared to the period discussed earlier, in most countries the period of nationalisation meant a reduction in the range of options under consideration, back to the kind of reasoning used when the railways were first invented. Just like nationalisation was literally unthinkable then, at least everywhere other than in Belgium, so too was it now all too often presented as the only alternative.

On the other hand, as in the previous period, choices often seemed to be driven by preferences that could not reasonably be described as government preferences. Bismarck's use of railway rates as a substitute for import levies may have been misguided, but it reflected a legitimate government policy choice, based on legitimate government preferences. The same cannot necessarily be said of the full nationalisation of the German railway system after World War I, which reflected a complicated interplay of post-war German and European politics.⁹¹²

Costly Information and Moral Hazard: As we have seen, nationalisation was often a response to the realisation on the part of the government and the general public of the true extent of the moral hazard involved in the state's involvement in the sector. For decades, authorities had pumped vast amounts of money into the industry, only to find it wanting more. Initially, the response was to professionalise and bureaucratise the state's railway authorities, but when that proved insufficient, moral hazard became a powerful argument for nationalisation.

Fundamental transformation: Comparing the "distinct structural alternatives" of a nationalised industry and some earlier form, the former is going to be characterised by the total absence of

⁹¹² Note also that nobody seems to have considered the possibility of privatising the German railway system as a way of raising money for reparations and avoiding the Prussian dominance in the sector.

any fundamental transformation while the nationalised regime endures. Even individual employees will be locked into their place, as they will often acquire the legal status of civil servants, with the lifetime employment guarantees that come with it. As for the three key organisations involved – unions, the railway “company” and the government – they are forever stuck together, without any possibility that any one of them might be replaced by a competitor. This rules out any model that relies on competition for the market as a necessary element. Why would the City of Oakland have to worry about companies submitting unrealistically low bids in order to win the tender if there was only one company bidding?⁹¹³ On the other hand, how can the city persuade the company’s management to provide it with realistic information about the true likely cost of the project if those people are essential for the completion of the project, while the city lacks the in-house knowledge necessary to make the calculations? The solution lies in negotiating, i.e. in game theory and reputation mechanisms⁹¹⁴, not in competition for the market.

Asset Specificity: There is nothing new to say about asset specificity at this point. The question of which assets are specific to the transaction does not change over time, which is problematic given that the model is meant to explain choices in governance structure which certainly did change over time.

Uncontractible investments: The most important category of uncontractible investments that a nationalised railway company might have problems with is the staff’s investment in customer service, i.e. in quality improvement. Nationalised companies have notoriously poor quality, which in the case of railway companies means poor customer service and poor punctuality above all. As long as there is some competition for the market, the railway company can be punished for its poor quality by being replaced. Moreover, as long as the railway company is a for-profit entity, the staff can be given bonuses connected to the company’s profit. In a nationalised company, it is much more difficult to incentivise the employees in this way.

Alignment Mechanisms: In addition to the evolutionary and competitive pressures already discussed before, a new kind of competition emerged for the first time in the early 20th century: intermodal competition, i.e. competition between different modes of transport. Ever

⁹¹³ Cf. Williamson (1976).

⁹¹⁴ Cf. Dixit (2004).

since the railways put the carriages and the “trekschuit”⁹¹⁵ out of business, and severely hurt the business of the canals for goods transport, the railways had functioned essentially in a vacuum, without any threat from a competing product. This changed when, from the 1860s onwards, the basic principles of electric and internal-combustion motors were invented. Once cars and buses had become commonplace, life would never be the same for the railways. The advantage of being able to run trains on electricity or diesel fuel instead of coal⁹¹⁶ was not enough to make up for this: the railways were losing some of their near-100% market share. From then on, any time the railways failed to get things right, they would be punished with a loss of market share in the transport market. Of course, among the things that are important to “get right” are the governance structures that govern the collaboration between the state and the railways.

Also, there were two major wars that forced a re-evaluation of existing practice where this would otherwise have been put off. While it would be incorrect to say that the choices made in war time must reveal the “true” optimum, free of rent-seeking interference⁹¹⁷, it certainly opened many people’s eyes to the merits of consolidation and nationalisation. For this reason, as much as because of the damage done by the war, both post-war periods proved fruitful times for careful design of governance structures. That is not to say that the choices made were necessarily correct, but it does mean that the theoretical optimum is somewhat more likely than any other possible outcome because the problem was considered by wise men working with as near a blank slate as it is possible to get.

Unfortunately, it is difficult to say what that theoretical optimum would have been. Uncertainty had continued to fall, as it has done continuously from the beginning of the railways to the present day. On the other hand, more than ever before the state was looking to promote a variety of different goals through the railways. They wanted to promote national unity, protect the working classes, promote passenger safety, protect their national industry against foreign competition, etc., etc. We may interpret the success of the liberalisation and privatisation efforts of recent decades as evidence that the governments of decades past were wrong to nationalise the industry in the first place, but as we will see in the next section this

⁹¹⁵ A trekschuit or tugboat is a mode of transport of people on canals used particularly in the Netherlands in the 17th – 19th centuries. The boat would hold up to 30 passengers and would be pulled by horses along a canal. The service was more comfortable than a coach, but also slower.

⁹¹⁶ Cf. Caron (2005a), p. 890-907, Veenendaal (2004), chapter 18 and Wolmar (2007), p. 191 and 235-240.

⁹¹⁷ During a war, the state’s priorities are different. As a result, quality is much more important than price, pointing to more hierarchy than in peace time. Cf. Hart, Shleifer & Vishny (1997), proposition 4 on page. 1142.

withdrawal by the government was only made possible by a still further decrease in the cost of information, as well as a decrease in the range of policy goals the government was concerned about.

3.4 Privatisation: Thatcher and Beyond

After the period of nationalisation was completed in the immediate aftermath of World War II, the countries of Europe were generally content to leave their railways in public hands. During this period, the process of electrification was completed, banishing the remaining steam locomotives to the railway museums.⁹¹⁸ While such technological changes can at times have consequences for governance choices, in the way that the emergence of satellite television and the internet has changed the analysis of the cable television industry discussed in section 2.4.3 above, they are not themselves the subject of this thesis. Instead, the goal is to identify the extent to which decision makers were motivated by expected transaction costs when reforming the railway industry. It follows that we should take up the story again with the election of Mrs. Thatcher as Britain's Prime Minister in 1979.

Subsequently, privatisation spread like wildfire across the continent. However, it was a very peculiar kind of wildfire, one that burned with much more enthusiasm in some places than in others. How is this possible? Well, it is possible because the proponents of some measure of privatisation came to outnumber the opponents in the EU's legislative bodies, the Council of Ministers and the European Parliament. These bodies passed a series of railway packages, which forced even those who were otherwise unconvinced to at least pretend to make an effort.⁹¹⁹

Because of this historically unusual history, this final section will focus mostly on the early developments in Britain and the history of the EU railway packages, leaving the developments in the national political arenas of France, Belgium, the Netherlands and Germany largely by the wayside. The idiosyncrasies of each country's policy choices and – more importantly – the arguments that underlie those policy choices can be understood in isolation now even less than before. The interactions between the policy discourse at the European and national levels is a fascinating field of study, but one that lies outside the scope of this thesis.⁹²⁰ Moreover, the sheer size of this discourse has grown enormously compared with the

⁹¹⁸ Cf. Veenendaal (2004), chapter 18.

⁹¹⁹ Cf. Communication from the Commission dd. 31.03.1998, COM(1998) 202F, on the Implementation and impact of Directive 91/440/EEC on the development of the Community's Railways and on access rights for rail freight.

⁹²⁰ Some literature is discussed in Nugent (2006), chapter 21.

time of nationalisation, a factor that should caution against too much desire for completeness for practical reasons alone.

3.4.1 Great-Britain

Section 3.3 concluded with the unexpected election victory of the Labour party under Clement Attlee in 1945, and the programme of nationalisation upon which his government embarked.⁹²¹ This led, amongst other things, to the enactment of the Transport Act of 1947⁹²², which created a British Transport Commission whose Railway Executive became the legal owner of the British railways.⁹²³ Even though the Tory party returned to government in 1951, Attlee's nationalisation of many industries was not reversed for many decades.⁹²⁴ This explains how Mrs. Thatcher came to inherit a system of economic governance that was – compared to Britain as we have come to know it in this thesis so far – extremely *dirigiste*.⁹²⁵ The Labour governments of Harold Wilson and James Callaghan had tried to tackle the crisis of the 1970s with such innovations of governance as the National Enterprise Board, which was meant to soften the blow of industrial bankruptcies and increase public ownership of industry.⁹²⁶ By the end of the 1970s, the failure of this approach became increasingly clear, even to the Labour party, since it failed to produce either economic growth, employment growth, or reduce inflation. This meant that the Conservative election victory of 1979 could reasonably be interpreted as providing them with a mandate to take a radically different course, consistent with their manifesto.⁹²⁷

Of course, in politics you can overdo it with the radical change. Even when the electorate is “suffering both from an economic crisis and a crisis of morale”⁹²⁸, this must be attacked with a sense of proportion, an understanding of how much real change the country and the government bureaucracy can handle. For that reason, the reforms proposed in the 1979 Conservative election manifesto, and the reforms implemented in the first years subsequent, were actually fairly modest compared to the reputation that Mrs. Thatcher has since acquired.

⁹²¹ Cf. Hennessy (1987).

⁹²² Transport Act (1947), 10&11 Geo. 6, c. 49.

⁹²³ Cf. Wolmar (2007), chapter 13.

⁹²⁴ Indeed, the National Health Service continues to exist to this day.

⁹²⁵ Cf. Yvrande-Billon & Ménard (2005), p. 682-684.

⁹²⁶ Cf. section 2 of the Industry Act (1975), c. 75. Cf. also Whitehead (1987).

⁹²⁷ Cf. Vincent (1987) and Riddell (1989).

⁹²⁸ 1983 Conservative Party General Election Manifesto.

Initially, the change was more one of attitude than of actual governance structures.⁹²⁹ To be sure, the process of privatisation and liberalisation started almost immediately, but without particularly ambitious goals. During the first Thatcher government, receipts from privatisation never exceeded £ 500m⁹³⁰ and the companies that were sold were generally those that only came into the government's hands because of bailouts and other assorted interventions, rather than because of deliberate government policy.⁹³¹ A notable early exception was Cable & Wireless, originally Britain's major telegraph company, which had come into the government's hands in 1947 as part of Attlee's nationalisation programme.

Looking at the reasons given for the Tory Party's enthusiasm for privatisation, we see an interesting new argument: Popular Capitalism.⁹³² The government wanted to encourage widespread share ownership, and one of the instruments it used to accomplish this was to privatise large government-owned corporations, and to privatise them in a manner that encouraged widespread share ownership. The benefits of popular capitalism were often phrased in terms of giving employees a stake in their companies, which would presumably make them less likely to go on strike and more likely to work hard.⁹³³ However, it seems clear from contemporary political statements and later academic analysis that Mrs. Thatcher and her cabinet colleagues considered popular capitalism to be something that was a *bonum in se*, a good thing that did not require any further justification.⁹³⁴ They emphasised the parallels between share ownership and home ownership, often simply talking about "ownership" *tout court*. From the point of view of economics, this is a curious way to elevate a governance structure to a goal as such.

⁹²⁹ Note that, because modern NIE models tend to incorporate game theory to model negotiations, credible commitments are an important part of the story. Cf. for example Williamson (1983). In politics, mere statements can be "credible" in a way that they would not be in an ordinary game theoretic analysis, because deviation tends to be punished at the polls. In that light, a mere change in attitude already goes a long way towards obtaining better negotiating outcomes for the government.

⁹³⁰ Cf. Riddell (1989), table 5.1 on p. 92.

⁹³¹ British Aerospace, formed in 1977 by the Aircraft and Shipbuilding Act 1977 (c. 1) as a consolidation of various government holdings in the aircraft industry, was turned into a plc under the British Aerospace Act 1980 (c. 26). 51,57% of the shares were sold in February 1981, and the remainder in 1985. Similar stories can be told about the pharmaceutical company Amersham (privatised in 1982), about the oil companies Britoil (1982 and 1985) and Enterprise Oil (1984), as well as car companies Jaguar (1984), Rolls Royce (1987) and Rover. Cf. Riddell (1989), p. 87-88.

⁹³² Riddell (1989), chapter 6.

⁹³³ Quoting the 1983 Manifesto: "We will continue to encourage workers to identify with the success of the firm for which they work, by the promotion of share-ownership and profit-sharing."

⁹³⁴ Riddell (1989) starts his chapter on the subject by describing it as the government's "moral vision". The Tory Party's 1987 calls up an image of "a capital-owning democracy of people and families who exercise power over their own lives in the most direct way." And later in the same document: "Owning a direct stake in industry not only enhances personal independence; it also gives a heightened sense of involvement and pride in British business."

British Rail, specifically, was not slated for privatisation until quite late in the day. The 1979 manifesto only said that “those industries that remain nationalised” should be interfered with less, while submitting them to “clearer financial discipline”. In 1983, they never once mentioned British Rail in the privatisation section of the manifesto, focusing on it only as part of their proposals for public transport. Regarding the governance of the railways, the Tory Party wrote: “Rail services are now facing vigorous competition from coaches and cars, and they need to respond with more innovative and more modern work methods. We shall examine ways of decentralising BR and bringing in private enterprise to serve railway customers.” In 1987 the claim is simply that the productivity in the public sector has gone up, as illustrated for example by the fact that “British Rail will cost the taxpayer 25 per cent less in subsidy this year than in 1983 and without any major route closures.”⁹³⁵ Only in the 1992 manifesto did the party – now under the leadership of John Major instead of Margaret Thatcher – commit to the ultimate railway reform: “We will end British Rail’s monopoly. We will sell certain rail services and franchise others.”⁹³⁶

The manner in which this was done was another novelty. Ever since the Stockton & Darlington’s abortive attempt at applying a toll road business model, the railways had always been vertically integrated, at least to the extent that the tracks, the stations and the rolling stock tended to be controlled – though not necessarily owned – by the same party. Legal technicalities aside, one company or government entity tended to be in control of all aspects of the problem of getting people from A to B via a given route. As we have seen, the only times the company needed to deal with others was when it wanted to sell tickets for travel to a destination it did not reach, and of course when it needed to deal with the government. Competition ensued when there was more than one way to travel from A to B, for example by using different routes or by using different modes of transport. However, the 1992 Tory plan envisaged a much more drastic vertical disintegration than anything hitherto contemplated.

Before considering the policy reasons for this choice, it is useful to consider one of the legal considerations: Directive 91/440⁹³⁷, which we will return to later. What this directive required was a degree of management and accounting separation between infrastructure and

⁹³⁵ That last qualifier was necessary because the last time a British government had tried to cut the railway budget, they ended up with Beeching’s massive track closures. Cf. Wolmar (2007), p. 282-288.

⁹³⁶ For more on franchise bidding, cf. the discussion of Williamson (1976) in section 2.4.3, above. The idea was pioneered by Demsetz (1968). Cf. also the definitions in section 23 of the Railways Act 1993 (c. 43).

⁹³⁷ Directive 91/440 on the development of the Community’s Railways, OJ L 237, d.d. 24.08.1991, p. 25-28.

operations.⁹³⁸ That is not the same, however, as an actual separation of ownership the way the Tories proposed it, something that is not even strictly required today, though subsequent legislation does favour it.⁹³⁹ The official government documents on the situation correctly explain this fact⁹⁴⁰ and it is clear that the European legal context of the problem played little or no role in their decision to do as they did.⁹⁴¹ And yet, the government and their supporters in the media certainly did have every incentive to exploit the traditional British mistrust of the EU in order to blame them for the unpopular aspects of the reform.⁹⁴² Invoking such seemingly irrelevant arguments is part of the process of winning popular support, a problem that is not part of our current analysis.

The most obvious reform, which the Thatcher government had already announced in 1983, was to “decentralise” BR by splitting it up into regional franchises, to be awarded by tendering.⁹⁴³ Without more, that would bring the governance structures of the industry back to the situation of the pre-war “big four” era. And since they already knew that that did not work as well as it might have, this one reform would necessarily have to spark another, and another, and so on. One could imagine the dominoes of reform falling as follows: If franchising is to work, there has to be genuine competition for the market. We have already seen in Williamson (1976) that that alone is not enough, but there is no question that the difficulties in organising a franchise bidding system are exacerbated if there are too few companies bidding. The solution is to make sure that – realistically – bidding is open to outsiders as well as incumbents. That means designing a pre-bidding process that allows would-be franchisors to do their due diligence without giving them access to too much sensitive information. It also means designing the object of the franchise to be as small as possible.

The smaller the franchise – few years, few miles of track, few assets – the lower the bids will be, and the smaller the financial risk for the would-be bidder. Of course, shrinking the object of the franchise along the lines indicated will always come at the expense of operating efficiency, which is yet a new representation of a familiar problem: trading off transac-

⁹³⁸ Cf. Crozier (2001).

⁹³⁹ Cf. section 3.4.3, below.

⁹⁴⁰ Cf. par. 58 of the 1992 White Paper, which will be discussed below, and page 3 of British Rail’s official response to that White Paper.

⁹⁴¹ Cf. Lodge (2003), p. 167 and Murray (2001), p. 14.

⁹⁴² Cf. Murray (2001), p. 14. Of course, the opposite argument works just as well. Cf. *The Independent*, 22 December 1994, ‘Railtrack sale “flouts EU law”’, reporting that the RMT Union had opined that EU law required Railtrack to remain in the government’s hands.

⁹⁴³ Cf. Yvrande-Billon & Ménard (2005), p. 684.

tion efficiency and production efficiency. So it follows that the strictly economic answer is that along each of these axes – time, space and money – the size of the franchise has to be shrunk until the marginal benefit of doing so, in terms of reduced transaction costs, is exactly equal to the marginal costs, which come in the form of reduced production efficiency. Then again, it is doubtful whether that answer would be very helpful in practice.

The first most obvious way in which the financial scope of the franchise can be reduced is by vertically separating the trains from the track. If companies can bid for a rail franchise without having to bring their own track, the opportunity is open to much smaller bidders than would otherwise be the case. While vertical separation of this kind had not been attempted since the Stockton & Darlington, it had the additional benefit of preventing hold-ups between successive owners. After all, there are many places where a franchise holder can buy the rolling stock it needs to carry out the contract, but only one place where it can get the tracks. If parties other than the incumbent are to be able to win the franchise and operate it successfully, there has to be a way for them to gain access to the tracks. Since forced purchasing is legally messy and undermines the incumbent's incentives to carry out long-term investments, vertical separation between track and franchise is appropriate not only because it makes the franchise smaller, but also because it makes the entire programme more feasible.⁹⁴⁴

This kind of split is familiar from other industries⁹⁴⁵, and was widely adopted in other countries as well.⁹⁴⁶ The same goes for the split between the freight and parcel divisions of British Rail, which were fully privatised, and the rest.⁹⁴⁷ However, the British government actually went several steps further to reduce the size of the franchises and make it as easy as possible to replace one company without affecting the industry at large. For starters, they also split off the ownership of the rolling stock into separate “rolling stock leasing companies” (ROSCOs), reducing the need for a winning bidder to contract for their purchase separately.⁹⁴⁸ Furthermore, they created not only 25 franchises, run by 25 “passenger train operating com-

⁹⁴⁴ Cf. p. 20 of the prospectus that accompanied the 1996 floatation of Railtrack Group PLC.

⁹⁴⁵ Cf. for example sections 3(3), 4(1), 6 and 66(1) of the Electricity Act 1989 (c. 29), which created transmission companies distinct from the electricity supply and production companies. Cf. Surrey (1996).

⁹⁴⁶ Cf. section 3.4.2, below. Note, however, that the 1987 privatisation of the Japanese railways maintained vertical integration between operation and track.

⁹⁴⁷ Cf. p. 22 of the prospectus that accompanied the 1996 floatation of Railtrack Group PLC.

⁹⁴⁸ Cf. Wolmar (2007), p. 304 and p. 22 of the prospectus that accompanied the 1996 floatation of Railtrack Group PLC. Note that for this scheme to work, it is not enough that there be more than one rolling stock leasing company. It is also necessary that they should have sufficient excess capacity and that it is possible to move rolling stock around relatively easily. In practice, none of these requirements were met. Cf. Wolmar (2007), p. 309.

panies” (TOCs),⁹⁴⁹ but also 7 “infrastructure maintenance companies” (IMCs or INFRACOs) and 6 “track renewal companies” (TRCs)⁹⁵⁰, most of which ended up in the building industry in one way or another.⁹⁵¹ All things considered, this added up to a pretty intimidating organisational chart for the industry⁹⁵², requiring many different contracts to be made between a great number of different parties, and to be monitored not only by the Office of the Rail Regulator, but also by a Franchise Director.⁹⁵³

This would be an absolutely perfect example of design at work, where every element of the governance structure fits perfectly with every other element, but for the fact that it was in fact a bit of a rush job. Mrs. Thatcher had always resisted privatising the railways in any manner as drastic as this, arguing that it was “a privatisation too far”.⁹⁵⁴ When John Major took over the Conservative Party at the end of 1990, there was finally a degree of consensus among the Tory front bench in favour of privatisation in some form. However, the party was wildly unpopular and an election victory looked unlikely. When the time came to write the 1992 manifesto, there was no detailed view of what a privatised railway industry should look like. That is why the manifesto does specify some elements of the ultimate arrangements – such as the franchising system and the privatisation of BR’s freight operations – but not others. It is silent, for example, on the issue of separation of track and train, or the creation of ROSCOs and INFRACOs, or the privatisation of Railtrack, the owner of the rails themselves.

When the Tories managed to snatch victory from the jaws of defeat in April 1992, they realised that they would have to hurry if they were to have any chance of completing the privatisation process within the space of a single parliament.⁹⁵⁵ After all, breaking up a company and organising a series of public stock offerings for the successor companies takes a long time, and the Tories did not want to take the risk that the project might go uncompleted,

⁹⁴⁹ Cf. p. 21 of the prospectus that accompanied the 1996 floatation of Railtrack Group PLC. and Yvrande-Billon & Ménard (2005), p. 685.

⁹⁵⁰ Cf. p. 20 and 22 of the prospectus that accompanied the 1996 floatation of Railtrack Group PLC. For the reader’s amusement, it might be interesting to remark that the difference between maintenance and renewal was defined based on the length of the track that needed to be worked on. Anything longer than 500 feet was renewal. Cf. Wolmar (2007), p. 337 (endnote 7).

⁹⁵¹ Cf. Murray (2001), p. 35-37.

⁹⁵² There is one available on p. 21 of the prospectus that accompanied the 1996 floatation of Railtrack Group PLC.

⁹⁵³ Cf. section 1 of the Railways Act 1993 (c. 43), as well as schedule I to the Act. The general duties of the Regulator are outlined in section 4, while the main duty of the Franchise Director – to negotiate with the passenger transport operating companies – is specified in section 23 of the Act. Cf. also Yvrande-Billon & Ménard (2005), p. 685-686.

⁹⁵⁴ Cf. Coleman (2005) and Lodge (2003), p. 165-166.

⁹⁵⁵ Under section 7 of the Parliament Act 1911 (c. 13), the maximum duration of a parliament is 5 years.

which would leave it vulnerable to being repealed by a future Labour government. In July 1992, only three months after the elections, the government published a White Paper on the future of the industry.⁹⁵⁶ In addition to the proposals already made in the manifesto⁹⁵⁷, the government now argued for a vertical split⁹⁵⁸ and for the eventual privatisation of Railtrack.⁹⁵⁹

Unlike a campaign manifesto, a government White Paper can offer at least some insight into the arguments that underlie the government's proposals. In the 1992 White Paper on rail the government offered six distinct benefits of its proposed reforms⁹⁶⁰:

1. The need to make the company profitable would force everyone involved to be more customer friendly. Presumably the idea was that, failing this, those customers would use other modes of transportation instead.
2. Competition for the market would force companies to innovate and strive for maximum customer satisfaction. It is difficult to see why privatisation and competition are listed separately, although perhaps the idea was to allude to a type of benchmarking behaviour, where the introduction of a service by one franchisor leads to pressure on the others to improve their quality as well.
3. More management to do what they are good at without interference from the government would, according to the Major government, lead to improvements in customer satisfaction as well. This is difficult to square with the insights of the previous chapter, which tends to argue that using contracts instead of vertical integration means that more things have to be fixed in advance. Maybe the interference referred to is not really government interference, in the sense of government action aimed at maximising its objective function, but rather government interference in the sense of action by individuals within the government in excess of, or contrary to, what would be necessary to maximise social welfare. Such intrusion is not without precedent in the UK, where Members of Parliament are elected separately in their own constituencies, making them vulnerable to rent seeking by their constituents, as well as by interest groups generally.

⁹⁵⁶ New Opportunities for the Railways: The Privatisation of British Rail.

⁹⁵⁷ The privatisation of BR's freight and parcel services is defended in par. 45-55 of the White Paper.

⁹⁵⁸ Par. 16 and 22-24 of the White Paper.

⁹⁵⁹ Par. 18 of the White Paper.

⁹⁶⁰ Par. 19 of the White Paper.

4. While the Thatcher government had already started managing the railway industry by objectives⁹⁶¹, the Major government now argued that privatisation would allow these goals to be made better enforceable. This is true, of course, only to the extent that the government's commitment to punish offenders is, in fact, credible.⁹⁶²
5. As a fifth point, the government argues for the importance of "regional or local identities" to boost the work ethic of the employees. This point, which also had its precedents in earlier government thinking about the railway industry, is difficult to incorporate in economic modelling.
6. Finally, the government argued that British Rail was subject to certain diseconomies of scale, meaning that it could be run more efficiently if it were smaller. A reduction in size would allow management to be more familiar with local circumstances, making it easier for them to identify "opportunities to cut out waste and otherwise reduce costs", as well as – presumably – to improve customer satisfaction by offering them a more tailor-made service.

On the other side of the ledger, of course, stood the one-time costs of reorganising the industry in this way, as well as the recurring transaction costs this market structure would involve. Both from the simple fact that they are not mentioned in the White Paper, and from the historic accounts of the period, we can conclude that there was a distinct optimism about the government's ability to keep these transaction costs sufficiently low, i.e. low enough that they need not be a real part of the discussion, i.e. at least an order of magnitude smaller than the gain in production efficiency.

In the end, the privatisation programme was completed with some time to spare. Following the enactment of the new Railways Act in November 1993, Railtrack was created on April 1, 1994.⁹⁶³ On the same day, the rolling stock was divided over three ROSCOs and the contracts with the IMCs and TRCs took effect.⁹⁶⁴ Following the breakup, the constituent parts were sold off piece by piece, culminating in the floatation of Railtrack in May 1996. The last significant privatisation, that of the Scottish TOC ScotRail was just weeks before the general election of May 1, 1997, which brought Tony Blair's Labour government to power. The new

⁹⁶¹ Cf. the 1983 and 1987 Conservative Party General Election Manifestos. For a pioneering discussion of Management by Objectives, cf. Drucker (1954).

⁹⁶² Cf. Schelling (1956) and Williamson (1983).

⁹⁶³ Cf. p. 20 of the prospectus that accompanied the 1996 floatation of Railtrack Group PLC.

⁹⁶⁴ Cf. p. 22 of the prospectus that accompanied the 1996 floatation of Railtrack Group PLC.

government did indeed find the new industry governance structure sufficiently entrenched that it did not consider it feasible to immediately undo what the Major government had done, its “commitment to a railway which is owned by and accountable to the public” notwithstanding.⁹⁶⁵

When the Blair government did act, starting in the year 2000⁹⁶⁶, the immediate occasion was a perceived reduction in safety, particularly because of a series of major railway accidents.⁹⁶⁷ Whether this perception was justified is a difficult matter⁹⁶⁸, but one that is mostly irrelevant for our analysis. What matters instead is the perceptions of the government of the time, and they saw not only an increase in railway accidents, but also two specific explanations for this problem. Firstly, the argument was that Railtrack’s private status gave it an incentive to underinvest in a way that the government could not prevent.⁹⁶⁹ Secondly, they argued that splintered nature of the industry meant that the contracting problems between any two companies added up to an enormous contracting problem. After all, it is problematic enough if the government is having difficulty writing a complete contract with Railtrack, in order to make sure it carries out its maintenance duties in the manner desired. But if those maintenance duties are then contracted between Railtrack and the IMCs, the problem is doubled. This is all the more true because, in this constellation, both the government and Railtrack will have difficulty maintaining the human capital necessary to negotiate and enforce these contracts. If neither the government nor Railtrack knows exactly how much maintenance is necessary when and where, who will make sure that the IMCs do what they are supposed to?⁹⁷⁰

The inevitable response to this problem was the partial renationalisation of Railtrack in 2002. By that time, the shares of the company had fallen significantly from their £3.90 opening price and their £17.68 maximum due to sprawling maintenance costs, as well as the penalties incurred in the aftermath of the Hatfield crash, when Railtrack initially ordered a

⁹⁶⁵ The quote is from p. 13 of the prospectus that accompanied the 1996 floatation of Railtrack Group PLC. Cf. Murray (2001), p. 16. In 1997 Labour General Election Manifesto said, however: “Our task will be to improve the situation as we find it, not as we wish it to be.” Cf. Wolmar (2007), p. 306.

⁹⁶⁶ Part IV of the Transport Act 2000 (c. 38) created the Strategic Rail Authority to replace the Franchising Director and supervise the government’s overall strategy for the railways. Cf. Murray (2001), p. 159-163. This body was abolished in the Railways Act 2005 (c. 14), section 1.

⁹⁶⁷ Southall in 1997, Ladbroke Grove in 1999, Hatfield in 2000 and Potters Bar in 2002. Combined, these four accidents involved 49 deaths.

⁹⁶⁸ Cf. Evans (2007) and Wolmar (2007), p. 307.

⁹⁶⁹ Cf. Murray (2001), p. 69-78, Whitehouse (2003) and Wolmar (2007), p. 309-311.

⁹⁷⁰ Wolmar (2007), p. 307-309.

network-wide shutdown, followed by a resumption of traffic with many speed restrictions.⁹⁷¹ The newly created company Network Rail was able to buy the shares back at £2.62.⁹⁷² Following the enactment of the Railways Act 2005⁹⁷³ in April 2005, the government supervises and controls the industry through only two bodies: the Department of Transport and the Office of Rail Regulation. These two carry out their duties with much less reluctance than originally envisaged by the Major government, resulting in an overall greater amount of government intervention in the industry. This is reflected not so much in a change in the formal structure of the industry – Britain continues to have the most splintered, liberalised and privatised railway industry of any European country – but rather in the level of detail with which the government contracts with its private counterparts.⁹⁷⁴

3.4.2 *The EU Railway Packages*

As noted earlier, the European Community's involvement in the railway industry only started in 1991, with Directive 91/440.⁹⁷⁵ The authority to enact such a law had existed since the original Treaty of Rome of 1957⁹⁷⁶, but the EC had never applied it to the railways in any drastic way.⁹⁷⁷ However, the combination of two factors created the perfect storm for the Europeanisation and liberalisation of the railways.

On the one hand, there was the Common Market programme, launched by the Delors Commission in 1985⁹⁷⁸ and adopted by the Milan European Council that same year.⁹⁷⁹ This programme was part of a general drive towards a strengthening of the European Communities

⁹⁷¹ Cf. Murray (2001), p. 116-118 and Wolmar (2007), p. 308.

⁹⁷² Wolmar (2007), p. 304. This new company was not to be a publicly traded company, but rather a company limited by guarantee, forbidden to pay out dividends. Cf. section 3(3) of the Companies Act 2006 (c. 46), and section 1(2)(b) of the Companies Act 1985 (c. 6) and Whitehouse (2003).

⁹⁷³ Railways Act 2005 (c. 14).

⁹⁷⁴ Cf. Department for Transport (2010) and KPMG (2010).

⁹⁷⁵ Directive 91/440 on the development of the Community's Railways, OJ L 237, d.d. 24.08.1991, p. 25-28.

⁹⁷⁶ Art. 75 EEC, subsequently renumbered as art. 71 EC and now art. 91 TFEU. Art. 84 (old) EEC, the Common Transport Policy covered "rail, road and inland waterway[s]", with a possible extension to transport by sea and air. Cf. Debande & Monami (1998), p. 349-352.

⁹⁷⁷ The most important acts include Regulation 1893/91 of 20 June 1991, OJ L 169/1, Regulation 3359/90 of 20 November 1990, OJ 326/1, Council Recommendation 84/646 of 19 December 1984, OJ L 333/63, Council Decision 82/529 of 19 July 1982, OJ L 234/5, Directive 75/130 of 17 February 1975, OJ L 48/31 and Council Decision of 3 April 1962, OJ 23/720. According to Knill & Lehmkühl (1998), the topic of railway reform "had been on the agenda of the Commission since the mid-1970s".

⁹⁷⁸ Completing the Internal Market, Commission White Paper d.d. 14 June 1985, COM (85) 310.

⁹⁷⁹ Milan European Council, 28-29 June 1985 (Bull. EC 6-1985, point 1.2.5).

after the difficult 1970s. After some failed attempts⁹⁸⁰, this led to the adoption of the Single European Act in 1986 and the Maastricht Treaty in 1992.⁹⁸¹ The net result of these developments was a significant increase in the width and depth of the Community's competences. While the actual competences in the area of transport remained relatively unaffected, clearly these Treaty amendments left the European institutions more willing to use them.

On the other hand, there was the ever spreading fashion of liberalisation at the national level.⁹⁸² In Germany, for example, Chancellor Kohl had been pushing his own – more modest – version of Thatcher's liberalisation programme since he came to power in 1982.⁹⁸³ In the Netherlands, similar moves were made by the successive Lubbers cabinets.⁹⁸⁴ Even the – initially – staunchly socialist French president Mitterrand was forced to move to the right during the 14 years of his tenure in office.⁹⁸⁵ As Marx and Engels might have put it: “Ein Gespenst [ging] durch Europa”⁹⁸⁶, only this time, it was not the spectre of communism, but the spectre of neo-liberalism.⁹⁸⁷

The connection between the philosophy – one might even call it an ideology – of neoliberalism and the theories discussed in chapter 2 is actually quite intriguing. Margaret Thatcher was famously a fan of Friedrich Hayek's 1945 essay *The Road to Serfdom*⁹⁸⁸, which Hayek later used as the basis for his 1960 book *The Constitution of Liberty*.⁹⁸⁹ These works, as well as others published by Hayek's Mount Pelerin society⁹⁹⁰ and the writings of yet another Nobel laureate, Milton Friedman⁹⁹¹, formed the core of neoliberalism. Also in the 1960s, Buchanan & Tullock invented public choice theory⁹⁹², which fit nicely with Hayek and Friedman because it explained why the government might do a less than perfect job of curing

⁹⁸⁰ Cf. Kapteyn & VerLoren van Themaat (2008), p. 28-29.

⁹⁸¹ Cf. Kapteyn & VerLoren van Themaat (2008), p. 29-34, Nugent (2006), p. 49 and Werner & Wessel (2010), p. 34.

⁹⁸² Cf. Engartner (2008), chapter 2.

⁹⁸³ Cf. Engartner (2008) and Gall & Pohl (1999), p. 362-378.

⁹⁸⁴ Cf. De Korte (1990).

⁹⁸⁵ This was not only the consequence of having to work with the gaulliste prime minister Jacques Chirac from 1986-1988, but also because of the economic realities with which he had to work. Cf. Singer (1996).

⁹⁸⁶ Cf. The Manifesto of the Communist Party, published on 21 February 1848 by Karl Marx and Friedrich Engels.

⁹⁸⁷ Cf. the work of John Williamson, the “inventor” of the Washington consensus: Williamson (1990, 2002), and more generally Harvey (2005) and Steger & Roy (2010).

⁹⁸⁸ “.. the most powerful critique of socialist planning and the socialist state which I read at this time [the late 1940's], and to which I have returned so often since [is] F. A. Hayek's *The Road to Serfdom*.” Thatcher (1995), p. 50.

⁹⁸⁹ Hayek (1960).

⁹⁹⁰ Cf. Barry (1979) for a discussion.

⁹⁹¹ Friedman received the prize in 1976, Hayek in 1974.

⁹⁹² Buchanan & Tullock (1962).

certain undeniable market failures.⁹⁹³ Apart from the work of Coase himself, however, the popularisation of Coasian institutional economics, as well as its application to the government, did not start until the middle of the 1970s; too late to influence the development of neoliberalism, assuming this would have been possible in the first place⁹⁹⁴, but early enough to affect the manner in which the privatisations mandated by neoliberalism were carried out. The question, as before, is how and how much decision makers relied on Coasian arguments, and whether the ultimate result is likely to have been close to the theoretical optimum. Despite the fact that the pillars of neoliberalism all focused on the shortcomings of the state, giving the ideology itself a sometimes jarring anti-*étatist* bias, that does not mean that the outcomes of neoliberalist policies were too far on the contracting side of the spectrum. It is very well possible that reality had a sufficiently *étatist* bias that the net result was exactly right.⁹⁹⁵ Similarly, it is possible that the neoliberalist politicians were restrained in their enthusiasm by the presence of sufficiently strong alignment mechanisms.

As we have already seen, the UK approach to railway privatisation failed to take into account the contracting difficulties involved in having two or three government regulators, an independent railtrack company, 25 TOCs, three ROSCOs and countless other companies work together to provide rail transportation services. And as the “experiment” progressed the UK government drew exactly that lesson. In other words, they learned about the importance of transaction costs the hard way, through trial and error.

The other European Member States, and to some extent the European Commission, tended to be more cautious.⁹⁹⁶ Moreover, their approach showed the signs of the political institutional framework within which they were working; because they were making policy at the European level before doing anything significant at the national level, they started by making laws that would most clearly help to “complete the Internal Market”. Completing the Internal Market is less about liberalisation or privatisation than about removing the barriers between the Member States.⁹⁹⁷ In the short term, that meant getting the national railway in-

⁹⁹³ For a proper introduction to Public Choice theory, cf. section 4.1, below.

⁹⁹⁴ Arguably, the work of Ronald Coase had much too much nuance for the hardcore neoliberals like Thatcher and Reagan, meaning that they would not have taken it into account even if they had been exposed to it.

⁹⁹⁵ Cf. the discussion of the UK government’s economic policies in the 1970s, above.

⁹⁹⁶ Cf. Hayek’s famous postscript to *The Constitution of Liberty*: “Why I Am Not A Conservative”. In a nutshell, the idea is that conservatism “by its very nature (...) cannot offer an alternative to the direction in which we are moving”, while Hayek, of course, was advocating a radical change to the status quo. Cf. Hayek (1960), p. 395-411. Viewed this way, continental Christian-Democrats were more conservative than Thatcher.

⁹⁹⁷ Cf. par. 8-9 of the 1985 White Paper *Completing the Internal Market*, Commission White Paper d.d. 14 June 1985, COM (85) 310. Cf. also Pelkmans (2006), p. 6-7.

cumbents to work together, for example on high-speed international services like the Eurostar.⁹⁹⁸ The first step in that direction was the aforementioned Directive 91/440.⁹⁹⁹

Article 1 of the Directive explains it all:

The aim of this Directive is to facilitate the adoption of the Community railways to the needs of the Single Market and to increase their efficiency;

-by ensuring the management independence of railway undertakings;
-by separating the management of railway operation and infrastructure from the provision of railway transport services, separation of accounts being compulsory and organizational or institutional separation being optional,
-by improving the financial structure of undertakings,
-by ensuring access to the networks of Member states for international groupings of railway undertakings and for railway undertakings engaged in the international combined transport of goods.

Contrary to what one might think, the most noteworthy aim here is not the second one, where a full vertical split is merely “optional”, but the first one, requiring “management independence”.¹⁰⁰⁰ The second and fourth aims have a clear connection to the goal of completing the Common Market; the other two, however, seem to have been at least partly motivated by EC paternalism.¹⁰⁰¹ As in Britain, by the 1980s the railways of the continent were in shambles, because their management had gotten away from the governments, with one subsidy seemingly requiring a limitation of free competition somewhere else, plugging one hole with another until “the accounts of the railway company lost all contact with reality”.¹⁰⁰²

This problem, as well as the Commission’s proposed solution, is a matter of (New) Public Management.¹⁰⁰³ Public Management is a distinct area of study from the institutional economics that is the topic of this thesis, examining a distinct set of questions. Putting it simply, we can say that institutional economics is about finding the right institutions, and (public) management is about trying to make sure that those institutions – and particularly hierarchy – function the way the institutional economists imagined it. Because public management is a

⁹⁹⁸ The Channel Tunnel first opened in 1994, with the Eurostar connection between London and Paris/Brussels starting later that year.

⁹⁹⁹ Directive 91/440 on the development of the Community’s Railways, OJ L 237, d.d. 24.08.1991, p. 25-28.

¹⁰⁰⁰ Cf. art. 4 and 5 of the Directive. Cf. also section II of the Commission’s 1998 evaluation of the Directive, COM(1998) 202.

¹⁰⁰¹ This being – of course – before the principle of subsidiarity was enshrined in the EU treaties by the Treaty of Maastricht. Under the current Treaties, the EU institutions would never enact a law simply because they think they know better what is in the best interests of the Member States than the Member States themselves. Cf. Davies (2006, 2008).

¹⁰⁰² Nash (1991), p. 3.

¹⁰⁰³ Cf. Starling (2002).

distinct topic, the woes of the European railway industry in the 1960s and 70s are not a suitable subject for this thesis, which is why they were largely ignored so far, except to the extent that they were relevant to the discussion of the Thatcher reforms in the 1980s and 90s.

As we have already seen, one of the New Public Management reforms implemented by Thatcher in the 1980s was management by objectives. The railways were organised as an explicitly separate agency¹⁰⁰⁴, its management was professionalised and given the freedom to get the job done¹⁰⁰⁵ and the success of management was measured with reference to explicit performance measures that were agreed upon in advance.¹⁰⁰⁶ This is the system that was now being legislated for the rest of Europe's railways as well.

From an institutional economic perspective, it is difficult not to be sceptical at such reforms. Superficially, they seem to place the principal at a greater distance from the agent, without any compensating increase in the ability of the government to monitor the railways, or in their ability to create competition for the market as an efficiency-enhancing device.¹⁰⁰⁷ Analytically, the only way that this seems to make sense is as a form of self-restraint on the part of the government, a credible commitment not to interfere with the day-to-day running of the railways. This is a question we will return to later in this chapter.

Following the adoption of Directive 91/440, progress was slow. With the exception of the UK, the European Member States did not seem particularly inclined to experiment with their railways, beyond the New Public Management reforms that is. The Netherlands, for example, did not get a proper new Railway Act until 2003.¹⁰⁰⁸ Until then, they worked with a patched up version of the Act of 1875.¹⁰⁰⁹ Only in 1996, five years after the enactment of the Directive, did the Netherlands move ahead with significant – and voluntary – reforms, creating distinct legal persons to administer the track and the rail transport services.¹⁰¹⁰ Germany, on the other hand, had to give itself a brand new set of Railways Acts in order to merge the railway companies of West- and East-Germany, so they conveniently also transposed Directive

¹⁰⁰⁴ Cf. Hood (1991), Table 1: Doctrinal components of new public management, No. 4.

¹⁰⁰⁵ Cf. Hood (1991), Table 1: Doctrinal components of new public management, No. 1.

¹⁰⁰⁶ Cf. Hood (1991), Table 1: Doctrinal components of new public management, No. 2 and 3.

¹⁰⁰⁷ Remember, it is called New Public Management for a reason.

¹⁰⁰⁸ Spoorwegwet, Stb. 2003, 264.

¹⁰⁰⁹ Spoorwegwet, Stb. 1875, 67.

¹⁰¹⁰ Law of 13 November 1997, Stb. 1997, 578. Cf. The Commission's 1998 evaluation of Directive 91/440, COM(1998) 202, and Knill & Lehmkuhl (1998), par. 3.2 and Veenendaal (2004), ch. 24.

91/440¹⁰¹¹, although they too missed the deadline of 1 January 1993.¹⁰¹² What Germany did not do, though, is go very far beyond the minimum reform required by the Directive.¹⁰¹³

In France, a somewhat more advanced degree of separation between the track and the transport branches of the railways was created by the Law No. 97-135 of 13 February 1997.¹⁰¹⁴ Under this law, the track was placed in a distinct legal person (“*épïc*”¹⁰¹⁵), but both the *Réseau ferré de France* and the SNCF, the *Société Nationale des Chemins de Fer* continued to be under the direct control of the state bureaucracy.¹⁰¹⁶ In Belgium, finally, there was little need for any implementing legislation, since the Belgian railways had been reformed only several months before the enactment of the Directive, in a manner that was broadly consistent with the Directive, albeit initially without any kind of vertical split¹⁰¹⁷, which had to wait until 1997.¹⁰¹⁸

In the rest of the 1990s, reforms enacted in various Member States – whether as a result of Directive 91/440, as a result of other EC legislation¹⁰¹⁹, or without any kind of EC compulsion – gradually paved the way for the notion of treating the railways just like this thesis has treated it so far: like any other market, to be regulated or otherwise intervened in only to the extent that it fails.¹⁰²⁰ This change in perspective, in turn, did not only make it eas-

¹⁰¹¹ Eisenbahnneuordnungsgesetz, BGBl. I S. 2378, Gesetz über die Eisenbahnverkehrsverwaltung des Bundes BGBl. I S. 2394, Allgemeines Eisenbahngesetz, BGBl. I S. 2396 and Deutsche Bahn Gründungsgesetz, BGBl. I S. 2386, all of which are dated 27 December 1993.

¹⁰¹² Cf. art. 15 of Directive 91/440.

¹⁰¹³ Cf. Gall & Pohl (1999), p. 409-416, Nash (1999), p. 13 and par. 11 and 15 of the Commission’s 1998 evaluation of the Directive, COM(1998) 202.

¹⁰¹⁴ Loi No. 97-135 du 13 février 1997 portant création de l’établissement public “ Réseau ferré de France ” en vue du renouveau du transport ferroviaire, JO du 15/02/1997, p. 2592-2594. Before that, the French government already enacted the Décret No. 95-666 of 09/05/1995 portant transposition de la directive 91-440 du Conseil des Communautés européennes du 29/07/1991 sur le développement de chemins de fer communautaires et relatif à la gestion et l’utilisation de l’infrastructure du réseau ferré national, JO du 10/05/1995, p. 7732 in order to comply with the Directive.

¹⁰¹⁵ Établissement public de caractère industriel ou commercial. Cf. art. 34 of the French Constitution and Cour des Comptes (2008).

¹⁰¹⁶ Cf. Rohou (1998) and par. 11 and 14 of the Commission’s 1998 evaluation of the Directive, COM(1998) 202.

¹⁰¹⁷ Wet van 21 maart 1991 betreffende de hervorming van sommige economische overheidsbedrijven, Staatsblad 27/03/1991, p. 6155

¹⁰¹⁸ Koninklijk besluit of 05/02/1997 tot uitvoering van de richtlijn van de Raad van de Europese Gemeenschappen (91/440/EEG) van 29/07/1991 betreffende de ontwikkeling van de spoorwegen in de Gemeenschap, Staatsblad 06/03/1997, p. 4826. Cf. also par. 11 and 15 of the Commission’s 1998 evaluation of the Directive, COM(1998) 202.

¹⁰¹⁹ Most importantly there were Directive 95/18 on the Licensing of Railway Undertakings, OJ L 143, 27.6.1995, p. 70–74 and Directive 95/19 on the allocation of railway infrastructure capacity and the charging of infrastructure fees, OJ L 143, 27.6.1995, p. 75–78. Cf. Dutch Law of 11 June 1998, St. 1998, 374.

¹⁰²⁰ For an early example, cf. the rapport of the Dutch Wijffels Committee: Wijffels, In ’t Veld & De Soet (1992), who discuss the implications of Directive 91/440 in their chapter 1. Knill & Lehmkuhl (1998) describe this as the “legitimisation” of national reform as a result of the European legislation.

ier for European government to get implementing legislation enacted, it also gave the European Commission more political space to get ambitious. In July 1998, they published their proposals for what would become the first railway package. It consisted of three directives: one to amend Directive 91/440, one to amend the licensing directive from 1995¹⁰²¹ and a new directive on the allocation of railway infrastructure capacity.

Starting with the first, it is safe to say that the Commission's proposals in this regard, which were ultimately adopted as Directive 2001/12¹⁰²², would be somewhat underwhelming to a reader unfamiliar with the nature of Brussels lawmaking. More than anything, the amendments seem to be intended to “clarify”¹⁰²³ what Directive 91/440 already meant to do. Yes, keeping separate accounts means having both a separate balance sheet and a separate profit and loss statement¹⁰²⁴ and yes, the “infrastructure manager” may be either a public body, or a public or private undertaking.¹⁰²⁵

The other three directives in the first package can be dealt with briefly. The licensing directive 2001/13¹⁰²⁶, like its predecessor, dealt with a fairly straightforward harmonisation problem: If railway undertakings are to operate in more than one Member State, there has to be some agreement on their licensing. Otherwise, a Member State might attempt to protect its incumbent railway company by setting very high standards for licensing. Much the same goes for the even-more-technical Directive 2001/14, “on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification”¹⁰²⁷ and Directive 2001/16 “on the interoperability of the trans-European conventional rail system”.¹⁰²⁸ None of these directives has any direct implications for the governance structures used by the Member States, except in the sense that they represented a “Europeanisation” of the problem.¹⁰²⁹

¹⁰²¹ Directive 95/18 on the Licensing of Railway Undertakings, OJ L 143, 27.6.1995, p. 70–74.

¹⁰²² Directive 2001/12 amending Council Directive 91/440 on the development of the Community's Railways, OJ L 75, 15.3.2001, p. 1-25.

¹⁰²³ Proposal for a Council Directive amending Directive 91/440 on the development of the Community's Railways, COM(1998) 480, p. 10, par. 6 and 9.

¹⁰²⁴ Art. 1(7) of Directive 2001/12, which reformulated art. 6 of Directive 91/440.

¹⁰²⁵ Art. 1(4)(a) of Directive 2001/12, which reformulated art. 3 of Directive 91/440.

¹⁰²⁶ Directive 2001/13 amending Council Directive 95/18 on the licensing of railway undertakings, OJ L 75, 15.3.2001, p. 26-28.

¹⁰²⁷ OJ L 75, 15.3.2001, p. 29-46.

¹⁰²⁸ OJ L 110, 20.4.2001, p. 1-27.

¹⁰²⁹ Cf. Barea, Dizy & Ruiz (2007), who discuss how the 2004 reform of the Spanish railway industry was inspired by – but not required by – the first railway package.

The impetus for a second package was created by a Commission White Paper published when the first package had only just made it to the official journal.¹⁰³⁰ As so often in EU politics, this meant that a set of proposals about railway legislation was rooted in a White Paper about transport policy, which, in turn, was a response to European Council conclusions talking about sustainable development.¹⁰³¹ Fortunately for European Citizens everywhere, this was not in fact as bad as it looks, since the European Council did explicitly invite the Commission to come up with further proposals:

29. A sustainable transport policy should tackle rising volumes of traffic and levels of congestion, noise and pollution and encourage the use of environment-friendly modes of transport as well as the full internalisation of social and environmental costs. Action is needed to bring about a significant decoupling of transport growth and GDP growth, in particular by a shift from road to rail, water and public passenger transport. To achieve this, the European Council:

- *invites* the European Parliament and the Council to adopt by 2003 revised guidelines for trans-European transport networks on the basis of a forthcoming Commission proposal, with a view to giving priority, where appropriate, to infrastructure investment for public transport and for railways, inland waterways, short sea shipping, intermodal operations and effective interconnection;

- *notes* that the Commission will propose a framework to ensure that by 2004 the price of using different modes of transport better reflects costs to society.

This last point is of course particularly intriguing for an economist. One of the problems facing the railways is that it has to pay for its own infrastructure, while its main competitors – road haulage and inland waterways – get their infrastructure free of charge. (With the exception of the occasional toll.)¹⁰³² To the extent that the European Council meant to say that the infrastructure charges imposed on the railway companies should better reflect the cost to society of using that infrastructure, such a proposal would be inconsistent with the goals stated, since it would only make the railways' problems worse. If, instead, they had some kind of kilometre-pricing system in mind for Europe's motorways, that was an ambitious suggestion indeed.¹⁰³³

¹⁰³⁰ European Transport Policy for 2010: Time to Decide, COM(2001) 370, d.d. 12.9.2001.

¹⁰³¹ Göteborg European Council, 15 and 16 June 2001, Presidency Conclusions par. 19-32 establishing a European Sustainable Development Strategy, and specifically par. 29 on sustainable transport.

¹⁰³² Cf. Nash (1991), p. 1.

¹⁰³³ Cf. Nash (2007), a paper we will return to in the next chapter.

In its response to the European Council's request, the Commission devoted part of its White Paper to the problem of "Revitalising the railways".¹⁰³⁴ Many important ideas it discussed again came from a "Completing the Common Market" angle, which is fine for freight transport, where the creation of a "trans-European rail freight network"¹⁰³⁵ almost inevitably implies a highly deregulated and privatised model at the national level, but for passenger transport the connection does not run quite so smoothly. In that area, the white paper was actually surprisingly unambitious, focusing mostly on technical issues and the old saw of collaboration between national railway companies and national regulators.¹⁰³⁶

This choice of emphasis was also clearly discernible in the Commission's proposals for what would ultimately become the second railway package. The amendments to Directive 91/440 dealt only with freight transport¹⁰³⁷, while the rest of the package dealt with such issues as safety, licensing and allocation of infrastructure capacity¹⁰³⁸, as well as interoperability.¹⁰³⁹ Finally, there was the creation of an Agency responsible for dealing with such matters.¹⁰⁴⁰ Notably absent from the European Railway Agency's given objectives is anything as politically sensitive as studying the relative merits of different possible governance structure for the industry.¹⁰⁴¹ In short, we can say that the second railway package did not break any new ground on either the reform of railway governance directly, or on the transfer of authority in this area to the European level. Even within the ERA's rather technical remit, it was limited to "address[ing] recommendations to the Commission" and "issu[ing] opinions to the Commission (...) and to the authorities (...) in the Member States"¹⁰⁴², meaning that even the creation of this Agency meant no further Europeanisation of railway policy.

¹⁰³⁴ European Transport Policy for 2010: Time to Decide, COM(2001) 370, d.d. 12.9.2001, p. 27-36.

¹⁰³⁵ Cf. p. 28 of the White Paper, as well as Directives 91/440 and 2001/12, *op cit*, as well as Directive 2004/51, OJ L 164, 30.4.2004, p. 164-172.

¹⁰³⁶ Cf. p. 33 and 35 of the White Paper. As for "technical issues", the White Paper provides a particularly eye-opening map on p. 31, which shows the different voltage used in different EU Member States, a map that looks more like an illustration of the Four Colour Theorem than like anything devised by sensible human problem solving.

¹⁰³⁷ Directive 2004/51, OJ L 164, 30.4.2004, p. 164-172. Cf. also the Commission's proposal for this Directive, COM (2002) 25, d.d. 23.1.2002.

¹⁰³⁸ Directive 2004/49, OJ L 164, 30.4.2004, p. 44-113.

¹⁰³⁹ Directive 2004/50, OJ L 220, 21.6.2004, p. 40-57.

¹⁰⁴⁰ Regulation 881/2004, OJ L 220, 21.6.2004, p. 3-15. For future reference: the European Railway Agency is located in Valenciennes in Northern France.

¹⁰⁴¹ Cf. art. 1 of Regulation 881/2004, which states the objectives of the ERA.

¹⁰⁴² Art. 2 of Regulation 881/2004.

The third railway package, enacted in 2007, did contain some advances in both respects. For example, it contained a Regulation on rail passengers' rights and obligations¹⁰⁴³, which secured for European citizens such rights as the right to bring a bicycle on the train¹⁰⁴⁴ and the right to have the Convention Concerning International Carriage by Rail of 1980 apply to them.¹⁰⁴⁵ Further Europeanisation was achieved by Directive 2007/59, which harmonised the requirements for train drivers' licenses.¹⁰⁴⁶

The Crown Jewel of the third railway package was Regulation 1370/2007.¹⁰⁴⁷ This Regulation, which had been in the making since 2000¹⁰⁴⁸, i.e. since before the second railway package was even proposed, repealed two Regulations from the early days of the Community.¹⁰⁴⁹ Its purpose was to navigate the difficult waters of the overlap between Union and Member State transport law, EU State Aid law and the law on Services of General Economic Interest (hereafter: SGEIs).¹⁰⁵⁰

The Commission's argument in support of its proposal ran roughly as follows: Assuming that passenger rail transport can never be made profitable as long as the public service obligation is maintained, and assuming that the policy reasons for favouring the railways that were outlined in the 2001 White Paper continue to hold, there is no escaping some state subsidy for the railways. That means that the regulatory problem facing the state is an order of magnitude more difficult than in the pre-War era, when the problem was simply one of re-

¹⁰⁴³ Regulation 1371/2007, OJ L 315, 3.12.2007, p. 14-41.

¹⁰⁴⁴ Art. 5 of Regulation 1371/2007.

¹⁰⁴⁵ Art. 4, 8, 11, 15 and 28 and Annex I of Regulation 1371/2007. The Convention already applied in every Member State except Malta and Cyprus, but at the time of writing the EU has still not ratified the Convention, even though the Vilnius Protocol of 1999 made it possible for a non-state entity like the EU to join. Cf. COM (2009) 441 of 31.8.2009, the Commission's most recent proposal for ratification, which was approved by the European Parliament on November 24, 2009.

¹⁰⁴⁶ Directive 2007/59, OJ L 315, 3.12.2007, p. 51-78.

¹⁰⁴⁷ Regulation 1370/2007, OJ L 315, 3.12.2007, p. 1-13.

¹⁰⁴⁸ COM (2000) 7, containing "a proposal for a Regulation (...) on action by Member States concerning public service requirements and the award of public service contracts in passenger transport by rail, road and inland waterway", was submitted by the Commission on 26 July 2000, cf. OJ C C 365 E, 19.12.2000, p. 169-178, together with "a proposal for a Regulation (...) concerning the granting of aid for the coordination of transport by rail, road and inland waterway", COM (2000) 5, OJ C 365 E, 19.12.2000, p. 179-183. Of the latter proposal nothing further was heard after the European Parliament voted its amendments in first reading, cf. OJ C 21 E, 24.1.2002, p. 256-257 and EP document no. A5-0096/2001. The former proposal was amended by the Commission in 2002, COM (2002) 107, OJ C 151, 25.6.2002, p. 146-183, and when that was not enough to convince the Council, the proposal was completely overhauled in 2005, COM (2005) 319, d.d. 20.7.2005. This final version of the proposal was the basis of the Council's Common Position, adopted on 30 November 2006, cf. Council document no. 13736/07, and of the subsequent second reading agreement.

¹⁰⁴⁹ Regulation 1191/69, OJ L 156, 28.6.1969, p. 1-7, English special edition: Series I Chapter 1969(I) P. 0276, and Regulation 1107/70, OJ L 130, 15.6.1970, p. 1-3, English special edition: Series I Chapter 1970(II) P. 0360.

¹⁰⁵⁰ At the EU level, these three areas of law are found in art. 90-100, art. 107-109 and art. 14 and 106(2) TFEU, respectively. The special importance of SGEIs is also emphasised in art. 36 of the Charter of Fundamental Rights and in Protocol 26 annexed to the Lisbon Treaty.

straining a would-be monopolist. Today, the problem is more one of finding a way to keeping the railway companies upright without completely undermining their incentives to innovate or to provide a high quality service. And given the stated goal of opening up the market for passenger services by rail as much as possible, this problem has to be dealt with at the European level.

Of course, the Commission as an institution had always had a pretty clear idea of what it wanted to do with SGEIs such as the railways. Ever since the 1980s, the Commission had been dominated by neo-liberals to a much great extent than most Member States, and – as we have seen – in many areas its reforms steadily moved the status quo in a neoliberal direction. The reason for this is not that Mrs. Thatcher had so much influence in the Berlaymont, or that the Eurocrats were such big fans of Mrs. Thatcher’s government. On the contrary, as anyone with a passing familiarity with EU history will know, Mrs. Thatcher managed to thoroughly antagonise just about everybody in Brussels while still signing on to the Single European Act.¹⁰⁵¹ Instead, the reason is that the usual reasons for caution with such potentially unpopular reforms do not apply to Commission *fonctionnaires*; if French farmers use their tractors to block the *périphérique*, the “European Mandarins” can still drive to work.¹⁰⁵² The “democratic deficit” that burdens the European Union in all its work also makes it easier for the European Institutions to champion policies that would be political suicide at the national level. This goes even for the directly elected MEPs and for the representatives of the Member States sitting in the Council.¹⁰⁵³

The Explanatory Memorandum of the Commission’s 2005 proposal started by reviewing the history so far, including the original 2000 proposal, which was summarised as focusing on using contracts as the preferred governance structure governing the transaction between the state and the railways throughout the EU, as well as focusing on using limited duration contracts and competitive tendering.¹⁰⁵⁴ The purpose of this approach, it said, was to improve legal certainty by removing the friction between Common Market principles and railway industry practice, as well as to “improve the efficiency and quality of services”. The review did not explain how the latter purpose would have been helped by the 2000 proposals.

¹⁰⁵¹ Cf. Rasmussen (1997), Riddell (1989), Thatcher (1995), Vincent (1987), and Wall (2008).

¹⁰⁵² Cf. Eppink (2007).

¹⁰⁵³ Cf. Hix (2008), ch. 4 and 5.

¹⁰⁵⁴ COM (2005) 319, d.d. 20.7.2005, p. 3 and 7-8.

The Barroso Commission had at least two arguments to cover its obvious backpedaling on this issue: Firstly, there was the 2003 Altmark ruling, which concerned exactly the overlap between transport, state aid, and SGEIs, giving the ECJ an impossible to miss opportunity to lay down the law.¹⁰⁵⁵ Secondly, there was the Commission's 2004 White Paper on Services of General Interest¹⁰⁵⁶, which – being written against the backdrop of the Treaty's explicit provisions in this area – is quite generous as to the Member States' right to legislate the governance structures they deem most appropriate.¹⁰⁵⁷

In paragraph 95 of Altmark, the ECJ spelled out what a Member State had to do in order to avoid seeing its payment for a public service obligation classified as state aid:

- *first, the recipient undertaking is actually required to discharge public service obligations and those obligations have been clearly defined;*
- *second, the parameters on the basis of which the compensation is calculated have been established beforehand in an objective and transparent manner;*
- *third, the compensation does not exceed what is necessary to cover all or part of the costs incurred in discharging the public service obligations, taking into account the relevant receipts and a reasonable profit for discharging those obligations;*
- *fourth, where the undertaking which is to discharge public service obligations is not chosen in a public procurement procedure, the level of compensation needed has been determined on the basis of an analysis of the costs which a typical undertaking, well run and adequately provided with means of transport so as to be able to meet the necessary public service requirements, would have incurred in discharging those obligations, taking into account the relevant receipts and a reasonable profit for discharging the obligations.*

While this had the benefit of finally clearing up whether, and if so when, the state aid rules applied to compensation for public service obligations, it should be obvious to the reader that such a detailed shopping list seriously limited the freedom of the EU legislature or the Member States to define the criteria differently.¹⁰⁵⁸

¹⁰⁵⁵ Judgment of the Court of 24 July 2003, Altmark Trans GmbH and Regierungspräsidium Magdeburg v Nahverkehrsgesellschaft Altmark GmbH, and Oberbundesanwalt beim Bundesverwaltungsgericht, Case C-280/00, ECR [2003] Page I-07747. Cf. Merola & Medina (2003) and Louis & Vallery (2004).

¹⁰⁵⁶ Cf. also the earlier Green Paper, COM (2003) 270 of 21 May 2003. According to Annex I of the White Paper on Services of General Interest, COM (2004) 374 of 12 May 2004, SGEIs are “services of an economic nature which the Member States or the Community subject to specific public service obligations by virtue of a general interest criterion.” By analogy, the European institutions use the term Services of General Interest to include “non-market services” as well.

¹⁰⁵⁷ Cf. Ross (2000). Note also that Protocol No. 26 of the current Treaties, regarding Services of General Interest, could not be more clear in this regard.

¹⁰⁵⁸ However, cf. Renzulli (2008).

As for the White Paper on Services of General Interest, it emphasised subsidiarity¹⁰⁵⁹ and collaboration between the Union and the Member States against the backdrop of the principles stated by the ECJ in *Altmark* the year before.¹⁰⁶⁰ As such, it formed somewhat of an intermediary step between the vague and generally aspirational statements in the Treaty and the specific and binding legislation of the third railway package. Compared to the 2000 proposal for Regulation 1370/2007¹⁰⁶¹, the 2005 proposal reflected a more genuine effort to strike an appropriate balance with regard to Subsidiarity.¹⁰⁶² In this new proposal, the *Altmark* criteria appeared – occasionally verbatim¹⁰⁶³ – in the operative part of the text. The resulting Regulation appeared in the Official Journal in December 2007 and mandated that henceforth the European passenger transport by rail would be governed through public service contracts, meaning:

one or more legally binding acts confirming the agreement between a competent authority and a public service operator [meaning any public or private undertaking or group of such undertakings which operates public passenger transport services or any public body which provides public passenger transport services] to entrust to that public service operator the management and operation of public passenger transport services [meaning passenger transport services of general economic interest provided to the public on a non-discriminatory and continuous basis] subject to public service obligations [meaning requirements defined or determined by a competent authority in order to ensure public passenger transport services in the general interest that an operator, if it were considering its own commercial interests, would not assume or would not assume to the same extent or under the same conditions without reward];

[D]epending on the law of the Member State, the contract may also consist of a decision adopted by the competent authority:

- taking the form of an individual legislative or regulatory act, or*
- containing conditions under which the competent authority itself provides the services or entrusts the provision of such services to an internal operator¹⁰⁶⁴*

The system thereby created is essentially a concession system, albeit one with both public and private concessionaries. In addition to the *Altmark* criteria, the Regulation gives detailed rules

¹⁰⁵⁹ Cf. par. 3.1 on p. 7.

¹⁰⁶⁰ Cf. par. 3.2 on p. 7-8.

¹⁰⁶¹ COM (2000) 7, OJ C 365 E, 19.12.2000, p. 169-178.

¹⁰⁶² Cf. COM (2005) 319, p. 11-12.

¹⁰⁶³ Cf. art. 4(1)(b)(i) of Regulation 1370/2007 and the second *Altmark* criterion: “the parameters on the basis of which the compensation (...) is to be calculated”).

¹⁰⁶⁴ Art. 2 (i) of Regulation 1370/2007, supplemented with some of the other definitions given in that article.

for the length of the concession – usually no longer than 15 years¹⁰⁶⁵ – and the manner in which these contracts are to be awarded.¹⁰⁶⁶

Looking back at the last 20 years of EU law in the area of railway transport, it is clear that this was not a process of policy learning, of a gradual tendency towards better alignment in any real sense. The various actors knew where they wanted to go, they just did not agree, nor did they know how to make such reforms politically palatable. Legitimation, not learning.¹⁰⁶⁷

3.4.3 Conclusion

Having considered these two radically different cases, representing as they do two radical extremes on the spectrum of radicality, on the one hand the British reforms which were implemented with a kind of radicality that is only possible in a winner-takes-all political system such as that of the UK, and on the other hand the extreme consensus model of the EU in all its slow-moving splendour, it is now time to return to the list of section 2.7.

Transactions: One could say that the very purpose of the modern reforms has been to bring the *quid pro quo* nature of the transactions that take place in the railway industry between the industry and the government back to the forefront. That is what Regulation 1370/2007 did, for example: it spelled out in significant detail what exactly the *quid* and the *quo* of the transaction were, and forced the Member States to be clearer still before entering into any public service contract.¹⁰⁶⁸ This is distinct from the fact that it also set limits on what kinds of *quids* and *quos* are permissible. The authors of such legislation are quite clearly working from a perspective of the railway industry as a market like any other, where a service is provided for consideration. While that is obviously also the perspective of this thesis, it is hardly self-evident.¹⁰⁶⁹

¹⁰⁶⁵ Art. 4(2)-(4).

¹⁰⁶⁶ Art. 5.

¹⁰⁶⁷ Cf. Knill & Lehmkuhl (1998).

¹⁰⁶⁸ Cf. art. 4.

¹⁰⁶⁹ The opposite perspective would involve viewing the relationship between the state and the railways as inherently bureaucratic, instead of having bureaucracy as only one possible governance structure governing an exchange. Such approaches tend to draw on Weber's theory of bureaucracy, cf. Weber (1920), p. 1046-1092, alt-

Efficiency: At the same time, the modern reforms assumed a more narrow notion of efficiency than before. The link between railway policy and the government's economic policy generally, which was first advanced by Bismarck in the 1870s and later "perfected" by the post-war Attlee government, has been replaced with the assumption that the greater good of society requires that the railways do what they do – moving goods and people from A to B – as efficiently as possible. This rules out any number of other goals that the government might pursue in its railway policy, up to and including the notionally inefficient but not illegitimate goal of making sure that "the whole of the national resources, in land, material and labour [is] fully employed".¹⁰⁷⁰

Costly Information: For a reform that is meant to replace bureaucracy with market, these new markets sure do cause a lot of paperwork. Consider the provisions of Regulation 1370/2007: In art. 3(3), the choice of a Member State to establish certain special maximum tariffs requires a notification to the Commission. The requirements given in art. 4 for the public service contracts mean that such a contract will always be a voluminous affair, among those one of the exceptions to the ordinary contract length requires more reporting to the Commission.¹⁰⁷¹ Art. 4(7) enters a new domain: requirements for the tendering documents, specifically the requirement that they "shall indicate, in a transparent manner, whether, and if so to what extent, subcontracting may be considered." This motif continues with the reference, in art. 5(1), to the public procurement directives 2004/17 and 2004/18, which, incidentally are of a combined length of 240 pages in the Official Journal.¹⁰⁷² Then there is a right of the Commission to ask for any and all information about the Member State's compensation for public service obligations¹⁰⁷³, the annual report that every competent authority has to write¹⁰⁷⁴, some further obligations to announce upcoming tenders and other awards in the Official Journal of the EU¹⁰⁷⁵, as well as various rights for private parties involved to ask for more information¹⁰⁷⁶ and to

though Weber himself was economist enough to be interested in markets as well, cf. particularly Weber (1920), p. 596-601. Note that bureaucracy revolves around power, rather than mutual exchange.

¹⁰⁷⁰ 1945 Labour Party General Election Manifesto.

¹⁰⁷¹ Art. 4(4), final sentence.

¹⁰⁷² Directive 2004/17, L 134, 30.4.2004, p. 1-113 and Directive 2004/18, L 134, 30.4.2004, p. 114-240. In fairness, 155 pages of that total consists of annexes.

¹⁰⁷³ Art. 6(2).

¹⁰⁷⁴ Art. 7(1).

¹⁰⁷⁵ Art. 7(2)-(3).

¹⁰⁷⁶ Art. 7(4)

appeal.¹⁰⁷⁷ Finally, no piece of EU legislation would be complete without an obligation for the Commission to evaluate the implementation of the Regulation, as well as its effectiveness.¹⁰⁷⁸

Of course, cynicism aside, this is exactly what we would expect from an NIE point of view: increasing your reliance on contracts instead of hierarchy means that any number of rules and understandings that previously could remain unstated or stated vaguely all of a sudden have to be written down, and written down precisely and in legally binding language, thus making them suitable of being the subject of (protracted) litigation. Moreover, compliance with all these formal rules also has to be checked, confirmed and verified, again in a transparent and suitably legalistic way. Of course, as pointed out in chapter 2.5, above, it is true of hierarchies and markets alike that principals have to check up on their agents, but it is nevertheless true that a shift towards market means more paperwork. It is this paradox that we have seen in the current section: an effort to do away with bureaucracy leads to more bureaucracy.¹⁰⁷⁹

Moral Hazard: Because rational decision makers weigh the marginal costs and benefits of paperwork just like they weigh the marginal costs and benefits of any other choice, this increase in bureaucracy does not only come at the expense of trees. It also means that the ability of the principal (i.e. the government) to supervise the agent (i.e. the railway companies) will decrease, meaning an equivalent increase in the scope for moral hazard. For every instance when a company gets caught, there must be countless others where it gets away with bending the rules in a way that goes against the spirit – if not the letter – of its agreement with the state in order to increase its profits.¹⁰⁸⁰ When they do get caught, this often leads to nothing more severe than one semi-public entity imposing a fine on another semi-public entity.¹⁰⁸¹

Asset Specificity and Fundamental Transformation: Any evaluation of the reforms discussed in this section has to begin by asking whether a genuine competition for the market has in fact

¹⁰⁷⁷ Art. 5(7).

¹⁰⁷⁸ Art. 11. This rapport has to be written “after the end of the transitional period”, i.e. somewhere in 2020. Cf. art. 8(2).

¹⁰⁷⁹ Cf. Vogel (1996), whose title says it all: “Freer Markets, More Rules”.

¹⁰⁸⁰ The Dutch rail infrastructure company ProRail was caught in an act of “penny cheap, pound foolish” by RTL news when it failed to switch on the heating systems for the points near The Hague central station in December 2009 even though snow had been predicted, resulting in a widespread railway shutdown the following day. Cf. RTL Nieuws, ‘ProRail veroorzaakte problemen spoor’, broadcast of January 14, available on <http://www.rtl.nl/actueel/rtlnieuws/home/>. (Last visited on May 25, 2010.)

¹⁰⁸¹ For example, on 26 June 2008, the Dutch railway regulator imposed fines totalling € 826.000 on ProRail for various infractions regarding the way it distributed railway capacity among the railway companies. Cf. NMa decision 200047/75 regarding “Capaciteitsverdeling 2007” and decision 200063/70 regarding “Netverklaring 2008”, both dated 26 June 2008.

been created. If this is not the case, then it is difficult to see where, in practical terms, the alleged efficiency gains are supposed to come from. At the same time, having transactions without any kind of competition is also problematic for Williamson's TCE: without a fundamental transformation, what is the difference between the period before the contract is signed and the period after? In such a situation, the moment of contracting would not alter the relationship between the parties in any significant way, nor would the moment of investment in a specific asset. Instead, there would be scope for opportunistic behaviour at all times, *ex ante* and *ex post*, meaning that the recommendation to consolidate would follow with or without asset specificity.

As so often when it comes to contestability¹⁰⁸², the problem is that it is difficult to observe these would-be entrants whose mere existence makes the chosen governance structure possible. Fortunately, in the railway industry it is possible to establish a lower bound for the number of companies competing for the market by looking at the total number of companies operating railway lines in a given country, as well as – where this information is available – the companies who competed in a tender but failed to win. With that in mind, we can appreciate the benefits of the UK having as many franchises as it does: Even though many of these franchisors have common shareholders¹⁰⁸³, there is still ample scope for competition for the market, and therefore for Fundamental Transformation.

Having made this observation, we should go no further. We should not confuse the question of whether a government has chosen to create a system of competition for the market with the question of whether it would be possible for it to do so. To do otherwise would be to confuse the result of applying TCE to the contemporary railway industry with the threshold question of whether it is capable of being so applied in the first place. It is clear from the last 15 years of experience in the UK, the Netherlands¹⁰⁸⁴, Germany¹⁰⁸⁵, France¹⁰⁸⁶ and Belgium that there is sufficient scope for competition for the market – should the relevant actors find such competition desirable – that a fundamental transformation does in fact take place when the government elects to contract with one railway company rather than another.

¹⁰⁸² The theory of contestable markets in neo-classical microeconomics was pioneered in Baumol, Panfzar & Willig (1982).

¹⁰⁸³ Murray (2001), p. 26 gives the example of the long-distance bus company National Express Group, which owns no fewer than nine TOCs.

¹⁰⁸⁴ Cf. OECD (2005), p. 213-229.

¹⁰⁸⁵ Cf. OECD (2005), p. 163-167.

¹⁰⁸⁶ Cf. OECD (2005), p. 155-161.

Uncontractible investments: The ultimate goal behind the most recent reforms, at least as far as passenger transport is concerned, is neatly summed up in the very first recital of the Passengers' Rights Regulation that was a part of the EU's third railway package in 2007:

*Whereas: In the framework of the common transport policy, it is important to safeguard users' rights for rail passengers and to improve the quality and effectiveness of rail passenger services in order to help increase the share of rail transport in relation to other modes of transport.*¹⁰⁸⁷

Speaking of the railway industry generally, the 2001 White Paper on Transport made a similar diagnosis:

*The priority must be to resolve the problems holding back its development: the lack of infrastructure suitable for modern transport and of interoperability between networks and systems, the constant search for innovative manufacturing technologies, the non-transparency of costs, and the patchy productivity and shaky reliability of the service, which is failing to meet customers' legitimate expectations.*¹⁰⁸⁸

In other words: if the railway industry is to compete with the other modes of transport, it has to deliver better quality, which is something that needs to be contracted for explicitly, since otherwise the subsidised railway companies would have no incentive to invest in improved quality. After all, as long as prices are fixed in the concession or agreed upon in negotiations with a government intent on keeping them low, the railway company will benefit from improvements in quality only through the resulting increase in the quantity of rail transport demanded, without the usual ability of a seller on a less-than-perfect market to maximise profits by translating this increase in demand into a higher price. It follows that the railway companies will always underinvest in quality improvements.

As noted in section 2.1, above, the mere fact that the rights and obligations of passengers are enacted in the form of a law that neither passengers nor railway companies directly consented to¹⁰⁸⁹ does not mean that the terms formulated in the Regulation have not become a part of the contract – in the institutional economics sense – between the passengers, the state and the railway companies. More notable for present purposes is the fact that all the rights listed in the Regulation are things that passengers can observe and, if necessary, enforce. This is of course a great solution for the government: many things that would otherwise be uncon-

¹⁰⁸⁷ Recital 1 of Regulation 1371/2007, OJ L 315, 3.12.2007, p. 14-41.

¹⁰⁸⁸ COM (2001) 370F, p. 28.

¹⁰⁸⁹ Obviously, representatives of both passengers and railway companies were exhaustively consulted during the consulting phase before the Commission formulated its proposal. Cf. the Commission's proposal for Regulation 1371/2007, COM(2004) 143F, p. 3-4.

tractible because they are impossible for the government to observe or verify¹⁰⁹⁰ are in fact quite easily observable for the third player in this game: the passenger. If the railway company offers poor quality by offering their passengers poor travel information in violation of art. 8, by having a cumbersome system for ticket sales in violation of art. 9, by being insufficiently accommodating to the disabled in violation of art. 19-25, or in any other way that violates Regulation 1371/2007, the government will know even if they cannot observe the breach of contract directly, because the passengers will complain or even sue.¹⁰⁹¹ However, enlisting the help of passengers only works in a small percentage of cases. In many other circumstances, it is either difficult for the government to be sufficiently precise about what they want, or it is difficult for the government to be certain about the additional subsidy the railway company deserves in return.

As we have seen, sometimes even investments in the basic infrastructure can be problematic. If the railway industry is completely privatised, how will the government know whether the infrastructure company is investing enough in maintenance? When an accident happens, it is already too late, not to mention that even then it can be difficult to agree upon the culprit.

On the whole, it is clear that the uncontractibility problems encountered by governments these days are exactly of the kind envisaged by GHM.

Alignment Mechanisms: The alignment mechanisms are stronger than they ever have been. In the UK, we have seen TOCs end up in great financial difficulties when the arrangements they had made with the government turned out to be flawed. One such a bankruptcy would simply be an evolutionary punishment for stupidity in the private sector, but to the extent that such problems are more common, they can act as an important source of pressure on governance structures: if the model chosen for the British railway industry works, the various companies involved are not supposed to go bankrupt except through their own error. At the moment, that does seem to be the case.

¹⁰⁹⁰ Cf. section 2.1.5, above.

¹⁰⁹¹ Art. 30 of Regulation 1371/2007 requires each Member State to “designate a body or bodies responsible for the enforcement of this Regulation”, which “shall take the measures necessary to ensure that the rights of passengers are respected.” These measures presumably include the penalties required by art. 32. Only a few private causes of action are created by the Regulation, being those familiar from other areas of transport law, i.e. for missing luggage, for serious delays, etc. This typically European preference for administrative enforcement instead of enforcement through torts is familiar from consumer law generally, as well as from competition law and many other areas of the law. Cf. for example Cafaggi & Micklitz (2007, 2009).

A more important alignment mechanism in contemporary transport law and practice is competition, and particularly competition between modes of transport. Whenever decision makers are talking about a “growing imbalance between modes of transport”¹⁰⁹², what they are saying is that the railways are losing the competition with the other modes. Of course, this does not necessarily mean that there is something wrong with the industry’s governance structures, but it does mean that everybody has every incentive to continue to look for improvements that can be made in this respect.

Finally, the scope for design is greater than at any time since the 1850s. The body of theoretical and empirical knowledge continues to grow, as lessons from the past are learned and further research is carried out.¹⁰⁹³ At the same time, the governments under consideration in this thesis show themselves more able than at any time since the early days of the railways to carry out drastic reforms in the industry. Perhaps ironically, this is the result of the vastly decreased importance of the sector relative to the economy as a whole. After all, it is difficult for a government to drastically alter the governance structures that tie it to the railways when that industry is “too big to fail”. The cost would be too high, especially if the reform should turn out to be a mistake. In John Major’s Britain, on the other hand, a total collapse of the railways would have been unfortunate, but hardly catastrophic. When the railways were almost entirely shut down after the Hatfield disaster, the country continued to function. Exactly because the railways have largely lost the competition with the other modes of transport, it is easier to reform them.

On the whole, it follows that if ever there was a time when the alignment between the transactions occurring in the railway industry and the governance structures that govern them were perfect, it would have to be now.

¹⁰⁹² 2001 European Commission White Paper on Transport, COM (2001) 370F, p. 22.

¹⁰⁹³ Cf. for example Barea, Dizy & Ruiz (2007), Friebel, Ivaldi & Vibes (2010), Laperrouza (2008), Laperrouza & Finger (2009), Lodge (2002, 2003), Van Ooststroom (2008), Weidmann (2008) and Yvrande-Billon (1999, 2000, 2006).

3.5 Conclusion

Looking back at the entirety of 185 years of railway history, it is clear that it will always be difficult to discern some kind of common thread. This is all the more true if that common thread is to be economic, since economic science by its nature does not do well with the vagaries of reality.¹⁰⁹⁴ While the true verdict for NIE will have to wait until we have considered the alternatives, it is appropriate to conclude this chapter with some tentative conclusions.

Even in the absence of a benchmark, we can formulate certain properties that it would be desirable for the application of the theory to display. Foremost among these is a high degree of responsiveness to changes in the exogenous variables. We would like for the theory to produce a different recommendation for different moments in history, for different countries, for passengers or goods transport, or even for different routes. After all, given how much the actual governance structures used have differed in this respect, a theory that always produces the same recommendation would be about as useful as a clock with no battery; it is even useless at the two moments each day when it does show the correct time.

Of course, the precise recommendations of the various versions of NIE is a problem that we have so far carefully avoided. Instead, at various points we have observed that the tendency was exactly opposite to what a generalised NIE model might suggest: in the initial period of high transaction costs, actors favoured the market. When the transaction costs became lower, they tended towards consolidation and nationalisation, a process that has only recently been reversed, albeit seemingly without any change in either transaction costs or production costs to explain it. This is an analysis which ignores the question of precise recommendation but simply considers how the changes in NIE exogenous variables would affect its endogenous variable, i.e. its recommendation of the optimal governance structure. Whatever the optimum might have been before, a reduction in transaction costs cannot cause the optimum to shift in favour of hierarchy. So if the observed transaction costs and chosen governance structures do simultaneously change in this way, something is wrong somewhere. Either the original governance structure was wrong, or the change in governance structure is wrong, or both. Regardless, the model does not come out looking good.

¹⁰⁹⁴ Such is the price it pays for seeking to say a little about a lot instead of saying a lot about a little. Cf. section 1.2, above.

Let us now look at each of the elements identified in section 2.7 in turn, elements which we argued were necessary and sufficient to make the model workable.

- Throughout the history of the railways, transactions properly so called were happening. In fact, at all times in all five countries those transactions occurred between legally distinct entities, using legally enforceable governance structures. Even when the state owned the railways, they were never fully integrated in the state apparatus, the way the roads are, for example. Instead, the old railway companies were either continued under new ownership, or merged into a single state-owned railway company.
- With the second element, we run into difficulty. Especially in the first and second periods, when the railway network was first introduced, it is not always fair to say that optimising behaviour was going on. Instead, issues of personal distaste, class warfare and national prestige also played a prominent role. In addition, during the period of consolidation the railway companies often managed to turn their personal interests into state policy through highly effective rent seeking, a topic we will return to in the next chapter. Only since the advent of modern democracy do the states under consideration genuinely appear to endeavour to maximise the efficiency of their railway transactions.
- Very little needs to be said about the costs of searching, negotiating and enforcing. While information became steadily less costly during the last 185 years, it is still hardly free of charge, as exemplified by the fact that railway (infrastructure) companies continue to blame others for delays.¹⁰⁹⁵ Another enduring problem are the cost overruns we discussed in section 3.1. Such uncertainty about the investment necessary in order to construct a given piece of railway infrastructure continues to plague the industry to this day.¹⁰⁹⁶ Such moral hazard problems really need to be dealt with if the industry is ever to be successfully deregulated.
- The next point, fundamental transformation, is particularly interesting, because a wealth of insight hides behind it. Looking at the narrative of this chapter, it is clear that such a transformation did not always occur. As we have seen in chapter 2, such a

¹⁰⁹⁵ During the presentation of its annual report over 2009, on May 28, 2010, Dutch rail infrastructure company ProRail said that “30% of delays in 2009 had been caused by people walking on the tracks, vandalism, [other] misconduct by passengers and theft on and around the track”.

¹⁰⁹⁶ The Dutch Betuwelijn, for example, was budgeted in 1992 at a cost of Hfl 5,2 bn/€ 2,4 bn, and ended up costing about € 4,8 bn. Cf. Centraal Planbureau (1995) and the findings of the Duivesteijn Committee, Kamerstukken II 2004-2005, 29 283, nrs. 5-7, particularly nr. 7 at p. 315, table 7.11.

transformation requires at least some competition for the market, since a transformation can hardly be said to occur if the parties are stuck with each other both before and after they decide to go into business together. One party is always the government, an entity that is not usually open to being replaced, so the question is whether the government genuinely had a choice of counterparties.

This is not the same as asking whether there was more than one railway company. Even when there was not, it is still possible that one could have been set up. As the most recent wave of reforms has demonstrated, a new railway company can always be created out of thin air. Nevertheless, it is probably fair to say that the costs of doing so would have been prohibitive for at least part of the history of the railways.¹⁰⁹⁷ Especially when access to large quantities of investment capital is difficult, or when all the relevant expertise is controlled by the incumbent, the government may be in a position where it has no choice but to work with the existing railway company if it wants to open a new line. This was the case, for example, in the France of the Second Republic, and in much of the rest of Europe during the end of the period of consolidation and the beginning of the period of nationalisation. It took until well into the 1950s and 60s before the improved liquidity of capital markets and the rise of other kinds of transport companies, such as bus and road haulage companies¹⁰⁹⁸, brought down the costs of creating a new railway company to a level where competition for the market was once again feasible.

- Investments in specific assets were present in their billions. Of course, the more pertinent question is whether any of those investments were made by the government in its capacity as counterparty to the railways. Obviously there were investments in railway construction, but whether there were also significant specific investments on the other side of the transaction remains difficult to say. A pragmatic analysis of the evidence would suggest that there are many occasions where an investment in political capital made the government vulnerable to opportunistic behaviour by the railway companies. However, reasonable minds can disagree whether that is the kind of investment that can sensibly be used in the Williamson model. Perhaps it would be better to file political capital under the rubric of Public Choice Theory.¹⁰⁹⁹

¹⁰⁹⁷ Cf. section 4.3, below, for a discussion of switching costs.

¹⁰⁹⁸ Cf. section 3.4.1, above. The British rail franchises almost always ended up in the hands of the major bus companies, because they were the only other parties – other than British Rail itself – who had the expertise to operate such a company.

¹⁰⁹⁹ Cf. section 4.1 and 4.2, below.

- As we have seen, the problem of uncontractible investments increased steadily as the 19th century progressed. At first, the government needed relatively little from the railway companies; the concession simply specified how and where the railroad was to be built, and what was to be done about the property rights of the landowners in the vicinity. None of that is particularly uncontractible, even in the mid-19th century, when communication was often difficult in the absence of a working railway system. Subsequently, the governments of the various countries under consideration became more concerned with the social implications of the railways, and started imposing various kinds of public service obligations, compliance with which was not always easy to achieve. Only during the second half of the 20th century did such investments in goodwill and other intangible assets become less costly to monitor, and therefore less uncontractible.

This variable is probably the most useful single variable we have encountered so far, when it comes to explaining the historical pattern of nationalisation and privatisation.

- Finally, there are the alignment mechanisms. The problem is that, put simply, it appears as if their combined strength is insufficient to push an unwilling government in a direction were it does not want to go, except over several decades. The present chapter is full of examples of government persisting in their chosen systems, despite its glaringly obvious shortcomings, most notably the French “muddle through” system of the second half of the 19th, where the railway industry was losing money by the billion, but the government remained unfazed until the political climate had caught up. At the same time, there were also examples of the government changing the system while it was actually running fairly well, including the change from Thatcher’s BR to Major’s fully unbundled and privatised version, and earlier Bismarck’s takeover of the railways of the various German states.

Evolution hardly ever played a major role, because states, governments and railway companies can afford quite a bit of inefficiency before they are forced out of business. Competition is more common, in the sense of the struggle of each country to improve its terms of trade. During some periods, in some places, we saw competition between railway companies connecting the same cities, either because of redundancy in the network induced by a lax government concession policy, as in the UK, or as a result of deliberate government effort to create such competition, as in the Netherlands. It is not obvious, though, that this kind of competition ever affected the institutional outcomes. The same can certainly not be said for competition between modes of transport. Such

competition clearly made life more difficult for the railways in the first half of the 20th century, forcing the sector and the government alike to rethink the industry's governance structure.

When it comes to rational design, finally, we have the problem that there is also such a thing as irrational design, and it is not always easy to tell the difference. It has long been a practice in good standing in politics to create a committee as a way to avoid actually doing something.¹¹⁰⁰ Also, just because a design is commissioned with the intention that it should actually make a difference, that does not mean that such a design is always free of ideological bias and other irrationalities. Often times, the resulting reforms conveniently aligned with the government's values, even if we cannot be sure that those values drove the reforms.

Given these tentative conclusions, it is now time to consider the benchmark models of institutional lock-in and public choice theory. Only when their views on the matter have been discussed is it really possible to say something conclusive about the merits of NIE.

¹¹⁰⁰ Cf. the discussion of the various committees set up in the second half of the 19th century in the Netherlands in section 3.2.2 above, or in Veenendaal (2004), ch. 8-15.

4. Railways: Alternative Approaches

Now that we have seen the story of the history of the railways, and have seen what NIE has to say about it, it is time to consider the benchmark against which NIE is to be judged. After all, as an economic model it is never going to be capable of perfectly explaining and predicting what is happening. Economic models, like human beings, are highly imperfect creatures. For this reason, one model's answer to a set of questions can never be judged unless there is an alternative model whose performance can be used as a benchmark.

In this case, we will examine not one but two alternative approaches, Public Choice Theory, particularly the part that deals with rent-seeking, and the theory of institutional lock-in or path dependence, the rationale being that between NIE and the two alternatives these models represent the bulk of what mainstream economics has to say about the matter.¹¹⁰¹ As we will see, these two alternatives differ wildly in their assumptions and their methodologies, not only between themselves, but also in comparison to NIE as we have discussed it in the previous two chapters. For this reason, it will not always be easy to make an objective comparison. However, once the work is finished it will allow us to say something about NIE beyond mere generalities.

In what follows, each of these models will be discussed, first in general terms, and then in its application to the case study of chapter 3. The theoretical sections 4.1 and 4.3 will of necessity be much simpler than the corresponding NIE chapter, i.e. chapter 2. At this stage, it is neither possible nor necessary to explore Public Choice Theory and Path Dependence with the level of detail of chapter 2. The present thesis is not set up to examine the relative merits of three theoretical perspectives, giving them each equal attention. Instead, the goal is more limited, to examine the value of NIE alone, and even that only in its application to certain types of problem. For this reason, our examination of the benchmarks need not be as detailed as our examination of that which is being benchmarked. The examination only needs to be good enough to give us an impression of the strengths and weaknesses of the two alterna-

¹¹⁰¹ Agency theory is, in my opinion, better viewed as a part of NIE, being a special case of the transaction costs based explanations of institutional outcomes. If it were 1975, we could consider approaching the problem from a Marxist perspective, but then we would probably end up concluding that Marxism is indistinguishable from Public Choice Theory; the capitalists derive power from their wealth, and they use this power to push the institutional outcomes in a direction that guarantees their continued wealth and power. This is an hypothesis that we will examine below, without any need to refer to Marx.

tive approaches. This means, most importantly, that there is no need to examine the history of each framework, or the roads not taken. Instead, the present discussion will be limited to the main thrust of the argument, often ignoring important discussions in the literature in favour of one simple answer. For example, rather than carefully examining the boundaries of Public Choice Theory, for example by discussing whether Gary Becker's work on the economics of the family¹¹⁰² is Public Choice Theory or merely the work of someone who is otherwise a Public Choice Theorist, we will prefer the simple top-down approach of defining Public Choice Theory as "the economics of politics"¹¹⁰³ and leave it at that. After all, we are not currently interested in examining non-political phenomena, so this perhaps overly restrictive definition can be chosen without creating problems for the present study.

In both of the application sections, i.e. in section 4.2 and in section 4.4, the general chronological approach of chapter 3 will be preserved, but without the same level of detail and without the same detailed sourcing. In other words, it is assumed that the reader is familiar with the previous chapter, and that the reader is prepared to go back to the relevant sections in chapter 3 in order to find the sources for the historical narratives mentioned in this chapter. Only when the argument in this chapter relies on materials not discussed in chapter 3, or not sufficiently discussed there, will we refer to sources in any great detail.

¹¹⁰² Eg. Becker (1981a).

¹¹⁰³ Van den Hauwe (1999), p. 604. Cf. section 4.1, below.

4.1 Public Choice Theory

4.1.1 Introduction

The essential difference between Public Choice Theory (hereafter: PCT) and the New Institutional Economics (or NIE) as we have seen it in the last two chapters is that PCT abandons the assumption that the government's objective function is some kind of social welfare function, or at least their perception of the social welfare function. Instead, public choice theorists assume that the individuals that make up the government act out of (enlightened) self-interest, for example by maximising their chance of re-election, by maximising their bureaucratic clout, or by maximising their incomes.¹¹⁰⁴ In this way, the black box of government is opened, something that NIE is capable of doing, but rarely does. After all, as long as a party to a transaction is an organisation rather than an individual, there are more transactions taking place within that organisation. So conceptually there is nothing stopping NIE from looking at the inner workings of government. In practice, however, such an approach is essentially the same as agency theory.¹¹⁰⁵ The problem is, however, that NIE models tend to study only one transaction at a time, making it difficult to explain the characteristics of the main transaction under consideration by referring to the transactions carried out inside the organisations that are the parties to the main transaction.

A second difference between PCT and NIE is that the former is explicitly public. Much of the endeavour of chapter 2, above, was to show how – theoretically – a theory of transactions in the private sector can just as easily be used to describe and understand transactions between the government and one or more private parties. We saw that that was a step that was by no means obvious, even though Williamson's original case study dealt with exactly such a situation.¹¹⁰⁶ After all, virtually every other example used by Coase, Williamson or others was set in the private sector, like the Fisher Bodies case study¹¹⁰⁷, the cases of the neighbouring farmers and the doctor who moved next to the confectioner that Coase discussed in his 1960 article¹¹⁰⁸, up to the luxury yacht discussed in Hart & Moore (1990).¹¹⁰⁹ Only later did such authors take the next step to argue that their earlier models would work just as well

¹¹⁰⁴ Cf. Mueller (2003), p. 1-2.

¹¹⁰⁵ Cf. the discussion of moral hazard in par. 2.7, above.

¹¹⁰⁶ Williamson (1976).

¹¹⁰⁷ Cf. par. 2.4.1, above.

¹¹⁰⁸ Cf. par. 2.2.2, above.

¹¹⁰⁹ Cf. par. 2.5.3, above.

when applied to the public sector.¹¹¹⁰ Public Choice Theory, on the other hand, has always been explicitly public in its focus. One of its inventors, Gordon Tullock, for example, begins his 2002 introduction of the topic by saying that “public choice is a scientific analysis of government behavior and, in particular, the behavior of individuals with respect to government.”¹¹¹¹ There is no suggestion, even, of extending the insights of PCT to the related topic of the behaviour of individuals with respect to management in and by private organisations. Public Choice theory is “the economics of politics”¹¹¹², or more elegant yet, “politics without the romance”.¹¹¹³

Finally, it is important to emphasise that, despite its name, Public Choice Theory is not, in fact, a theory. Instead, it is a collection of theories and models connected by a common perspective, as described here. Different authors have worked on different problems using different methods, sharing only the desire to use economic methods to examine “government failure”. As a result, PCT includes a much greater range of theories than NIE, from the highly analytical analysis of Arrow’s Impossibility Theorem¹¹¹⁴ to the almost sociological studies of government bureaucracy.¹¹¹⁵ However, in line with modern practice in economic science, they do tend to develop mathematical models whenever possible.

Because of the fragmented nature of PCT, it will not be possible to give a complete oversight of the field. For that, the reader is referred to more specialised resources.¹¹¹⁶ Instead, we will consider the Public Choice analysis of a number of distinct areas that are particularly relevant to the railway industry, most notably the theory of rent seeking, but also models of decision making under varying levels of democracy, as well as the issues of bureaucracy and interest groups. The theory of social welfare functions will not be discussed here, since it was already covered in section 2.2.3, above.

¹¹¹⁰ Eg. Hart, Shleifer & Vishny (1997) and Williamson (1999).

¹¹¹¹ Tullock (2002), p. 3. Cf. Mueller (2003), p. 1, who writes: “Public choice can be defined as the economic study of nonmarket decision making, or simply the application of economics to political science.”

¹¹¹² Van den Hauwe (1999), p. 604.

¹¹¹³ Buchanan (1979).

¹¹¹⁴ Cf. Arrow (1963) and Mueller (2003), chapter 23.

¹¹¹⁵ Mueller (2003), chapter 16, discusses how economists have attempted to turn Weber’s theory of bureaucracy into an economic model.

¹¹¹⁶ Cf. in particular Mueller (2003), Rowley & Scheider (2004) and Shugart & Razzolini (2001).

4.1.2 Rent Seeking

Economic rent is defined as “the excess amount earned by a factor [of production] over the sum necessary to induce it to do its work.”¹¹¹⁷ As such, it is neither good nor bad¹¹¹⁸, but simply a concept akin to the producers’ surplus in neoclassical microeconomics.¹¹¹⁹ Of course, public choice theorists are only interested in economic rents to the extent that their size depends on government policy choices.¹¹²⁰ Even then, there are good rents and bad rents, good rent seeking and bad rent seeking. A textbook example of beneficial rents are the rents created by intellectual property rights, such as patents and copyrights. Such rents are entirely the result of government intervention in the economy and, as long as the government does not overdo it, they serve “to promote the Progress of Science and useful Arts”¹¹²¹ by giving inventors and artists an incentive to pursue their activities.¹¹²²

Copyrights are a good example of the difficulty of distinguishing between beneficial rent seeking and the other kind. In recent decades, both European and American lawmakers have extended the duration of a copyright. Following the world’s first copyright law, the 1709 UK Statute of Anne¹¹²³, the US Copyright Act of 1790 granted the right for a period of 14 years from the moment of registration, subject to the possibility of having it renewed once.¹¹²⁴ By 1976, this had become 75 years or the duration of the life of the author plus 50 years, without any need for registration¹¹²⁵, which was extended by another 20 years in 1998.¹¹²⁶ When the European Union first got involved in the area of copyright law, the result was a similar extension from life plus 50 years to life plus 70 years.¹¹²⁷ These extensions were the result of aggressive lobbying efforts by the music industry and other beneficiaries, with the American act even referencing its most important sponsor in its official name. For a musician like Sonny Bono to seek to have the term of copyright “last forever”¹¹²⁸ is undoubtedly rent

¹¹¹⁷ Wessel (1967), p. 1222.

¹¹¹⁸ Cf. Tullock (2002), p. 43-44.

¹¹¹⁹ Cf. Katz & Rosen (1998), p. 132-133.

¹¹²⁰ Although note that the government aspect is not always incorporated in the definition. Cf. Brooks & Heijdra (1988), p. 2-3 for a survey. In fact, Buchanan (1983) deals with rent seeking and inheritances.

¹¹²¹ US Constitution Article I, section 8, clause 8.

¹¹²² Cf. Aghion & Howitt (1992).

¹¹²³ Act for the Encouragement of Learning, 8 Anne c. 19.

¹¹²⁴ Act of May 31, 1790, §1, 1 Stat. 124 (repealed 1831). Cf. Lessig (2001), p. 1061.

¹¹²⁵ Pub. L. No. 94-553, title I, § 101, 90 Stat. 2573 (1976). Cf. Breyer (1970).

¹¹²⁶ Sonny Bono Copyright Term Extension Act, Pub. Law 105-298, codified as 17 USC 302.

¹¹²⁷ Directive 93/98, OJ L 290, dd. 24.11.1993, p. 9-13, since re-enacted as a part of Directive 2006/116, OJ L 372, dd. 27.12.2006, p. 12-18.

¹¹²⁸ Lessig (2001), p. 1065, quoting 144 Cong. Rec. H9952 (1998) (statement of Congresswoman Mary Bono).

seeking, but it is difficult to say much more than that. It is impossible to say, for example, what the optimum length of a copyright would be, from society's point of view.

Or actually, it is possible to say one thing more. In *Eldred v. Ashcroft*¹¹²⁹, the appellants argued that the 1998 Act was unconstitutional. To be precise, they claimed that

the Copyright Clause must mean something when it says terms must be limited; and that to understand what it means, you must read it in light of its purpose. [It follows that] the terms "limited times" could not be read in the abstract, but must instead be read in light of the command that Congress exercise this power to "promote . . . Progress."'¹¹³⁰ And the simplest way to assure that the power only promotes progress is to require that the power only be used in exchange for progress. In exchange, that is, for producing something that has not already been produced. Thus, it is improper, under this reading, to extend the term of copyright to already existing works, since such an extension promotes nothing except the bottom line of (mainly) publishers.¹¹³¹

Unfortunately, the plaintiffs lost both before the District Court¹¹³², before the Court of Appeals¹¹³³ and before the Supreme Court.¹¹³⁴ However, from a public choice point of view, the argument is unimpeachable: What can be the economic benefit that society is to derive from an extension of the copyright on works that have already been created? Instead, there is only the welfare loss of 20 additional years of monopoly power.

This story of copyrights tells us everything we need to know about rent seeking. It not only gives examples of beneficial and harmful rents, but it also shows how a small but motivated interest group can move mountains. In these kinds of arguments about rents, both sides tend to suffer from collective action problems, where all actors have an incentive to free-ride off everyone else's efforts. Both the copyright holders and those paying royalties are many, and their individual stake in the outcome of the debate is not always large enough to warrant a great deal of attention. However, the former have a distinct advantage in that they are fewer, have higher stakes on average, and have access to pre-existing and well organised interest

¹¹²⁹ US Supreme Court, *Eldred v. Ashcroft*, 537 US 186, 123 S. Ct. 769, 154 L. Ed. 2d 683.

¹¹³⁰ US Constitution Article I, section 8, clause 8[, which states in full that: "The Congress shall have power (...) To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."]

¹¹³¹ Lessig (2001), p. 1066-1067 (Footnote in the original). Harvard Law professor Lawrence Lessig appeared as counsel for Eldred throughout the litigation.

¹¹³² *Eldred v. Reno*, 74 F. Supp. 2d 1 (D.D.C. 1999).

¹¹³³ *Eldred v. Reno*, 239 F.3d 3

372 (D.C. Cir. 2001); rehearing and rehearing en banc denied, 255 F.3d 849 (D.C. Cir. 2001).

¹¹³⁴ US Supreme Court, *Eldred v. Ashcroft*, 537 US 186, 123 S. Ct. 769, 154 L. Ed. 2d 683. It is noteworthy that the dissenting opinion was written by Justice Breyer, who had already voiced considerable unease in his 1970 law review article about the efficiency and morality of some aspects of copyright law, including the discussions at the time concerning what would become the 1976 extension of the copyright term. Cf. Breyer (1970), p. 323-329.

groups who speak for them. It is easy to see that this will often be the case when the question concerns a possible rent of \$ X – be it in the form of a direct cash payment, like a royalty, or otherwise – that is to be transferred from a large group to a smaller one. In that case, the smaller group will automatically have a higher average stake in the question, and therefore it will be easier for them to overcome the collective action problem.¹¹³⁵

4.1.3 Politicians

Now why is it that the US Congress should enact a law that is so obviously not conducive to “the general Welfare of the United States”?¹¹³⁶ As so often in economics¹¹³⁷, the answer lies in the incentives of the players, i.e. the senators and representatives. As for all elected officials in developed democracies, we assume that their overriding goal is to get re-elected.¹¹³⁸ How does an American politician get re-elected? By voting in a way that pleases the electorate, especially on issues that are of high importance to voters, and by voting in a way that pleases those that contribute to his re-election campaign. In other words, the way in which politicians are assumed to vote deviates from a hypothetical referendum vote in two ways: the politician gives more weight to voters for whom the issue is particularly salient, and the politician gives more weight to those who – not to put too fine a point on it – are willing to put their money where their mouth is. It is important to realise that, in a zero-transaction cost world, this would be social welfare enhancing, since it allows for certain Coasian bargains to be made. Even in a world with non-zero transaction costs, we can say that by mostly giving voters their way on issues that are particularly important to them, the politician is likely to achieve an outcome that is similar to what the voters would have bargained for in a zero-transaction cost world.¹¹³⁹

Unfortunately, the use of the government in this way is not without costs. The sums spent by politicians on their campaigns to win favour among voters, and the sums spent by private parties competing for favour among politicians are being wasted, in the sense that they

¹¹³⁵ Cf. Olson (1965), the classic work on collective action.

¹¹³⁶ US Constitution Article I, section 8, clause 1.

¹¹³⁷ Cf. Elster (2007).

¹¹³⁸ Cf. Downs (1957). In developing democracies, there might be a tradeoff between maximising the amount of money obtained through theft and corruption during one’s current term of office, and maximising the likelihood of winning another term, i.e. between being a “roving bandit” and being a “stationary bandit”. Cf. McGuire & Olson (1996) and Olson (2000). Like all bandits, the corrupt politician would like to be a stationary bandit, but unfortunately this usually means being somewhat less corrupt in the short term. Hence the trade-off.

¹¹³⁹ Cf. Tullock (1970) and Wilson (1969).

could have been deployed in a more productive manner.¹¹⁴⁰ These rent seekers are caught in a kind of prisoners' dilemma, wherein they would all benefit from cutting their expenditure in half, if only they could find a way to agree to do so. After all, all that matters for any individual lobbyist or campaign manager is their spending relative to the spending of their competitors. Halving all spending leaves everyone in the same position, except that they now have a sum of money available for more productive use.¹¹⁴¹

4.1.4 Bureaucracy

If politicians want to get re-elected, what do civil servants want? What are their incentives? A common answer among economists and other social scientists, going back to Weber (1920), is that civil servants want power.¹¹⁴² A problem emerges, however, when we consider what this so-called "power" actually is. In this thesis, we have already encountered power in the legal sense, meaning "a legal position of a person vis-à-vis other persons having a legal relation that can, under the applicable legal rules, be changed by a fact under the volitional control of the first-mentioned person."¹¹⁴³ While this definition certainly covers the power that a superior has over his underlings in a bureaucracy, it is far from clear that it is something that can be maximised the way market power – or at least market share – can be.¹¹⁴⁴ In addition, the definition does not cover other ways that someone in an organisation might come to have power, most notably the power that comes from having (scarce) expertise or the bureaucratic control over scarce resources.¹¹⁴⁵ A simple fix for all these power problems is to simply as-

¹¹⁴⁰ Cf. Congleton (1988) and Cowen & Tabarrok (1999) for a discussion of the difference between the sums spent on rent seeking and the actual welfare losses caused.

¹¹⁴¹ Cf. Van Winden (2004).

¹¹⁴² Cf. Niskanen (1971) and Mueller (2003), ch. 16.1.

¹¹⁴³ Ruiter (2008). Cf. section 2.1.1, above.

¹¹⁴⁴ Of course, for an economist any such problem can always be solved by postulating a suitable utility function, but even then the problem remains of how to map power relationships onto utility values.

Alternatively, we could start from the concept of state space that was introduced in section 2.1, above, and postulate that person X has power over person Y if X is able to unilaterally – i.e. without Y's consent – alter Y's utility in one or more states of the world. The power of X over Y would then be proportionate to the size of the utility differential that X is able to achieve in a given state of the world, and to the probability that that state of the world will be realised. However, not only is it dubious whether this model accurately describes what people actually maximise when they maximise power, it is also highly abstract and nearly impossible to operationalise. For a general discussion of the "messy" concept of power, cf. Zimmerling (2005).

¹¹⁴⁵ Cf. Mueller (2003), p. 361.

sume that the bureaucrat wants to maximise his budget, an assumption that brings us back to the safe zone of bargaining over money.¹¹⁴⁶

Another plausible assumption about bureaucrats is that they seek to maximise “slack”, which can be defined as “the difference between [their] revenue and the minimum cost of production”.¹¹⁴⁷ In other words, what they want is to be left alone, to minimise the risk that they are going to get in trouble for going over budget, or for slouching around just a little bit too much. All else equal, an increase in budget would, of course, mean an increase in slack. However, in practice an increase in budget usually comes at a price, in the form of added responsibility. Being given more money is not necessarily advantageous if you are also asked to do more with that money. Unfortunately, the problem with the assumption of “slack-maximisation” is that, unlike the budget, the minimum cost of production is unknown to anyone but the bureaucrat.

To a point, each of these assumptions is open to empirical verification. Since they predict a negotiation between principal and agent, between the bureaucrat and his superior, however, the empirical implications are not very different from those offered by agency theory or by Williamson’s assumption that the parties to a transaction act opportunistically, making the public choice theory of bureaucracy of limited use as a benchmark for the NIE of bureaucracy.

For completeness, it should be noted that many scholars who have studied public bureaucracy empirically have tended to take a more forgiving view of the actors who work within its domain.¹¹⁴⁸ Phrasing it in the language of economics, they have generally found that civil servants have, to some extent, internalised the public norms that govern their work. The social welfare function, or their perception of it, is an element of civil servants’ utility function. Or, in normal English, civil servants care about doing a good job. This is of course a most interesting perspective, from the point of view of Public Choice Theory, because it amounts to saying that civil servants are being selfishly unselfish. It is left as an exercise to the reader whether such a finding would refute PCT or confirm it. In any event, we shall see in section 4.2, below, whether there is evidence of selfish unselfishness in our narrative of railway history.

¹¹⁴⁶ Cf. Mueller (2003), ch. 16.2. Cf. Dahl (1957) and Dowding (1991) for attempts to be more precise about the “concept of power”.

¹¹⁴⁷ Wyckoff (1990), p. 35.

¹¹⁴⁸ Eg. Wilson (1989) and Hill (1992).

4.1.5 Dictators and Monarchs

If politicians want to get re-elected, and bureaucrats want to be left alone, what do dictators want? The straightforward answer – at least for an economist – is that a dictator wants to steal whatever he can. This leads to Mancur Olson’s analysis of dictators in terms of “roving bandits” and “stationary bandits”¹¹⁴⁹, depending on how likely they are to be able to hang onto power. This – Nobel Prize winning – analysis is a prime example of what Lionel Robbins meant by “deductions from simple assumptions reflecting very elementary facts of general experience” a non-obvious conclusion.¹¹⁵⁰ In fact, it leads Olson to conclude that in some circumstances a country is better off with a dictator who is firmly entrenched than with a dictator who might be overthrown, a result that is simply marvellous in its counterintuitiveness.

There is just one problem with this assumption: it does not seem to fit very well with the historical monarchs we encountered in chapter 3. Even though they were the ultimate stationary bandits, in the sense that they expected their dynasties to last for centuries more¹¹⁵¹, they appropriated a much smaller part of their countries’ surplus than a strictly individualist present-value-of-theft maximisation model would suggest. They could have gotten away with so much more, without either risking the stability of their regime, or causing a noticeable decline in the country’s overall wellbeing. They did not have Swiss bank accounts, like modern dictators do. Even lavish palaces had gone out of fashion.¹¹⁵² Of course, to some extent this had to do with the greater constraints within which these monarchs had to work, but even when we take that into account, we are forced to conclude that the historical accounts of the personalities of most 19th century monarchs in no way justify describing them as “bandits”.

This means that – in economic terms – they were altruistic agents for their subjects, albeit only to a point. The wellbeing of the people enters into the monarch’s utility function with some non-zero weight. While this is not fundamentally problematic, it goes against the

¹¹⁴⁹ Cf. McGuire & Olson (1996) and Olson (2000).

¹¹⁵⁰ Cf. chapter 1, above.

¹¹⁵¹ The obvious exception is, of course, the house of Bourbon in France, where neither Louis XVIII nor Louis-Philippe were ever really secure in their positions.

¹¹⁵² Fyffe (1895), in discussing the smaller German states before the Napoleonic wars, writes: “Forty years earlier, Versailles, with its unfeeling splendours, its glorification of the personal indulgence of the monarch, had been the ideal which, with a due sense of their own inferiority, the German princes had done their best to imitate. (...) Yet even on the Rhine and on the Moselle the influence of the great King of Prussia had begun to make itself felt. The intense and penetrating industry of Frederick was not within the reach of every petty sovereign who might envy its results; but the better spirit of the time was seen under some of the (...) princes in the encouragement of schools, the improvement of the roads, and a retrenchment in courtly expenditure.”

natural cynicism of public choice theory. It also causes a problem of indeterminacy: it is one thing to assume that an individual will know how to maximise his own utility, but quite another to assume they will know how to promote the wellbeing of a nation. It is obvious that the monarch will make mistakes in his attempts to maximise social welfare, but an assumption that his utility function consists of a combination of personal wellbeing and national wellbeing does not tell us anything about the kinds of mistakes he will make. Also, for many monarchs it is perhaps more appropriate to compare them to bureaucrats. French kings like Louis XVIII, Charles X and Louis-Philippe were seemingly less interested in theft or national improvement than in maximising their power over the country, in competition with the various other government bodies and institutions, and in competition with the people from whom all legitimate authority derives.¹¹⁵³

4.1.6 Conclusion

Of necessity, this discussion of Public Choice Theory was rather brief and fragmented. It should, however, suffice for the moment, in order to provide a basis for the renewed examination of the history of the railways in the next section. Suffice it to say that when interest groups, politicians, bureaucrats and monarchs all seek to maximise their individual utility, instead of working to maximise social welfare, there is no reason to believe that the resulting policies will be conducive to the wellbeing of the nation. That is not to say that they will not be, but only if voters vote rationally, if the interest groups roughly cancel each other out¹¹⁵⁴, if bureaucrats are sufficiently supervised and if monarchs are either powerless or benevolent.

So how can we determine whether Public Choice Theory is a suitable explanatory framework for the railways? It seems difficult to make a list – like the one of section 2.7 – of elements that are necessary and sufficient for the theory to work. After all, the requirements of PCT do not reach much further than the requirement that all actors are selfish enough to focus on their own utility and their own incentives. This does not seem like a very unreasonable assumption.¹¹⁵⁵ However, for PCT to have more explanatory power than NIE, it needs to yield

¹¹⁵³ Note that the French Constitution of 1814 was careful to reiterate that “all authority in France resides in the person of the king” and that “our first duty towards our subjects was to preserve, in their own interest, the rights and prerogatives of our crown”. Of course the truth is that after 25 years of revolution, not even their authors believed that such statements corresponded to the political reality.

¹¹⁵⁴ Cf. The Federalist No. 10.

¹¹⁵⁵ Albeit an assumption that can be neither proved nor falsified. Cf. Boland (2002), p. 15-16 and section 4.2.5, below.

different predictions, meaning that cynicism needs to be relatively common among those actors who choose the governance structures.¹¹⁵⁶

In those cases where Public Choice Theory suggests that actors have an incentive to act towards the common good, it does little more than borrow the definition of “the common good” from NIE or some other framework. On the other hand, it is exactly when the theoretical optimum is not achieved that PCT might contribute – either as an alternative approach or in combination with NIE – by explaining why this should be so. In that sense, PCT shines when NIE fails. However, it is important to emphasise that the two theories are not perfect inverses: it is not necessarily true that any failure to achieve the optimum is explained by PCT. Like NIE, PCT still implies a high level of rationality, only at the level of individuals instead of organisations. Later in this chapter will we relax that assumption, to discuss the ideas known as institutional lock-in or path dependence.

¹¹⁵⁶ For example, it is well known that many in Mrs. Thatcher’s government were big fan of the television show ‘Yes, Minister’. Thatcher herself even did a short sketch with the actors once, on the occasion of presenting them with an award. The fact that she clearly agreed with many of the assumptions behind PCT makes PCT a much more obvious choice for explaining her behaviour in office.

4.2 Public Choice Theory and Railways

In this section, we will retell – in a more concise form – the story of chapter 3, from the perspective of Public Choice Theory. Instead of explaining the behaviour of actors in terms of parties looking to find efficient governance structures to govern their transactions, we will now focus on the individual incentives of the decision makers.

4.2.1 Beginnings

As the reader will recall, we identified two distinct manners in which the story of the railways can be begun. It can be told as a story about technology, or it can be told as a story about institutions. As it happens, both perspectives have a distinct public choice component.

The question of incorporation is not as such within the ambit of public choice theory, being almost entirely of a private nature. However, the same cannot be said about the government's involvement in the early innovations in company law. We know, for example, that one of Van Oldenbarnevelt's concerns in creating the Dutch East India Company in 1602 was that he wanted to avoid a situation where the Dutch trade with the East Indies would be controlled by a small group of powerful merchants. In his estimation, such an outcome would have undesirable political consequences, something he sought to prevent by creating a company with more widespread stock ownership.¹¹⁵⁷

More generally, these early companies were extremely beneficial both to the merchants who were fortunate enough to control them and to those in control of the state. Their positions were protected at the expense of those who might seek to challenge their commercial or political dominance. The merchants were given various legal monopolies, protecting them against domestic competition. In return, the actors controlling the state were given large sums of money, as their share of the profits. Even when such cash does not flow into the politicians' personal coffers, having control over what it is spent on is still an extremely useful way to buy political support in a way that tax income is not. After all, taxpayers tend to want some measure of influence over public spending in return for their contributions.¹¹⁵⁸ When the

¹¹⁵⁷ Cf. Den Tex (1962), p. 384-420.

¹¹⁵⁸ As the Rev. Johnathan Mayhew put it in a sermon in 1750: "No Taxation Without Representation."

money is seemingly “free”, on the other hand, the beneficiaries of such spending tend to be the only voices heard.¹¹⁵⁹ It is mostly for this reason that states that are blessed with significant endowments of natural resources are likely to be much less democratic than their overall welfare levels would suggest.¹¹⁶⁰ In 17th century England, James I and Charles I used these monopolies as a source of income so that they would not have to ask Parliament for permission to levy taxes, a problem that the 1623 Statute of Monopolies¹¹⁶¹ failed to fully address.

This conflict was only resolved in the relatively unsettled political climate following the Glorious Revolution, when it was decided that monopolies would henceforth require an Act of Parliament. This explains why James Watt, having perfected the steam engine, turned to Parliament for the rent that he felt was his due as a reward for his ingenuity. In a series of private Acts starting in 1769 he got his due: the exclusive right to his invention up to about 1800.¹¹⁶²

In the UK, the Great Reform Act of 1832¹¹⁶³ functions as something of a natural experiment when it comes to the relative influence of various lobbying groups on the one hand, and the pandering behaviour of politicians on the other hand. However, in our case study the evidence is inconclusive. After all, while it is clearly true that the railways were received in a more welcoming manner after 1832 than before that year, this can be more plausibly ascribed to the greater familiarity of the various actors with the new technology, as a result of the success of the Stockton & Darlington and the Liverpool & Manchester. That said, it is clear that many of the difficulties encountered by those early ventures fall comfortably within the ambit of Public Choice Theory: We have members of the House of Lords working to protect their interests as land owners or canal proprietors, we have other politicians noticeably unhappy about the technological uncertainty that so suddenly invaded their tranquil lives, thus interfering with the slack they had so carefully won in the pre-railway era, and we have various interested parties, ranging from coal miners to would-be railway entrepreneurs and customers¹¹⁶⁴ trying to overcome this inertia. In the end, the proponents of the railways prevailed, but why?

¹¹⁵⁹ Of course the hidden source of such income is the consumer who pays the monopoly price for the imported goods. As such, a government monopoly is akin to a highly inefficient consumption tax.

¹¹⁶⁰ Most recently Tsui (2010), who finds that “discovering 100 billion barrels of oil (approximately the initial endowment of Iraq) pushes a country’s democracy level almost 20 percentage points below trend after three decades”.

¹¹⁶¹ Statute of Monopolies, 21 Jac. 1, c. 3.

¹¹⁶² Cf. Muirhead (1858), ch. XIII-XIV and Marsden (2002), chapter VII.

¹¹⁶³ Representation of the People Act 1832, 2&3 Wil. IV c. 45.

¹¹⁶⁴ One of the key backers of the Liverpool & Manchester was Joseph Sandars, a wealthy corn merchant. Cf. Hylton (2007), p. 37.

Answering this question from a PCT perspective is quite difficult, because the relevant government actors had incentives that were quite complex. None of the discrete alternatives that were discussed in the previous section were already completely developed in the 1820s and 30s. Even after the Great Reform Act, the accountability of members of the House of Commons to their constituents was still highly imperfect.¹¹⁶⁵ As a result, they could expect to keep their seats for long periods of time, as long as they did not displease the – often aristocratic – local and national leadership of their party too much. In other words, they were a lot like civil servants, only with greater protection against dismissal.

This makes it difficult to understand why these politicians ever did much of anything. The Great Reform Act was to a large extent motivated by fear of political upheaval, such as the Peterloo massacre of 1819, an obvious enough explanation from the point of view of Public Choice Theory. But why would MPs and peers chose the interests of urban factory owners and merchants, who were greatly underrepresented in parliament even after 1832¹¹⁶⁶, over the interests of the landed gentry and nobility? The only explanation seems to be that politicians and bureaucrats alike do – at least part of the time – seek to promote the general welfare as they see it. After all, it would be difficult to argue that the promotion of the railways was necessary to avoid civil unrest the way the enactment of the Great Reform Act arguably was. In any event, there is no historical evidence for the proposition that the politicians of the time were motivated by any such concern. Instead, they seem to have been motivated by a desire to promote the wellbeing of city dwellers in a manner well out of proportion to their electoral significance. The rural aristocracy ended up with enormous payoffs and little else.¹¹⁶⁷

This observation cannot be generalised immediately to the other imperfect democracies we have studied. In those countries, parliamentary representation did not favour the rural counties and the nobility like it did in the UK. Rather, they were imperfect because they tend-

¹¹⁶⁵ The vote was still restricted to a minority of men, about 650,000 in number, on a total population of about 32, 5 million. Cf. Evans (1994), p. 38. Of course, as such a small franchise does not mean that a healthy system of political competition cannot exist. However, evidence suggests that for much of the rest of the 19th century, the aristocracy continued to control parliament. Because of a combination of feudal loyalty, patronage and outright corruption, the middle classes tended to vote according to the preferences of their betters. Cf. Whitfield (2001), p. 203. Remember, the secret ballot was only introduced by the Ballot Act of 1872, 35&36 Vict, c. 33.

¹¹⁶⁶ Even after 1832, great industrial cities such as Manchester and Leeds, which now for the first time formed distinct constituencies, elected as many MPs as an ordinary rural county. To make matters worse, the criteria for admittance to the franchise were more generous in counties than in urban boroughs until the distinction was abolished in the Representation of the People Act 1884, 48&49 Vict., c. 3.

¹¹⁶⁷ Cf. Wolmar (2007), p. 63: “Between them, the Grand Junction and the London & Birmingham paid more than £ 700,000 in payments to the landowners to acquire what in effect were narrow slivers of land used, if at all, for agriculture. This enormous sum represented a fifth of the two railways’ combined initial share capital of £ 3.5 m.”

ed to preserve a significant amount of power for the king, who did not have any immediate reason to favour the interests of land owners over city dwellers. Instead, these kings tended to favour those who supported them, especially when the king's position was less than perfectly secure. In restoration France, that meant an arch-conservative alliance between the King, the nobles and the Catholic Church, whereas after 1830 the new King's alliance lay more with the liberal bourgeoisie that controlled parliament. That same urban class was highly influential in Belgium and the Netherlands, though with vastly different consequences for the railways. In Belgium, as we have seen, the alliance between the Catholic Church and the liberal bourgeoisie, which had supported the movement for independence, lasted long enough for the Church to consent to an ambitious plan for economic development through railway construction. In the Netherlands, on the other hand, the wealthy tended to be suspicious of the King's initiatives, and reluctant to support any great tax-funded state expenditure at a time when the treasury was already in such poor shape.

In Germany, finally, the railways had the least chance of being able to marshal democratic forces in their support. In most German states, there was no democracy to speak of, and the hereditary rulers were little inclined to listen to the advice of mere merchants. As a result, the railways were often dependent on the whim of the monarch. A stark example of this is the difference in policy between the later German emperor Wilhelm I, while he was still crown prince of Prussia, and his father, King Friedrich Wilhelm III. In the 1830s, Prince Wilhelm served as General Governor of the Rhine Province and Westphalia¹¹⁶⁸, and in that capacity he was very active in promoting the railways in that part of the country.¹¹⁶⁹ At the same time, his father in Berlin was seemingly reluctant to commit any great amount of money or political capital to the new technology. Given how much their material and political fortunes were connected, it is difficult to see how the son could be exposed to different incentives than his father. Even if economic science sometimes assumes that individuals have an attitude of *après moi, le déluge*¹¹⁷⁰, this assumption seems uniquely implausible for prominent aristocratic and noble families such as the Hohenzollern royal family of Prussia. It follows that one or both of two things must have been the case: either a) father or son or both incorrectly perceived the interest of the family in the development of the railways, or b) either father or son or both did

¹¹⁶⁸ Cf. Steitz (1974), p. 124.

¹¹⁶⁹ Cf. Steitz (1974), p. 109, discussing the opening of the Deilbachbahn in 1838, which was named Prince Wilhelm Bahn in his honour, as well as p. 153n, 159, 179, and 202, all of which give examples of the Crown Prince's efforts promoting the railways by lobbying in Berlin or by encouraging the efforts in his own provinces.

¹¹⁷⁰ Cf. Samuelson (1958), who introduces a model of solidarity between overlapping generations in order to avoid the conceptual problems caused by assuming that people only care about what happens during their own lifetime.

not put the interests of the family first to the extent that Public Choice Theory would tend to assume.

Overall, it appears as if the groups that were hurt by the introduction of the railways, and by the industrial revolution generally, did not resist as much as PCT would have predicted. During the relevant time period, they were still in almost complete control of most of the countries under consideration, but failed to use their power to defend the status quo. The Great Reform Act of 1832 did not change the power of the hereditary House of Lords to block any and every piece of legislation, a power that they would continue to enjoy until 1911.¹¹⁷¹ In France, the *Chartre* of 1814 established a similar Upper House of Parliament¹¹⁷², although in those days the French parliament of course had significantly less power than its counterpart across the Channel. In Prussia, there was a pitched battle between different government ministries and departments, with the department for the Mails, supported by the Ministry for Public Works, standing up for the interests of the network of carriages they operated¹¹⁷³ and on the other side the Ministries for Finance and Defence.¹¹⁷⁴ However none of this resistance rises to the level of an organised resistance against a potentially wide-ranging reform, which is the least of what one might expect from a relatively small group of wealthy and politically savvy aristocrats. The most plausible explanations are that these individuals did not put their own interests first to the extent assumed by the models and that they failed to fully appreciate how much their position was under threat.

4.2.2 Consolidation

During the period of consolidation, the railway companies are no longer fighting against the vested interests of those who control the state. Instead, the railways had become associated with those same vested interests. After the first English railway boom in the 1840s, virtually anyone with money owned shares in the railways¹¹⁷⁵, and the railways in other countries were not far behind in attaining such dominance. As we have seen, in Belgium much of the railway

¹¹⁷¹ Parliament Act 1911, 1&2 Geo. 5, c.13.

¹¹⁷² Art. 24 – 34.

¹¹⁷³ Cf. Holterman (2009), p. 113 and Steitz (1974), p. 73-75, 169.

¹¹⁷⁴ Cf. Holterman (2009), p. 113.

¹¹⁷⁵ That is, if they did not own their own railway. The Duke of Sutherland, for example, had a railway built in the north of Scotland so that his seat at Dunrobin could be reached by rail. The Duke's railway, which connected Golspie and Helmsdale over a distance of 19 miles, was donated to the Highland Railway in 1884 and proved exceedingly useful for the supply of the Royal Navy base at nearby Scapa Flow during World War I, thus proving once again that the whims of Victorian noblemen can have their uses. Cf. Wolmar (2007), p. 127.

system ended up in the hands of two powerful financial consortiums. In all countries under consideration, with the exception of the UK, the government had a financial stake in the railways, although that financial interest usually did not translate into the power to make decisions. It will be convenient to start by considering the rent-seeking activities by the railways and their owners, leaving the discussion of government-initiated reforms for the next section.

The key rent that was up for grabs throughout the history of the railways is, of course, the concession to build a railway connecting two places over a given route. After the failure of shared use of the Stockton & Darlington, this automatically also implied the concession to operate a railway service on the tracks so constructed. As discussed in chapter 3, this concession could only be obtained from the government; even if it was legally permissible to build a railway without a concession, it was almost impossible to do so except for very short connections. For this reason, railway construction inevitably meant some rent-seeking.

In Britain, the initial procedure for awarding concessions was extremely vulnerable to rent-seeking, because it was ad hoc and offered many distinct points of access for rent seekers. Every individual peer or MP was a potential lobbying target. Part of the purpose of Gladstone's streamlining efforts in 1844 was to reduce the number of people involved in the examination of applications, so that something might actually get done amidst all the talking of all the groups looking to influence every decision. Instead of having proposals examined by Parliament immediately, the 1844 Act arranged for them to be examined by an advisory board first, where an objective assessment of their desirability could be made.¹¹⁷⁶ Even at that point, the railways were too influential for Gladstone to dare suggest making the board's recommendation in any way binding, and after it started its work, the railway companies grew successively more displeased with the result. In July 1845, little more than a year after it had started its work, they succeeded in having the advisory board abolished.¹¹⁷⁷

Even at that early stage, the incumbent railway companies were already looking for a second kind of rent, in addition to the prize of winning a concession. Among the instructions of the Advisory Board was that it should "discourage unnecessary competition"¹¹⁷⁸, including – but not limited to – proposals that were only motivated by the desire to interfere with other,

¹¹⁷⁶ Railway Regulation Act 1844, 7&8 Vict., c. 85.

¹¹⁷⁷ Cf. Wolmar (2007), p. 89-90.

¹¹⁷⁸ Wolmar (2007), p. 89.

more beneficial, railways. What the incumbent railway companies wanted, of course, was for the board to protect them against each other and against new entrants, while maintaining their right to discourage legitimate new entrants from entering the market by whatever means necessary.

As the century progressed and the market matured, new railway construction was less of a threat to the major incumbents.¹¹⁷⁹ Instead, they looked to the government for protection against each other, i.e. against unprofitable competition. The simplest way to avoid such threats to profit is of course to merge, a method that had an honourable pedigree in the empire building of George Hudson in the 1840s. While competition law as such did not exist yet in the 19th century, the British railway companies, having been created by private act of Parliament, needed a further private act to allow them to merge.¹¹⁸⁰ Such private acts were, of course, highly valuable, and therefore the reason for much rent seeking and other lobbying.¹¹⁸¹ Taking this power away from Parliament and placing it with some kind of railway regulator would of course have been more efficient, since it minimises the opportunities for rent-seeking, albeit at the cost of an increased likelihood of capture. However, throughout the 19th century the precedent of 1845 remained in place; everyone in power benefited too much from being lavishly wined and dined by the railway interests for them to simply give up their leverage over the industry. The crown only gained some measure of influence over the sector in the post-World War I reforms.¹¹⁸²

In the rest of Europe, the legislature had never had quite so much to say about the railways. Generally, they had been called in to ratify proposals for reform, as well as to appropriate funds for subsidies, but little else. The initiative was with the executive branch of government. That is where the civil servants were who had regular dealings with railway executives. That is where the railway companies had to go for a concession, as well as for a variety of other assistance.

As we have seen in section 3.2.2, the history of the Dutch railways in the second half of the 19th century did not involve a great deal of rent seeking. Quite the contrary, it was the

¹¹⁷⁹ The exception to that rule is, of course, the construction of new main lines, such as the construction of the Settle & Carlisle in the 1870s, which offered a third main line connection between England and Scotland, in competition with the East Coast and West Coast main lines.

¹¹⁸⁰ Cf. London & North Western Railway Act (1846), 9&10 Vict., c. cciv, which allowed the merger between the Grand Junction, the London & Birmingham and the Manchester & Birmingham railways.

¹¹⁸¹ Cf. Tomlinson (1915), who discusses the many mergers and acquisitions that formed the basis of the London & North Eastern Railway.

¹¹⁸² Ministry of Transport Act (1919), 9&10 Geo. 5, c. 50 and Railways Act (1921), 11&12 Geo. 5, c. 55.

government who, from time to time, tried to get the railway companies to build more railways, to run more trains, to innovate. However, the railways were only inclined to do so if they were given significant inducements, in the form of government participation in the equity capital, minimum return guarantees, and even government construction of certain infrastructure, such as bridges. Much to the surprise of absolutely no one, these inducements tended to be one-off; as soon as the railway companies had held up their end of the bargain, they sunk back into blissful inactivity.

In France, on the other hand, the railway companies sought whatever rents they could get their hands on. Arguably, government rents had been the basis of the French railway industry almost from the start, with trains as a source of income of only secondary importance. The first reform, which took place in the 1850s, set the stage for what was to follow by showing the railway companies that they would be bailed out if they got in trouble. This is of course the ultimate moral hazard: the railways take the profits, the government takes the losses, but without any corresponding power of the latter over the former to make sure they run their business with an eye on losses as well as profits. The weak lines were merged with their profitable cousins to form vaguely profitable companies, and if they went bankrupt anyway, they were purchased by the government. This made it clear to everyone involved that the secret to high profit was to persuade the government to let you get rid of the “lemons”, or to sell you a company free of them.¹¹⁸³ Alternatively, you could take advantage of the moral hazards involved by using a high risk-high reward business model, for example by building a railway whose profitability was heavily dependent on high economic growth along the route, or on keeping construction costs unrealistically low.

In the period after the grand agreements of 1859 there are too many examples of successful lobbying by the big six to recount here. The most important one was the debate over deconcentration in the 1870s, when the suggestion was made that the appropriate lesson from the defeat of 1870 was that France should have smaller railway companies that relied less on government support and direction. In the new democratic French parliament this suggestion was easily defeated by the concerted opposition of the big six. When the sector finally was reformed in 1883, it was in a manner that made the lives of the big six easier, and even then the government had to add plenty of enticements.

¹¹⁸³ Note how the *nouveau reseau*, which contained many lines of dubious profitability, was not built by the big six, but by separate companies using government guarantees.

It is not necessary to say much about Belgium at this point, since the peculiarity of its railway history in this time period has been sufficiently discussed in chapter 3. The two big banking houses of *Société Générale* and *Banque de Belgique* owned both a large part of the railway industry and a large part of parliament and government. As a result, it was never likely that the government would do something to jeopardise these private parties, and when they ultimately ended up in trouble, their railway assets were purchased by the government at a generous price.

Germany, finally, continued to have a relatively undemocratic system of government, which still caused problems for the lobbying options of the railway companies. It is important to remember that this was true for almost all German states, regardless of whether they chose to entrust the railways to private companies or not. The French Emperor Napoleon III was to a significant extent forced to compromise with the railway companies, since he needed their support for the economic and military development of his country, a support that he could not obtain through less friendly means. The same was not true for many German Princes, including most notably the King of Prussia.

To the extent that the German Princes left the railway companies alone, or even used subsidies to convince them to expand, this reflects the fact that they were more interested in slack maximisation than in power maximisation. Secure in their positions against domestic threats¹¹⁸⁴, and honour bound to resist any temptation to engage in “banditry”, they were not interested in controlling every aspect of society. Arranging for private companies to build and operate the railways meant that the authorities were spared the burden of raising the funds to do so themselves, the burden of developing the organisational capacity to operate and supervise the railways, and the financial risk that the project might go wrong. Instead, they could simply lend whatever support they were capable of putting together without too much stress, and leave the rest of the worry to the managers of the company. Of course, this exposed them to claims from the railway interests saying that the government could and should do more, but the rulers would give in to such lobbying only when giving in was less trouble than resisting.

Concluding, we can say that Public Choice Theory works well enough for the period of consolidation, as long we overlook the uncertainty regarding the appropriate perspective. After all, for almost all of the relevant players an argument can be made that they are best

¹¹⁸⁴ With the notable exception of the revolutionary year 1848, the German Princes of the 19th century would have been most worried about threats against their position stemming from Prussia or Austria, rather than threats coming from within their own states.

treated as bureaucrats instead of the politicians or monarchs that they are supposed to be. This is particularly problematic because there seems to be no easy way to decide *ex ante* which frame of reference to use; the only thing to do is to see which one works best.

4.2.3 Nationalisation

Historically, the initiative for nationalisation has come from either the government or the railway companies. Other stakeholders, such as the competitors and customers of the railways, have tended to stay on the sidelines, at least in the countries examined in this thesis.¹¹⁸⁵ Only when one of both sides had decided to push for a policy of nationalisation did the other stakeholders sometimes speak out in support.¹¹⁸⁶ It should be obvious that any railway-driven reform quite comfortably fits in the public choice narrative. If the reform is backed primarily by the government, its Public Choice pedigree remains to be determined.

The obvious example of a reform where the impetus came from the sector is Belgium, of course. The banking houses that controlled the railways got into financial difficulties, so they sought to off-load their railway interests to the state, who were happy to accommodate because they were concerned about the strategic importance of the railways, particularly from a military point of view, because they were concerned about the profitability of the lines they already owned¹¹⁸⁷, and because they were concerned about protecting their relations with the *Société Générale* and the *Banque de Belgique*.

In all other countries under consideration, nationalisation happened to a greater or lesser extent over the objection of the railway companies. The amount of trouble each government ran into is consistent with a Public Choice narrative: the bigger and more profitable the railway companies were, the more they were able to resist and postpone nationalisation. In Germany, the railways were still small when the government started taking them over. Very few consisted of more than a trunk line with a few branches. The Köln & Minden Railway, for example, which was already discussed at the beginning of chapter 3, combined its basic trunk line connecting Cologne and Minden with additional lines to the Dutch border at Em-

¹¹⁸⁵ For an example of customer agitation leading to a serious increase in government regulation of the railway industry, cf. the history of the Interstate Commerce Commission in the United States, which was created in response to protests by Western farmers in the so-called Grange movement. Cf. Dempsey (2003), part V.

¹¹⁸⁶ Cf. eg. Devys (1910), p. 27.

¹¹⁸⁷ The one notable example of the government going against the wishes of the railway interests, the nationalisation of the *Grand Central* in 1897, was motivated by the desire to add a connection to the port of Antwerp to the state-owned network. Cf. Devys (1910), p. 29.

merich (NRS) and at Venlo (SS), south towards Siegen and Gießen, and – most importantly – north, parallel to the original trunk line, via Münster, Osnabrück and Bremen to Hamburg. This gave them a powerful position in the Ruhe area, local competition notwithstanding, but hardly any kind of real political weight with regard to the whole of Germany, or even Prussia. What is more, these construction projects had been financed with generous equity participation from the Prussian government, who had in any event made sure to retain significant power through share purchases on the open market. As a result, when the KME was nationalised in 1879, the Prussian government already owned 74% of the shares and encountered little difficulty.¹¹⁸⁸

In the Netherlands, the railways were relatively easy to nationalise because they were not very profitable.¹¹⁸⁹ In the UK, the problems were somewhat greater, since some of the British railways were big enough and profitable enough to be able to wield significant political influence, sufficient to enable them to put off nationalisation until after World War II, while in France the big six allowed the formation of the SNCF to go ahead, but only after decades of dismal profitability, and even then at an exorbitant price.

If we observe that nationalisation occurs at the insistence of the government, the degree of difficulty they encountered is naturally a part of any PCT explanation of that decision. However, there has to be more. As long as the government actors incur a non-zero cost as a result of nationalisation, in the form of a reduction in slack, an increase in uncertainty, a decrease in various forms of corruption, or quite simply the effort cost of persuading everybody involved to cooperate, the decision to nationalise has to involve an offsetting benefit for the actor that makes it all worth while.

The most obvious benefit that the government actors will obtain through nationalisation is power. This raises the question of whether nationalisation actually resulted in an increase in power for the same actors who were instrumental in making this reform happen, as well as the questions of whether these actors were motivated by the desire for power for its own sake and, if not, whether their gain in power resulted in other benefits for these actors. In other words, if the actors who pushed for nationalisation did not gain any power, the desire for power cannot explain the decision to nationalise. Similarly, if they did gain power, but did

¹¹⁸⁸ Law of 20 December 1879. Cf. Ziegler (1996), p. 211-229.

¹¹⁸⁹ Cf. Veenendaal (2004), p. 232-234.

not desire power for its own sake, the only way the desire for power can act as an explanatory factor is if these actors used their newfound power to get something they did desire.¹¹⁹⁰

Consider, for example, the Prussian policy of gradual nationalisation of the 1870s and 80s, which included the aforementioned decision of the Prussian government in 1879 to take over the Köln & Minden railway. *Cui bono?*¹¹⁹¹ Absent any method for measuring power, the answer must necessarily be somewhat imprecise. However, it seems reasonable to say that the main gain lies not with the top-level political players, but with the railway ministry. For powerful actors such as the Emperor or the Chancellor, the added power of having direct control over the railways would barely register. Instead, the main beneficiaries would be the key players at the next highest level of government: the Prussian Ministry for Trade, Industry and Public Works.¹¹⁹²

The problem is that there is no evidence that these officials were the ones pushing for nationalisation. Rather, the evidence suggests that the main proponent was the Chancellor Otto von Bismarck, who saw it as a step towards an Imperial German railway system, after his first attempt to set up such a system in 1874 failed because of the resistance of the other German states. For the British nationalisation of 1945, the “power-grab” hypothesis is similarly unconvincing, while for the others it is at the very least implausible; where there is no clear initiator of the reform, and no one gains significantly in power, the desire for more power is at best only a partial explanation.

As a final example of the difficulties involved in using Public Choice Theory to explain the decision to nationalise, we can take the Dutch case. As we have seen in section 3.3.3, the Dutch had maintained their intermediary approach – halfway between competition and collaboration – until the high point of World War I. Focusing on individual government actors instead of on the government as a whole hardly adds any clarity: Who exactly are we meant to focus on and what did they supposedly gain by advocating nationalisation?

¹¹⁹⁰ Cf. what was said previously about the assumption that monarchs such as the Kings of France and Prussia would behave as “stationary bandits”.

¹¹⁹¹ Cf. M.T. Cicero, *Pro Roscio Amerino*, ch. XXX, s. 84, “*L. Cassius ille quem populus Romanus verissimum et sapientissimum iudicem putabat identidem in causis quaerere solebat 'cui bono' fuisset.*” Similarly in *Pro Milone*, 32.3.

¹¹⁹² By Act of 7 August 1878, the department of Public Works became a separate ministry, responsible for the railway system and other public works.

Regarding the first question, it is clear that nationalisation was urged by many throughout the second half of the 19th century, both inside the government and outside, including most notably seven members of the commission of inquiry of 1890. As such, the ultimate decision was somewhat of a collective effort, where it is impossible to pinpoint the pivotal player, players or class of players. It is not even possible to say that it was the politicians, civil servants or railway interests as a class who were pivotal in moving nationalisation forwards. Absent such a pivotal role, any self-interest that we might be able to identify for a person, group of persons or a class is of itself insufficient to explain the choice to nationalise, because the resulting game would be like a prisoners' dilemma: actors would be tempted to free ride off the effort of others as a way of obtaining the benefits of nationalisation – whatever they may be for a particular actor – without incurring the costs of promoting this reform.¹¹⁹³ Of course, the temptation to free ride is all the greater if the benefit is not that great to begin with.

Just like the absence of any noticeable ability to effect change is a problem for PCT, similarly the absence of any willingness to do so is problematic. As we have seen, there are many conceivable objectives for the various categories of actors involved, including money, power, slack, and re-election. But what does the model say about decisions where none of these objectives are at stake to any significant degree? Then we will have to conclude that most actors will favour the status quo, which is a result that we will examine further in the next section. Alternatively, we could assume that in the absence of any immediate self-interest, the government actors will behave like Weberian bureaucrats, impartially executing their assigned tasks.¹¹⁹⁴ Of course, this assumption is the one underlying the analysis of chapter 3. In the specific case of the nationalisation of the Dutch railways, both assumptions have some merit: the reform was exceedingly slow, and it involved a significant measure of impartial policy and policy analysis, in fact more so than in any of the other states under consideration.

All of this forces us to conclude that PCT is not a very useful theoretical framework for explaining this part of the history of Dutch railway regulation. In the other countries, the model fares somewhat better, as we have seen, though even then it leaves somewhat to be desired.

¹¹⁹³ This argument is essentially the same as the model examined by Dixit & Olsen (2000), who noted the shortcomings of the Coase Theorem in a setting with voluntary participation. Cf. section 2.2.3, above.

¹¹⁹⁴ Cf. Weber (1920), part III, ch. VI, which contains his famous discussion of bureaucracy.

4.2.4 Privatisation

Compared to the earlier periods, the time of privatisation poses its own problems for Public Choice Theory. Today, both politicians and directors of corporations find it much more difficult than it once was to get away with various dubious practices. While it is still as reasonable as ever to assume that actors behave as they do for self-serving reasons, even if they are government bureaucrats or politicians, they can no longer defend their choices with mere *ipse dixit*.¹¹⁹⁵ Some plausible-sounding defence needs to be possible for a reform to go ahead, unless of course one is willing to gamble that one will not be discovered. That said, it is important to remember that this defence need do no more than to convince one's key constituents. For a bureaucrat, that means his superiors, and for a politician this means her political base, i.e. the 50% + 1 of voters most likely to vote for her.

When the former US Senator Ted Stevens was attacked for lobbying for a “bridge to nowhere”¹¹⁹⁶ in 2005, the only people he had to convince were his voters in Alaska. Since the Federal government was going to shoulder the bulk of the \$ 398 million cost of the bridge, the Senator actually had very little convincing to do.¹¹⁹⁷ Only when Stevens' Republican Party tried to convince the American people to elect Senator McCain President in 2008 did the project become a problem, not in the least because McCain's running mate – Alaska governor Sarah Palin – had actively campaigned for the bridge during her campaign for governor in 2006. While there is no question that the project would have been a waste of money, this only became a political problem for Senator Stevens and Governor Palin when their political constituency changed.

A lot of the time, the solution for this problem is to construct a political discourse that obscures the transition from the “is” of economics to the “ought” of politics. Of course, nothing could have helped Senator Stevens' bridge – that was nothing short of legalised corruption, buying votes with Federal money – but in many other cases the redistributive consequences of a proposed reform can be swept neatly under the rug by any politician or lobbyist

¹¹⁹⁵ *Ipse dixit* is lawyer-Latin for “because I say so”.

¹¹⁹⁶ The bridge did not literally go to “nowhere”. It connected the town of Ketchikan, Alaska (Pop. 7.368) with the nearby island of Gravina (Pop. 50), which contains Ketchikan International Airport. The road that was supposed to connect the town with the bridge, however, was built even though the bridge was cancelled, since the \$ 25 million in Federal money had already been earmarked. As a result, it quite literally goes to nowhere, stopping a little short of the shore where the bridge was to begin. Cf. <http://www.propublica.org/article/palin-defends-construction-of-road-to-nowhere-925> (last visited August 26, 2010).

¹¹⁹⁷ At least not when it came to defending his position and protecting his chances for re-election. He did, of course, have to convince Congress and the President to accept the earmark. As it happens, Congress voted to spend the money on Hurricane Katrina relief instead.

worth his salt, for example by discussing ideology as if it was fact, or by using Kaldor-Hicks Efficiency instead of the Pareto principle.¹¹⁹⁸ The difference is of course that under Kaldor-Hicks it only has to be possible for those who gain to compensate those who lose, whereas under Pareto this compensation actually has to occur. In practice, politicians avoid talking about the redistributive consequences of their policy choices by talking about GDP growth a lot.

Applying Public Choice Theory to the problem of railway reform under these conditions suggests that railway executives and private investors will seize on theoretical developments in economics and public administration to lobby for a privatisation of the railways, or – in the case of the former – for any kind of decrease in direct government intervention. Regardless of the overall efficiency effects of such a reform, it is clear that it involves a potential transfer of wealth from the passengers to the railway's owners and directors. If the efficiency gain is large enough, it is possible that all parties will end up with a net gain, in which case the reform would be Pareto efficient. If the gain is smaller or completely non-existent, the passengers' loss due to the increase in the railways' monopoly power¹¹⁹⁹ will exceed their gain due to increased quality, meaning that the passengers' interests will conflict with the interests of the railways and with the goal of maximising efficiency *tout court*. In that case, the reform is possibly Kaldor-Hicks efficient but not Pareto efficient. Assuming that some gain to society nonetheless exists, the Pareto principle – as well as honest politics – would require that some of the railways' gain be returned to the passengers in the form of lower prices, i.e. lower than the monopoly price given the regulatory framework remaining in place.

So what actually happened? In Britain, as we have seen, there was great resistance from within the railways against Mrs. Thatcher's reform plans, and against Mr. Major's privatisation scheme. But who was it exactly that objected? It is only natural that the unions would be unhappy¹²⁰⁰; any reduction in government influence over the railways immediately translates into a reduction in their ability to lobby the government for rents for their members. The same could be argued for railway passengers' lobby groups, although their objections are more plausibly troubling: an efficiency drive of the kind that one would expect following privatisa-

¹¹⁹⁸ Cf. Kaldor (1939) and Hicks (1939).

¹¹⁹⁹ There has to be an increase in monopoly power. Without it, no one would have an incentive to lobby for reform.

¹²⁰⁰ Cf. Murray (2001), *passim*.

tion would normally hurt the workers, while being relatively irrelevant for the passengers.¹²⁰¹ Passengers lobby for more trains and more destinations, something they would still be able to do under the new scheme. It follows that the passengers' concerns can quite plausibly be taken at face value: they feared a decrease in efficiency due to the complex market structure that was being imposed.

But what about the ones who were supposed to benefit? Wolmar (2007) discusses how the railway managers were kept out of the conversation for fear that they might be critical, noting particularly "one meeting with [Transport Secretary] MacGregor at which the heads of the profit centres unanimously warned that the scheme was unworkable".¹²⁰² Of course, just because the scheme is unworkable, that does not mean that there is no money to be made. Given their obvious expertise in railway management, the upper management of British Rail was encouraged to get involved in the bidding process for the various franchises, and three such franchises were actually won by BR managers. Most of the other franchises, as well as the ROSCOs and INFRACOs, went to established players; the former were usually run by bus companies such as Arriva¹²⁰³, while banks took the ROSCOs¹²⁰⁴ and construction companies ran the infrastructure maintenance companies.¹²⁰⁵ This is of course exactly what Public Choice Theory would predict: in a subsidised industry such as the railways, rent-seeking is a company's primary source of income. This was true in France in the 1850s, and it is still true today. Basic theory concerning the market for corporate control¹²⁰⁶ would therefore suggest that the companies fall in the hands of those parties most skilled in persuading the government to increase the subsidy. Given the economies of scale involved in lobbying, that means that we would expect consolidation within the railway industry, and takeovers of railway companies by outside parties from other highly regulated and consolidated industries. The bus, banking and construction industries fit that description to a t.

However, while that suggests an answer to the question of "*cui bono?*", it does not establish that railway privatisation in Britain occurred in response to rent seeking. Absent some kind of "smoking gun", we have no reason to dispute the general wisdom about the British

¹²⁰¹ Layoffs could hurt quality, although it is unlikely that the most important metric of quality – punctuality – would be affected very much. On the other hand, some of the cost savings could be passed on to passengers, making the net effect of layoffs on passenger satisfaction more or less neutral, depending on the railway company's freedom of manoeuvre.

¹²⁰² Wolmar (2007), p. 304.

¹²⁰³ Cf. Murray (2007), p. 25-30.

¹²⁰⁴ Cf. Murray (2007), p. 32-34.

¹²⁰⁵ Cf. Murray (2007), p. 35-36.

¹²⁰⁶ Cf. Manne (1965).

railway privatisation: that John Major needed one more thing to privatise, because that was what Tories were supposed to do.¹²⁰⁷ Mrs. Thatcher had already privatised everything else, so the Major government was forced by the political dynamic of Tory party politics to privatise the one major public corporation that was left: British Rail. There is no reason to doubt that Mr. Major sincerely believed that privatisation would be beneficial, but it is also true that it is a programme that came with a distinct personal benefit in the form of enhanced influence over his party.

At the European level, Public Choice Theory is even more difficult to apply, due to the level of complexity of the EU decision making process; even less than at the national level is it possible to identify a pivotal player or group of players. The only way to explain the passage of any piece of legislation through the lens of PCT is to examine the personal incentives of each and every veto player. The EU has a lot of veto players. Initially, there is little difference with the national legislative procedures: a proposal get formulated within a Directorate-General and passed up the hierarchy until it reaches the *chefs de cabinet* and finally the Commission itself.¹²⁰⁸ Up to that stage, the proposal has encountered about as many hurdles as it would have at the national level, although in most countries there is less friction between the bureaucracy and the political players than exists within the European Commission.¹²⁰⁹ However, once the proposal is officially made, it has to run a veritable gauntlet of independently minded Member State representatives and MEPs, where focused opposition of just a few players is often enough to kill it in its infancy. Each of these players has a unique set of political incentives, ranging from the participants in Council working groups whose position is largely bureaucratic, to the lobbying Walhalla that is the European Parliament, to the high-publicity political minefield of the European Council summits.

This complexity means that it is not immediately clear which frame is most appropriate for any given European legislator. Of course, the well-known democratic deficit¹²¹⁰ means that a simple politician-frame will rarely be appropriate. In fact, PCT would suggest that in the absence of adequate democratic controls, politicians would tend to respond more to their

¹²⁰⁷ Cf. Murray (2007), chapter 1 and Wolmar (2007), p. 298.

¹²⁰⁸ Cf. Nugent (2006), ch. 9.

¹²⁰⁹ Cf. Eppink (2007). The main explanation for this phenomenon is that there are more Commissioners than Directorates-General, so that there is no unequivocal line of authority from an individual Commissioner to the DG or DGs that support him. This means more freedom of manoeuvre for the Directors-General and more distrust between the Directorats-General on the one side and the Commissioners and their *cabinets* on the other side.

¹²¹⁰ Cf. Hix (2008).

other incentives, such as the desire to placate lobbyists, the desire for slack and the desire to perform their duties to the best of their ability. What we actually observe is that politicians are noticeably more positive about privatisation and liberalisation when they are acting at the EU level than they are at home.¹²¹¹ Whether this is due to rent-seeking or honest conviction we cannot know.¹²¹²

To make matters worse, we cannot even rely on the question of *cui bono*. As we have seen in chapter 3, European railway law tends to leave a lot up to the Member States, offering moral support for any privatisation effort, nothing more. While it does require a degree of liberalisation – presumably to the benefit of the railway managers – it does so with a firm eye on consumer interests, limiting whatever slack such managers would otherwise obtain through liberalisation.

As a result, we can conclude no differently than that the European legislative process is too diffuse and multifaceted to lend itself to easy PCT analysis, especially in the absence of any obvious source of rent-seeking. In other policy areas, the picture can be clearer, for instance because the European Parliament may not have a role, or because the stakes of the various actors are clear. In the area of railway regulation, however, the outcome is determined by the complex interplay of the incentives of various power players at the national level, key members of the European Parliament, and various Eurocrats. Who will prevail depends on subtle linkages between issues, as well as on the amount of political capital that each player is willing to invest. In such a setting, it is still possible to look back and explain why the outcome was the way it was, and this is done by political scientists the world over.¹²¹³ However, as long as it is nearly impossible to generalise from any one case study, the exercise is largely one of compiling an historical account, without any hope of creating a usable positive model to predict which decision will come out in the end.¹²¹⁴

¹²¹¹ This is evidenced, for example, by the widespread practice of blaming “Brussels” for policies the speaker actually supported as a member of the Council. Cf. Wallström (2008).

¹²¹² In an optimistic mood, we could hypothesise that voters do not elect politicians they disagree with. However, Public Choice Theory is not generally so optimistic about the honesty of politicians and the ability of voters to discern politicians’ true convictions. Cf. Downs (1957).

¹²¹³ Notably by Moravcsik (1998).

¹²¹⁴ The model of Moravcsik (1998), for example, is certainly valuable in that it provides a framework for analysis, but it has far too many “variables” to be a useful predictive model. Essentially the only way that you can gather all the information you need to specify the exogenous variables is if you already know the outcome.

4.2.5 Conclusion

Looking back at the analysis of the past few sections, it is clear that Public Choice Theory is seriously flawed when it comes to predicting the outcomes discussed in chapter 3. For a large number of critical actors it is unclear which frame should be deployed in discussing their actions. Of course we can have the tail wag the dog by deploying whatever frame works best, but that is no way to use a predictive model, since it is an approach that only works if we have access to 20/20 hindsight. This problem of frames is less severe for contemporary problems, which more closely resemble the kind of situation the creators of PCT had in mind when they did their original work, but even then there is still the problem of numbers. As we have seen, PCT works fine as long as there are only a few pivotal actors, or groups of actors or even classes of actors. However, in many real life situations opening up the black box of government in the manner proposed by PCT means considering the political games played by dozens of actors, each with their own incentives and constraints. While it is certainly possible to survey each of the relevant actors, this task is extremely labour-intensive.¹²¹⁵ Moreover, the result is often more a work of history couched in the language of PCT than a test of the model, something we have explicitly tried to avoid here. Given the many possible sources of “utility”, it will always be possible to identify the self-interest that each of the relevant actors was serving by acting as he did, even if the self-interest in question turns out to be nothing more than the desire to do a good job. Like utility maximisation itself, PCT in such a form can be neither proved nor falsified.¹²¹⁶

This is true even of otherwise excellent economic papers such as Nash (2007), who argues that, in “the case of responses to traffic congestion, public choice provides a greater explanation for the emergence of property rights than does economic efficiency.”¹²¹⁷ This is a conclusion that is stunning in its redundancy, since it is quite obviously true that public choice theory as a whole will always provide “a greater explanation” than any one of its ingredients. Even if we assume that the author did not intend for the utility derived from doing a good job to be included in the incentives analysed by Public Choice Theory, it is still not a very remarkable conclusion. After all, it quite comfortably fits with our intuition about politicians’ tendency – or lack thereof – to introduce congestion charges. However, it is also a fact that congestion charges, like railway reforms, do occasionally happen, and regardless of whether

¹²¹⁵ Even Moravcsik (1998) limited himself to studying Germany, France and the UK.

¹²¹⁶ Cf. Boland (2002), p. 15-16.

¹²¹⁷ Nash (2007), p. 1.

one insists on a strict separation between PCT and economic efficiency or not, it would be nice if together these approaches could indicate when such a reform might happen, and ideally without requiring massive amounts of information about dozens of private actors, politicians and civil servants.

For a congestion charge, this is actually quite viable: Given that such congestion is a city issue, we can imagine a candidate for Mayor, let's call him Livingstone, predicting that the time is right for such a charge, which is why he makes it a key campaign issue.¹²¹⁸ As it turns out, a large number of voters in the city have concluded that such a charge would be an improvement for them, so the candidate gets elected, based on this as well as other promises. Because his office offers him sufficient power and influence to push such a proposal through, and because he would not like to face voters during the next election after having broken such a key campaign promise, the Mayor sets about creating a congestion charge, spending whatever political capital he needs to in order to get this done. Viewed this way, it is clear that the enactment of a congestion charge can be predicted by starting from the question of whether it would be efficient to do so, followed by a series of adjustments to take into account the relative influence of winners and losers¹²¹⁹ and the manner in which this issue is likely to become linked to other issues to form a single political platform¹²²⁰, to end up with the likelihood that voters will vote for a candidate promising to introduce such a charge, and the likelihood that voters will care if this promise is kept. However, for a policy field that is as quintessentially national and obscure as railway reform, the picture will always be much messier, with more actors, more different assessments of the situation¹²²¹, and more money on the table. In such circumstances, the result is either a beautiful historical narrative or a simplistic piece of economic theory that offers few testable predictions and even fewer predictions for future developments in the law.

¹²¹⁸ This could be based on the assessment that such a promise is likely to get the candidate elected, which would correspond to the pure and cynical version of PCT, on the assessment that the city would benefit from a congestion charge, or on a combination of the two.

¹²¹⁹ For example, not all the winners and losers may live in the city. Many will work there but live elsewhere, making them ineligible to vote for Mayor.

¹²²⁰ Even though many forms of congestion charging are market-based, the idea has traditionally been rejected by the political right. As a result, it has become linked with other pro-environment policies more than with other transport policies. It follows that the introduction of a congestion charge becomes more likely if the city suffers from increased smog, whereas increased traffic jams will simply lead to calls for more asphalt.

¹²²¹ Note that Nash (2007), in contrasting PCT with efficiency, relies heavily on voters' inability to accurately ascertain their true interests.

With that in mind, we can move on to the real question that this section was meant to tackle: What can we learn from comparing NIE with PCT? Which strengths and weaknesses of NIE have become apparent here?

As so often, the answer is that the strengths and weaknesses of NIE are flip sides of the same coin. NIE as we have examined it in chapter 2 is clear-cut and analytically tractable in a way that PCT can never hope to be. In every NIE model it only takes a few variables to define the optimum. For PCT, on the other hand, the relevant variables are not at the level of society, but at the level of individuals or groups, meaning that the number of variables will almost always be higher and the analysis correspondingly more complex.

The flip side of this is that NIE tends to overlook the role of the individual. We encountered many examples of that in chapter 3, from the Prussian king Friedrich Wilhelm IV in 1840s Prussia to John Major in 1990s Britain. Obviously, there is a place for such factors, but not a very useful one: it all gets hidden away in the all-knowing “objective function” of the government. Much as we are able to be more specific about this objective function if and when we need to be, it would be nice if we could do this in a more systematic manner, by extending the NIE approach to include this problem.

In section 2.4.2, above, we already saw a hint of how this might work: One way to summarise NIE is to say that a firm is a “nexus of contracts”¹²²², something that must also be true for firms that act as parties to a contract at arm’s length. This suggests that NIE would be more successful at understanding such a contract if it also considered the transactions going on at either side. If our focus is on the transaction between the government and a railway company, the analysis should nevertheless involve the transactions taking place within the government and within the railway company as well. A very evident observation is, for example, that the board of the railway company will be much less worried about making transaction specific investments if they do not have to pay for those investments out of their own pockets. Likewise, the government’s behaviour makes a lot more sense if we consider that the contract between democratically elected politicians and their voters is highly incomplete.

¹²²² Cf. Cheung (1983), Gindis (2009) and Hendrikse (2003), par. 5.5.

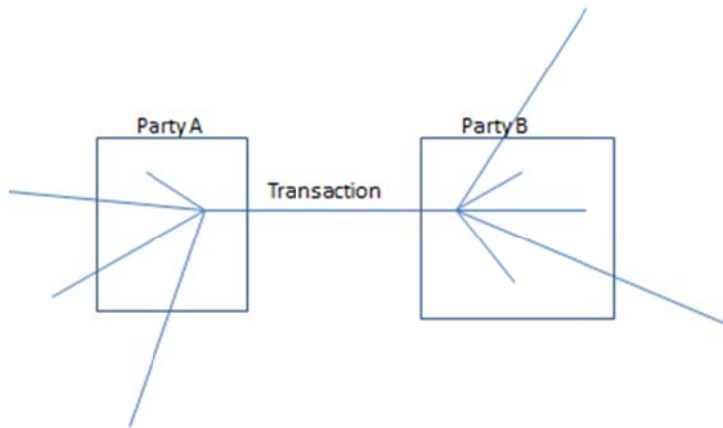


Figure 9: A nexus of contracts

In this way, something of a meeting of the minds might be possible between NIE and Public Choice/Corporate Governance theory. In each case, though, there will be an almost one-to-one trade-off between the realism of the analysis and its tractability.

4.3 Path Dependence

Having considered Public Choice Theory, we will now move on to the other alternative approach for explaining and predicting developments in railway regulation: Path Dependence. This approach, which deals with such concepts as Status Quo Bias, Institutional Lock-In and Increasing Returns¹²²³, emphasises not the relative efficiency of alternative modes of regulation, but rather the costs involved in going from the status quo to any alternative. The result is a simple proposition: History Matters, especially when it involves feedback mechanisms, meaning that choices made today affect the options available tomorrow.

The simplest story¹²²⁴ that illustrates the concept is the Polya Urn.¹²²⁵ George Polya was a Hungarian mathematician, who proposed the following problem: Imagine an urn with r red and g green balls. You draw a ball at random, put it back, and add c additional balls of the same colour.¹²²⁶ How many red and green balls will you have after n iterations? The answer is, of course, that the proportion of each colour after n iterations is distributed in a beta distribution around the initial proportions, regardless of the size of c . However, this is also true if we ask the same question a few iterations later. At any given time, the best predictor for the “final” proportions of red and green is the current distribution of red and green.¹²²⁷ This also means that the first few draws are disproportionately important; after starting with one ball of each colour, drawing red the first five times pretty much guarantees that red will continue to be in the majority indefinitely.¹²²⁸

It is easy to see how we might go from this story of balls and their colours to a story of people and where they might live, i.e. to a story of agglomeration effects in economic geography. Bleakley & Lin (2010), for example, tell a story about portage: In the 1730s, the American colonials started to explore and develop lands further away from the sea than they had so far been accustomed. Significantly, they moved further inland than the rivers’ last rapids be-

¹²²³ Note that these terms are not strictly synonymous, since they refer to slightly different lines of argument.

¹²²⁴ The use of the word story here is not coincidental; because in path dependence history matters, explanations and predictions will always come in the form of a narrative with an explicit time dimension. This idea was already presaged by the discussion of alignment devices in chapter 2, but only comes to full maturity here.

¹²²⁵ Cf. Pierson (2004), p. 17.

¹²²⁶ For convenience, c is usually set at one.

¹²²⁷ In other words, the system has a bias towards the status quo.

¹²²⁸ The probability of red losing its majority if the starting position consists of 6 reds and 1 green is 1,56%.

fore the ocean. This means that the cotton and other agricultural products that were grown inland had to be transported down river in lighter boats, before being carried past the rapids to bigger, seafaring ships. The practice of carrying goods past the rapids in this way is called portage, and the places where this was done quickly grew into trading towns where the traders met the farmers. What Bleakley & Lin have shown is that even today, long after these rapids ceased to be a problem, the towns that were founded there continue to be important. In the present day, for instance, there is no reason why the capital of Virginia, Richmond, should be where it is. One place is as good as another. But 250 years ago, the city was located at the last set of rapids of the James River before it reached the Atlantic Ocean. That small initial advantage is the only reason why this vibrant city continues to be vibrant more than 200 years after the first locks were built to bypass the rapids, and more than 150 years after the railroads completely solved the problem.¹²²⁹

One final example, and perhaps the most famous illustration of the phenomenon, is the history of the QWERTY keyboard. David (1985) discussed the question why we continue to use QWERTY keyboards instead of the allegedly superior Dvorak arrangement.¹²³⁰ The answer is as simple as it is depressing: keyboard makers use the arrangement that their customers have already learned to use, and keyboard users learn to use the keyboard that is used by keyboard makers. Each side's decision is optimal given the practices of the other, and the only real explanation for any of this is that 100 years ago type writer engineers thought it might be a useful sales gimmick if salesmen could write "type writer" using only the top row of the keyboard.¹²³¹ To put it in Polya Urn terms: Given that the first 60 years of ball drawing involved nothing but QWERTY balls, the introduction of Dvorak balls in the 1930s was never going to be a great success; the market was biased in favour of the status quo.¹²³²

¹²²⁹ Bleakley & Lin (2010) in fact identify a whole string of cities at a line roughly parallel to the Atlantic coast, all with origins in portage. Moreover, they draw similar lines around the Great Lakes, and around the Mississippi river and its tributaries.

¹²³⁰ "Allegedly" because the superiority of the Dvorak arrangement has been disputed, most notably by Liebowitz & Margolis (1990). However, the value of the hypothetical as an expositional tool does not depend on its veracity.

¹²³¹ Cf. David (1985), p. 333. Apparently, an earlier arrangement had the R in a different position. Obviously, the original purpose of the arrangement was to minimise the chance that the typebars would clash. The alleged superiority of the Dvorak system is in the typing speed that it allows, not in the likelihood of clashes.

¹²³² In a blog post dated February 8, 2011, economist Paul Krugman gives another wonderful example of path dependence: the strange habit in economics of drawing supply and demand curves with the independent variable on the wrong (vertical) axis. Cf. <http://krugman.blogs.nytimes.com/2011/02/08/supply-and-demand-and-qwerty>. "[E]veryone has been drawing the diagram this way for more than a century, so that anyone trying to put out a textbook that did it the other way would lose a lot of potential adoptions from professors who don't want to change their notes."

It is important to realise that this is a result achieved only through decisions that were both freely made, and made with perfect rationality. At no point does the narrative involve anyone being forced to do anything, or anyone acting against their personal best interest. Instead, the source of the suboptimality here is similar to the familiar prisoners' dilemma: individually rational yet socially suboptimal. The only difference is that here the game is not simultaneous but sequential.

Some authors, however, have argued that it is meaningful to speak about path dependence in other circumstances as well. Liebowitz & Margolis (1995), for example, distinguish between first, second and third degree path dependence, whereby first degree path dependence corresponds to a situation where the chosen path is optimal, though not necessarily uniquely optimal.¹²³³ In other words, first degree path dependence implies a decision maker with perfect foresight. If that assumption is relaxed, we find ourselves in a situation of second degree path dependence, where the path is inefficient, but in a manner that was not foreseeable when it was chosen. Second degree path dependence is still rational, at least for a given amount of information. Third degree path dependence, on the other hand, involves path dependent inefficiencies that can be fixed today. In the sense of Williamson (1993b), the inefficiency is remediable.¹²³⁴ It should be obvious from the thesis so far that this last category of path dependence is extremely problematic for economic theory, in the sense that rational actors should not allow such a situation to persist, meaning that economic theory would not normally lead one to expect to observe a case of third degree path dependence.¹²³⁵

Following this brief introduction, it is time to focus on the source of these path dependencies. What is it, exactly, that causes certain choices to reverberate through the ages? Why is it that European countries have inherited their choice of which side of the road to ride on from the time of horses and carriages¹²³⁶, but have long since stopped using local time¹²³⁷ and local measures of weight and volume?

¹²³³ Liebowitz & Margolis (1995), p. 206-208. Cf. also Liebowitz & Margolis (2000), p. 985. A related, but not identical, typology is proposed by Roe (1996), p. 646-653.

¹²³⁴ Williamson (1993b), p. 140.

¹²³⁵ Cf. Liebowitz & Margolis (1990).

¹²³⁶ Eg. the Highway Act 1773, 13 Geo. 3, c.78 and section 78 of the Highway Act (1835), 5&6 Will 4, c. 50, not to mention art. 9(1) of the 1949 Geneva Convention on Road Traffic.

¹²³⁷ In no small part due to the railways. Cf. Wolmar (2007), p. 104-105. Cf. also Statutes (Definition of Time) Act 1880, 43 & 44 Vict., c. 9.

The issue here is the simple monetary costs and benefits of switching over at any given time. A situation where there are two (or more) alternatives, it costs an X amount to switch from one to the other, and this sum changes over time, often growing.¹²³⁸ Moving an entire city to a different location costs a lot of money. Retraining everyone with a keyboard to get used to a different arrangement of the keys costs a lot of money. Introducing standardised time or standardised weights and measures is much less costly, and certainly less costly relative to the benefits. This is path dependence framed in terms of the costs associated with deviating from the path, once selected.

However, there is another way to look at it. Many examples of path dependence can also be framed by focusing on the economies of scale involved, the increasing returns associated with making a choice and sticking with it. Once you draw a red ball the first time, this leads to more red balls and more red balls. QWERTY keyboards lead to QWERTY typists who lead to more QWERTY keyboards. This is not a competing approach, but a complementary one; it is a different language that sometimes fits better, and sometimes fits more poorly. Instead of emphasising switching costs, it emphasises economies of scale, including network economies.¹²³⁹

The history of railway technology, of course, also has a famous example of increasing returns path dependency. George Stephenson, for reasons that are still not entirely clear¹²⁴⁰, decided to make his railway 4 ft. 8½ inches wide. A little later, Isambard Kingdom Brunel carefully studied the issue and settled on 7 ft for his Great Western Railway.¹²⁴¹ Many other railways had different gauges still.¹²⁴² The ensuing gauge war was certainly an example of switching costs path dependency: it costs a lot of money to change the gauge of an existing railroad, once built. However, it was also an example of increasing returns path dependency: the more railroads are built with a given gauge, the more it will be efficient to build new railroads in that same gauge, especially if they are

¹²³⁸ On Wikipedia we read, for example, that one of the reasons why the Swedish parliament pushed through the switch to driving on the right side of the road despite significant opposition in society was that they realised that with the introduction of more specialised headlights on cars the switchover was set to become much more expensive in the future. Cf. http://en.wikipedia.org/wiki/Dagen_H (last visited 21 September 2010).

¹²³⁹ The classic economic analysis of this kind of technological lock-in is Arthur (1989, 1994).

¹²⁴⁰ Cf. Hylton (2007), p. 166.

¹²⁴¹ Cf. Hylton (2007), p. 167.

¹²⁴² The narrowest was the Garnkirk & Glasgow, at 4 ft. 6 inches, while none went near Brunel's 7 ft. Cf. Hylton (2007), p. 167. On the continent, most railways tended to use Stephenson's gauge, with notable exceptions including the Dutch NRS, which used Brunel's wide gauge.

meant to connect with the first one. When railways were one-off, there was no problem with different companies using different gauges, but as soon as they started to form a country-wide network, the network economies of scale involved practically assured that they would settle on a standard gauge sooner rather than later.¹²⁴³ Regardless of the objective merits of various gauges, Stephenson's initial advantage was enough to secure him the victory.

Finally, it is important to briefly discuss the connection between sunk costs and path dependency. Sunk costs, i.e. costs that have already been incurred and cannot be recovered¹²⁴⁴, have a tendency of setting the players on a certain path that they will have difficulty leaving down the line. As Woerdman (2002) puts it, once the sunk setup costs have been incurred, "continuation of the status quo is for free."¹²⁴⁵ The question is, though, whether sunk costs actually qualify as a distinct "frame" for path dependence. Given most people's difficulties in dealing with sunk costs, there is plenty of reason to be hesitant.

The problem with sunk costs in everyday life is that people tend to assign them continuing significance even though they are, well, sunk. No matter how many times economists insist that sunk costs do not matter, that decisions should be based only on forward-looking variables, people stubbornly insist on throwing good money after bad.¹²⁴⁶ In other words, when faced with sunk costs, people overestimate the extent to which their options are path dependent. For this reason alone, we should be wary about the combination of sunk costs and path dependence: it plays into our innate biases.¹²⁴⁷

On the other hand, it is undeniable that sunk costs cause at least a first degree path dependence. However, when you come right down to it, this is caused by the fact that any attempt to leave that path means that the setup costs will have to be incurred again. Of course the story of Bleakley & Lin (2010) about early American cities is also one of sunk

¹²⁴³ As early as 1846, the UK Parliament decided that all future railway companies would have to use Stephenson's standard gauge. Railway Regulation (Gauge) Act (1846), 9&10 Vict., c. 57. However, this statute did not prevent the GWR from continuing to build new railways using the broad gauge, which they continued to do until 1863. They only finished converting to standard gauge in 1892. Cf. Hylton (2007), p. 170.

¹²⁴⁴ Definition taken from Besanko, Dranove & Shanley (2000), p. 631.

¹²⁴⁵ Woerdman (2002), p. 71.

¹²⁴⁶ Cf. Arkes & Blumer (1985) and Arkes & Hutzler (2000) for an analysis from a psychological perspective.

¹²⁴⁷ How innate? Weatherhead (1979) discusses the phenomenon in the context of birds defending their (partly finished) nests.

costs: the fact that many of the costs incurred in building a town at the location of a portage were sunk is exactly equivalent to saying that there would have been switching costs involved in moving the town. To the extent that the setup costs of the status quo are less sunk, the switching costs of moving to an alternative are also lower.¹²⁴⁸ On the whole, it is difficult to see what would be gained by recognising sunk costs as a separate frame in addition to switching costs and increasing returns, and we will not do so in the remainder of this thesis.

As for the present research, it is worth noting that these issues of path dependence were explicitly considered in the research design. Part of the purpose of tracking the institutional changes in the railway industry all the way back from its inception was to try, as much as possible, to control for these path dependencies. The assumption was, for example, that institutional lock-in would be less of a factor at the time when railways were first introduced¹²⁴⁹ and that it would be easier to control for path dependencies by following the path from the beginning. We have seen the results of that approach in chapter 3, but in order to be able to evaluate them, we must now see how much explanatory power path dependence has without the transaction costs framework. As with public choice theory, the essential thing to keep an eye on while we do so is the operationalisation of the key variables. Just like the public choice approach to railway history occasionally involved defining self-interest so broadly that it became almost meaningless, a path dependence approach might be vulnerable in its operationalisation of switching costs in particular. As soon as we extend switching costs beyond monetary outlays, towards effort, discomfort, and such like, we run the risk of ending up with a theory of everything again. If anything, no matter how ephemeral, qualifies as a switching cost, we are back to writing history instead of economics.

¹²⁴⁸ N.B. the correspondence is not one-on-one. Cf. Woerdman (2002), p. 73-74 for a more detailed analysis.

¹²⁴⁹ Cf. Holterman (2009).

4.4 Path Dependence and Railways

Unlike Public Choice Theory, applying path dependence to railway history does not require any great amount of new facts. Path dependence and NIE are applied at the same level, treating organisations as a whole instead of peaking inside the black box. For this reason, the present section can be briefer than section 4.2, above. Instead of discussing new facts, we will mainly be highlighting things already mentioned, often in the passing, in chapter 3.

4.4.1 *Beginnings*

In the beginning of chapter 3 we saw that while the technology of the railways may have been new, in most cases the institutional choices were not. But does that mean that they were path dependent in any meaningful sense? If so, what exactly was the nature of the relevant economies of scale or switching costs? There is some indication that there were path dependencies at play. In every country except Belgium, the old institutional setup – or lack thereof – of the canals and turnpikes was applied as a matter of course to the new railways. Belgium was the exception, of course, because it was an entirely new country that came into being just at the time when the railways were first created. But why should that matter?

We already know the answer for some countries. In the UK and the Netherlands, there were intense political objections against government involvement in any kind of private enterprise. In the UK, government involvement was not even considered, whereas in the Netherlands its support was limited to the King. This explains why both the old and the new infrastructures had the same institutional structure: they shared a common ideological rationale. In statistical terms: they correlated not because the one caused the other, but because they both had a common cause.

That leaves France and Germany. In Germany, we see a whole range of outcomes, varying from Prussia's benign neglect to the enthusiasm in Saxony and Brunswick. None of the German states seem to have been "captured" in the choices they made in the past. In France, it was only natural that the government would seek to avoid the cost of setting up a new agency to deal with the railways when a perfectly good agency existed in the

form of the *Corps des ingénieurs des ponts et chaussées*. However, this hardly amounts to a serious example of path dependence, given that the existence of that *corps* does not seem to have influenced the government's decision concerning the degree of state influence in the railways.

A more generous approach to applying path dependence in these circumstances is to make it more explicitly political.¹²⁵⁰ That is to say, by focusing on the costs of individual actors by combining path dependence with Public Choice Theory. For example, we could run a model that links the institutional choices made with the power to make further institutional choices.¹²⁵¹ For example: the existing governance structures for canals and turnpikes in Britain made a number of people very rich and powerful, which in turn allowed them to influence the governance structures chosen for the railways. However, this approach has several problems.

First of all, it involves incorporating the difficult notion of “power”, which already proved so difficult to model earlier in this chapter. If we cannot model power, we cannot model how institutions cause some persons to have power but not others, nor how power can be used to affect institutional choices. Secondly, it is difficult to see what this approach would add to a simple Public Choice Theory model, which would simply take the *ex ante* division of power as an exogenous variable. Admittedly, this extension of PCT adds something to the descriptive value of the model, but at the price of using two empirically dubious causal connections instead of one. This empirical vulnerability itself is the third objection; there is little to no evidence that those actors who derived power from the old institutional arrangements were able to use that power for anything other than short term rent extraction. Even something as simple as the rule for level crossings, which initially favoured the roads over the railroads, was quickly altered to favour the railways.¹²⁵²

¹²⁵⁰ Eg. Woerdman (2002).

¹²⁵¹ Cf. Pierson (2000), p. 259.

¹²⁵² Railway Regulation Act (1844), 7&8 Vict., c. 85.

More generously still we could talk about path dependence and policy learning.¹²⁵³ Learning is clearly a path dependent process, whereby the probability that a given lesson will be learned depends on the lessons that have already been learned in the past. The benefit of using policy learning as an operationalisation of path dependency is that it fits well with the impression one would get from a casual consideration of the case: governments had to learn to deal with this new problem.¹²⁵⁴ However, the problem with this approach is that it fit poorly with our explicit starting point: that path dependency should involve individuals acting rationally. Unless we can model the way people gather information, bringing in policy learning simply introduces too much indeterminacy. Moreover, to some extent policy learning is already incorporated in the model of chapter 2; all three of the alignment mechanisms that were discussed there involve the possibility of learning in some form.

It follows from the foregoing that path dependence was not a major factor in the early days of the railways, just as we hypothesised.¹²⁵⁵ What remains is to see if institutions, once created, are as tenacious as is often claimed.

4.4.2 Consolidation

As was to be expected, path dependence became much more of a factor in the development of the railway sector once the initial institutional arrangements were in place. A company, and to a lesser extent a government agency, has the same fear of dying as a human being; despite being only legal persons, they cling to their immortality. Long after the wild inlands of Ontario and Manitoba ceased being quite so wild, the Hudson's Bay Company is still with us, though now in the incarnation of a department store. A famous example among international organisations is the Bank for International Settlements, originally set up in 1930 to facilitate Germany's World War I reparations payments, which is still with us in its modern day incarnation of a forum for banking regulators. The question is whether this phenomenon was important in the railway history as well, specifically in the period of consolidation. Also, we have to examine whether this quali-

¹²⁵³ Cf. Pierson (2000), p. 259-262.

¹²⁵⁴ Cf. Holterman (2009), p. 118-119.

¹²⁵⁵ Cf. Holterman (2009), p. 102 and chapter 1, *ante*.

fies as path dependence the way we have defined it earlier. Where are the switching costs? What are the increasing returns? The positive feedback mechanisms?

Well, to begin with that seems to depend on the direction of the change we are observing. In Belgium, it was trivially easy to get the government out of the railway business. All they had to do was stop building new state railways and start giving concessions to private companies. In government, a switch from action to inaction will almost always be easier than the other way around. After all, action requires consensus between all the veto players, whereas inaction requires no consensus of any kind. In the case of Belgium, the official policy had always been to limit the state railways to the core network, so no the switch to private rail upset no one's legitimate expectations. Also, this switch came early enough in the history of the railways that there was not yet a constituency within the state bureaucracy whose value consisted only of their ability to construct and run a state railway system. There was no mirror image of the situation of the early railways in many other countries, where the organisational capacity did not exist to allow the state to take on the task of railway construction itself. In 1850s Belgium, the available expertise allowed a range of possible alternatives, none of which having deteriorated as a result of the path initially chosen. In other words, the switching costs involved in this move were virtually non-existent, and there was no lock-in due to increasing returns or sunk costs either.

The other countries under consideration all moved towards greater government influence during the period of consolidation. The way this happened was often highly path dependent. Countries that had had state ownership or other state financial involvement in the sector tended to get more state ownership. A notable example is France, which initially got involved in the railway sector at the lowest ebb of path dependence: when the railway companies went bankrupt. It should be obvious that bankruptcy is often a good way to overcome path dependencies; all investments in physical and human capital have become valueless overnight and everyone has a renewed motivation to consider all available alternatives. In these circumstances, very few people ever resist the opportunity of a government bailout, and certainly not on general principle.

So for France, the question is why bailouts led to more bailouts, i.e. to continued government involvement. Obviously, this is not always the case. Looking, for example, at

the American TARP programme of 2008¹²⁵⁶, we see that many of the recipients of funds got out of the programme as quickly as possible. The main companies that are likely to remain in the programme for any length of time are the car companies General Motors and Chrysler, which were on pretty poor footing even before the crisis hit.¹²⁵⁷ Needless to say, the railway companies of 1850s France were more like General Motors than like Goldman Sachs, in that their business models were wholly unrealistic and in urgent need of drastic reform. It is interesting to see that the French government tried to do exactly that. Like the current Obama administration, it tried to get the companies to add a touch of realism to their projections of cash inflows and outflows. The question is what should be done if a realistic examination of the business suggests that it is not possible to run it in a profitable manner.

The answer to that question turns on a notion that we have so far been careful to avoid: the public interest.¹²⁵⁸ Van Genugten (2008) gives an “economic” definition of public interests: “Public interests are societal interests that the market cannot safeguard [due to their] complex external effects”.¹²⁵⁹ While this is certainly an improvement compared to the “political definition” that she also discusses – which essentially states that a public interest is whatever the political system says it is¹²⁶⁰ – it still leaves open the question of which interests are “societal interests”.¹²⁶¹ In line with what we said about social welfare maximisation in earlier chapters, there is no need to resolve that question here. Instead, we can simply reformulate the pertinent question as: Are there market failures that have a significant detrimental effect on social welfare and that the government is

¹²⁵⁶ Troubled Asset Relief Program, cf. title I of the Emergency Economic Stabilization Act, Public Law 110-343, 122 Stat. 3765, enacted October 3, 2008.

¹²⁵⁷ At the time of writing, the following companies are still being supported by the original TARP: GM, AIG, GM’s old in-house automotive financing company Ally Financial, Chrysler and the Regions Financial Corporation. Of course, there are countless other companies still receiving support from the government under various other acts, such as other parts the Emergency Economic Stabilization Act or the American Recovery and Reinvestment Act, Public Law 111-5, i.e. the Stimulus Act.

¹²⁵⁸ Cf. Van Genugten (2008), p. 4-6, Teulings, Bovenberg & Van Dalen (2003) and WRR (2000).

¹²⁵⁹ Van Genugten (2008), p. 5. This definition is her summary of the analysis of Teulings, Bovenberg & Van Dalen (2003).

¹²⁶⁰ Cf. WRR (2000), p. 19-21.

¹²⁶¹ In the English translation of WRR (2000), the Dutch Scientific Council for Government Policy writes that “interests are *social interests* if their promotion is desired by society as a whole.” (p. 12, emphasis in the original). Clearly, that definition is either tautological or outright incorrect, depending on how one understands the phrase “society as a whole”.

The Dutch original, which appears in WRR (2000), p. 20, is slightly less ethereal: “Interests are social interests if their promotion is desirable for society as a whole.” (“Belangen zijn *maatschappelijke belangen* als hun behartiging voor de samenleving als geheel gewenst is.”) Still, such a definition will always be troubling for an economist, who is not used to distinguishing between the interests of “society as a whole” and the interests of the individuals that make up that society.

likely to be able to (partially) fix? This version of the question at least gives an answer to the Obama administration's problem with the American car industry: Permanent state intervention is not the answer.

In 1850s France, the answer was less obvious. Up to that point, there was no precedent for the kind of large scale bailout that they were considering. There was also no precedent for any kind of serious government involvement in any form of transport. Nevertheless, and probably for good reason, the French government of Napoleon III decided that an important public interest was involved and bailed out the industry. As we concluded in the previous section, that decision as such was not the result of any real path dependency. However, the method that was chosen for the bailout did have the effect of creating path dependencies for the future.

As we have seen in chapter 3, the French government elected to provide a rate of return guarantee, an instrument also used by the Dutch government in later decades. The problem with such a guarantee is that, unlike the instruments of equity participation and fixed annual subsidy, the government's liability is essentially open-ended. There is no way to know in advance how much the guarantee will cost. Every year again, the government is at the mercy of the railway companies. Every year again, the government's expenditure will end up higher than expected, so every year again there will be pressure from parliament or – in the case of Second Empire France – from elsewhere in the political system, for the government to get more involved in the management of the industry. At the same time, there will never be a majority for anything drastic. If such a majority had existed originally, the government would have chosen a more hierarchical instrument, one with more control rights for the state. Given that in the Netherlands and France a rate of return guarantee was as far as a majority was willing to go, that is not suddenly going to change in subsequent years. As a result, you end up with a patchwork of little fixes enacted in successive years, each in response to yet another year of disappointing railway profitability.

So in what sense is this phenomenon path dependent? Firstly, it is difficult to see how there are switching costs involved. The costs of setting up a different instrument for government support were the same before and after the instrument of rate of return guarantee was chosen. There are no sunk costs involved in providing such a guarantee, not even political costs. Secondly, it is equally difficult to see what the increasing returns

might be. There are no economies of scale involved, no clear feedback mechanisms, be they economic or political. We are left to conclude that the path dependence observed is rooted in nothing other than the irrationality of the relevant political actors, who failed to see that they were trapped in a perpetual Groundhog Day of regulation. Calling this second degree path dependence would be extremely generous, since the problems were almost certainly remediable – assuming of course a social welfare maximising government.¹²⁶² A diagnosis of second degree path dependence would only follow if we combine path dependence with the Public Choice Theory of earlier sections; in that case, the political costs of getting all the veto players on board for a different institutional regime would count as switching costs.¹²⁶³

In the UK we have perhaps the quintessential example of a positive political feedback phenomenon, whereby profits made by railway companies translated into more seats in parliament and therefore more power and more profit. Whereas initially the power and money of the turnpike and canal owners was not enough to overcome the economic advantages of the railways, in later decades, the railway interests did manage the money-power-money feedback loop. However, as the century progressed, their ability to keep this scheme going steadily eroded in the face of a series of electoral reforms, extending the franchise and removing opportunities for corruption. When the railways were first introduced, seats in the House of Commons were for sale in any of a number of different ways.¹²⁶⁴ In a series of statutes, starting with the Great Reform Act¹²⁶⁵, these were all removed, paving the way for a genuine one-man-one-vote system.¹²⁶⁶ These reforms made it possible for the Labour party to rise to prominence starting in the 1890s, the party that would end up as the main proponent of nationalisation in the 20th century. Already in earlier decades, the reform of the election laws strengthened the position of the parties relative to those who were previously in a position to sell or give away seats, putting spe-

¹²⁶² Cf. Liebowitz & Margolis (1995) and above for a discussion of the different degrees of path dependence.

¹²⁶³ It would still not be first degree path dependence, because the path chosen is suboptimal from the perspective of a decision maker looking forward at the time of the initial decision.

¹²⁶⁴ Cf. May (1863), chapter VI.

¹²⁶⁵ An Act to amend the representation of the people in England and Wales (1832), 2&3 Wm. IV, c. 45.

¹²⁶⁶ Or as close to one as it is possible to get in a single seat constituency system. Cf. United States Supreme Court in *Reynolds v. Sims* (1964), 377 US 533, 84 S. Ct. 1362; 12 L. Ed. 2d 506. In the UK, obviously, this issue was never litigated, except very recently in the litigation about the Channel island of Sark. *R. (ex parte Barclay) v. Secretary of State for Justice*, [2008] EWCA Civ 1319, [2009] 2 WLR 1205, par. 120, upheld on appeal to the Supreme Court of the UK in *R. (ex parte Barclay) v. Secretary of State for Justice*, [2010] 1 AC 464, [2009] UKSC 9.

cial interests with and without money on more equal footing. In this way, it appears to be possible to use a political path dependence model to explain why the railways were able to obtain *carte blanche* in the middle of the 19th century, but ever less so as time went on.

In Germany, on the other hand, it is difficult to see how such political barriers might arise. This is especially true for Prussia and similarly authoritarian states, where political opposition can easily be swept away in the face of a determined monarch.¹²⁶⁷ Instead, the key players in Prussia and Germany generally preferred a great deal of state involvement in the railways, for military-strategic reasons as well as for reasons of efficiency, and on the whole they steadily worked towards that goal. In fact, in periods and places where this preference was lacking, this was immediately reflected in the governance choices made.¹²⁶⁸ As for other kinds of lock-in, it is clear that German railway companies were always small relative to the state, so that the tail never ended up wagging the dog, as it did in France, Belgium and the UK. This was also the case because it would be very difficult for the railway sector to organise and lobby effectively in an undemocratic state like Prussia, which not only offers relatively little access to stakeholders, but also tends to display a profound disdain of the money-making bourgeoisie by the nobility involved in the higher callings of military and government.

In conclusion, it is clear that the only path dependencies we observe are political in nature. Companies that are large relative to the economy, as in France and the UK, or otherwise influential, as in Belgium, are difficult to ignore once created. In such circumstances, interfering with their operation can be profoundly disruptive for the economy, making them too big to fail. At the same time, their money buys a lot of influence. Bailouts, which are supposed to be clean and one-off, degrade into annual haggling between the companies and the state. Similarly, regulation that affects the incumbent major railway companies in a negative way will miraculously disappear in an ocean of red tape, if it is even allowed to fail in such a graceful manner.¹²⁶⁹

¹²⁶⁷ As the example of 1848 shows, sometimes it can take a few months to marshal the necessary determination. Cf. Rapport (2008). Note that, in describing Prussia as “authoritarian”, one must remember that in actual practice compromise was generally preferred over a too direct exercise of the power of the Crown. Cf. Dunlavy (1994), 21-27.

¹²⁶⁸ Cf. Ziegler (1996), chapter 3.

¹²⁶⁹ Note how Gladstone’s screening committee for new railway proposals in the 1840s was not even allowed the honourable defeat of being allowed to continue to exist without real power. Instead, it was summarily abolished.

The other two countries under examination, the Netherlands and Germany, show the benefits of having less powerful railway companies. In the Netherlands, the railways were never as important for the economy as elsewhere, given the dominance of river transport for goods. For this reason, the government could afford to care little about the profitability of the railway companies or their investments in new track and rolling stock. On the occasions where an overhaul of the industry was contemplated, the main opposition came from the liberal side of parliament – on ideological grounds – rather than from the railway companies themselves.

In Germany, on the other hand, the railway companies were too small to put effective pressure on the government, all the more so because the undemocratic nature of the Prussian and German governments. Being an able politician, however, Count Bismarck generally preferred to avoid confrontation and applied a divide-and-conquer approach, consolidating and nationalising the sector by gradually buying up the equity of the companies, one at a time.

In other words, such a path dependence approach adds precious little to the Public Choice approach of section 4.2.2, above, while a non-political path dependence analysis, i.e. one that does not rely on the relative power of various groups, does not appear to show any signs of actual path dependence. As a result, the only kind of path dependence remaining is the third-degree kind, which we have already ruled inadmissible in section 4.3 above, on the grounds that it violates our rationality assumption and could just as easily be considered a violation of economic theory, generally.

4.4.3 Nationalisation

For nationalisation even more than for consolidation one would hypothesise that path dependence would be a prominent factor. After all, consolidation demands little from the government directly, or from its relationship with the railway companies. Theoretically, government may elect to be relaxed about railway companies consolidating or not, it may encourage this development or not, without ever running into the limits of its ability to oversee the industry or make contracts with the companies. The switching costs of changing the level of consolidation in the industry would normally be dwarfed by the switching

costs involved in nationalisation. As a result, path dependence will cause the decision to nationalise – which is discontinuous and therefore by definition not path dependent – to be postponed, and when this decision is taken, the manner in which it is carried out will bear the traces of past institutional choices.

There is no country in our sample where this pattern does not occur. In Germany, we observe difficulties in wresting control over the railways from the level of the states to the level of the Empire, something that proved difficult even after the former were nearly obliterated in the wake of World War I. The only thing that simplified the negotiations was the fact that the railways had become nearly worthless in many places due to war damages and their role in the scheme of reparations payments.

In Belgium, we could say that nationalisation was only partly the government's choice, and partly something they were forced into by their earlier choice to allow the two main financial conglomerates to control the industry. As a result, they ended up bailing out the banks by buying up the railways, even those that had little military-strategic significance. However, just like the bank-controlled railways had never been run in any kind of unified way, but rather as a collection of individual lines, so too were the nationalised railways. Only in 1912 did the Belgian government consider creating anything like a “national” railway company for the first time, an idea that was greatly encouraged by the destruction of World War I, but only completed in 1926, when the NMBS/SNCB was created. In other words, both in Germany and in Belgium true national consolidation followed the beginning of state ownership by more than half a century. It was consolidation that was delayed, not nationalisation.

In the Netherlands, we have seen how both consolidation and later nationalisation were put off time and again. After the process of consolidation had – rather grudgingly – resulted in two companies, these companies continued to exist until the creation of the NS in 1917 and their actual merger in 1938. During the intervening period, nationalisation occurred gradually, because the state purchased new shares in the two parent companies whenever the railways needed new financial support. Only in 1938 did the Dutch railways system become entirely state-owned and entirely consolidated, three quarters of a century after the Dutch government first started building railways in the 1860s.

In France, the story is similar. The big six, having become through reorganisations and mergers four private and two public companies¹²⁷⁰, continued to exist as independent companies until 1938, when the SNCF was created. However, even then the old companies continued to hold 49% of the shares, and the new company continued to be divided into the same old geographic areas of operation. Only when the convention of 1937 expired in 1982 did the company become fully consolidated and nationalised.

The UK, finally, also based its process of nationalisation on the pre-existing railway giants. Of the so-called Big Four that were created in 1921, only the Southern Railway did not have one dominant pre-existing company.¹²⁷¹ The other three were clearly set up by adding the smaller companies to the Great Western, to the London & North Western and to the Great Eastern, respectively, in order to form the new Great Western, the London, Midland & Scottish and the London & North Eastern. After nationalisation, in 1947, these four companies became the six regions of British Railways, with new regions being created for Scotland and for London and the Midlands. In this incarnation, they continued to exist as operational units until the 1980s, remaining dormant afterwards before being revived as railway franchises in the 1990s. This explains why Isambard Kingdom Brunel's brain child, the Great Western Railway Company, still or again – depending on how one wishes to look at it – exists today.¹²⁷²

Concluding, we can say that in all five countries the decision to nationalise was put off until some calamity made it more politically palatable. In Belgium, the first such calamity was the Franco-Prussian war of 1870 and all five countries one or both of the World Wars fulfilled this function. Also, in all five countries the division of the industry into distinct companies survived the moment of nationalisation by decades. It follows that these institutional choices appear to be path dependent. However, what remains to be discussed is whether they are path dependent in an economic sense, and whether they are

¹²⁷⁰ The former being the Nord, the PLM the PO-Midi and the Est, and the latter were the Ouest-État and the railways of Alsace-Lorraine, which had become property of the French state after World War I. The Compagnie de l'Ouest had been purchased by the state in 1908 and merged with its existing network in western France, and the Compagnie du Midi had been purchased by the Paris-Orléans in 1934. Cf. chapter 3, above, and Caron (2005a, 2005b).

¹²⁷¹ Instead, it consisted of three roughly equal partners: the London, Brighton & South Coast, the London & South Western and the South Eastern & Chatham.

¹²⁷² Of course the franchise currently held by First Great Western Ltd is in many important respects different from the network created by Brunel and his successors, but – as its name would suggest – it still connects London with Bristol and environs, and it still uses largely the same track as in the 19th century, albeit in standard gauge. Compared to that earlier company however, the current franchise is much smaller, since it only covers intercity services, and in a much smaller area.

path dependent in any sense that is distinct from the political considerations and learning effects that were already discussed earlier.¹²⁷³

As before, the increasing returns frame does not seem to be particularly helpful. Institutions are not like railway tracks, where a standard gauge will inevitably emerge. France, for example, first introduced a state railway company in some of the western parts of the country, yet this choice in no way affected the state's options with regards to the rest of the country. On the contrary, it gave the government a convenient instrument for the future, giving it the capacity to nationalise the other major railway companies if and when this should prove necessary, as it did for the *Ouest* in 1934. The fact that part of the network was run with a different governance structure than the rest made matters easier for the government, not more difficult. In such circumstances, it is difficult to see what the economies of scale-induced path dependencies might be.

Switching costs, on the other hand, seems like a more promising approach. Even if it is no problem for different governance structures to exist side by side, it is undoubtedly quite costly to change the governance structure for any given case. Nationalisation, for example, requires myriad decisions to be made about who is going to be in charge of what how, arrangements need to be made with the previous owners, etc. These are all classical transaction costs: they are the costs of pushing paper. The fact that these costs are dwarfed by the overall costs of running a railway network is irrelevant. The only thing that matters is how they compare to the efficiency gains associated with the switch.¹²⁷⁴ And as we have seen in earlier chapters, while it is likely that there are significant efficiency differences between the various possible governance structures, it is exceedingly difficult to calculate their size or even to say which institutional setup should be preferred. In such circumstances, it is not unreasonable for the government to treat various alternatives as being economically equivalent, which, in turn, means that even the smallest amount of switching costs can cause a significant degree of path dependence, at least until the underlying uncertainty about the efficiency of governance structures is resolved.

¹²⁷³ Cf. section 4.2.3 and 2.6, respectively.

¹²⁷⁴ Cf. fig. 2.2 on p. 74 of Woerdman (2002).

It is interesting to consider where this version of path dependence falls on the Liebowitz & Margolis scale of first, second and third-degree path dependence. Clearly a case can be made for any of the three. If we decline to equate rationality with a God-like level of omniscience, we can reasonably say that in many instances the relevant decision-makers accurately analysed the efficiency implications of the various alternatives they were faced with, and chose a path that they considered better than the others or as good as the others, without subsequently being proved wrong.¹²⁷⁵

The case for second-degree path dependence is even more straightforward. Liebowitz & Margolis write: “[If] the inferiority of a chosen path is unknowable at the time a choice is made, but we later recognise that some alternative path would have yielded greater wealth, [we speak of] second-degree path dependence.”¹²⁷⁶ In other words, whether we should speak of first- or second-degree path dependence depends on the extent to which the choices made were subsequently proved wrong.

Third-degree path dependence, finally, implies not only the existence of an alternative that would have been superior if it had been implemented from the start, but one that would still be superior today. That is to say, it involves a remediable inefficiency, like nationalising the railways without actually setting up a workable system for supervising them. In actual fact, most other cases of alleged third-degree path dependence will probably be debatable, depending on which kinds of switching costs one elects to include in the analysis and on the perceived superiority of the alternative.

Those categories of switching costs are still the elephant in the room, as in the previous sections. While it may be possible to explain why nationalisation should be delayed and why it should be done in a manner that maintains some pre-existing governance structures by relying on the near unknowability of the efficiency effects of nationalisation and on the relatively small plain vanilla transaction cost of switching, the historic evidence shows that these were certainly not at the forefront of many decision makers’ minds.

¹²⁷⁵ Remember that it is often difficult to distinguish between subsequent increases in understanding and the contemporaneous shifts in the preferences of many key actors. Cf. section 3.3, above. Needless to say that if a later generation comes to appreciate state control for its own sake, this does not have any bearing on the earlier generation’s decision to keep the railways private.

¹²⁷⁶ Liebowitz & Margolis (1999), p. 985.

In Germany, the argument was about power and money for the Imperial government. In Belgium, the issue was one of patronage, with members of parliament being reluctant to give up their power to exert pressure on the railways for the benefit of their constituents. In the UK and the Netherlands, railway nationalisation was delayed until the debate about the desirability of greater government influence over the economy was resolved. This debate was about preferences much more than it was about efficiency. In France, finally, the government of the Third Republic was chronically unable to tackle any of the major problems it was faced with. Following the Third Empire, the framers of the French constitution of 1871 sought to design a more restrained form of government. To make matters worse, throughout the existence of the Third Republic France suffered from sharp political divisions that made compromise difficult. As a result, the French governments of the time were often paralysed, unable to resolve important problems for years or decades on end.¹²⁷⁷

In summary, while it is certainly possible to explain the pattern of nationalisation using a path dependence framework, the actual historical evidence does not support such an approach. Instead, it suggests a greater role for political considerations, not unlike the ones discussed in section 4.2.3, above.

4.4.4 Privatisation

As with nationalisation, it is natural to distinguish between the act of privatisation and its timing and manner. The act of privatisation is not only a – more or less – discrete event, but it is also a rather drastic break with tradition. Ever since the 1840s, 50s and 60s, the trend in all five countries was towards greater government control over the railways. As late as 1982 France enacted a further increase in state control over the SNCF, turning it into a public company without private shareholders. Of course, path dependence provides us with a framework for talking about such changes in direction, something NIE and PCT do not, but that does not mean that it is a particularly comfortable conversation. The theory suggests that we should expect a break with tradition if and when the inefficiency of the status quo relative to the best alternative exceeds the switching costs involved in

¹²⁷⁷ Cf. for example Lehning (2001).

abandoning the status quo.¹²⁷⁸ As such, path dependence is a model of change as well as of continuity. Whether that is true in practice as well as in theory, however, remains to be seen.

When the railways were nationalised, the institutional choices made in the past had a tremendous influence on the timing and manner of nationalisation in many places. Many railway companies were influential enough to be able to resist being nationalised until such time as they were certain they would benefit from the deal, and even after nationalisation the state railway company would often bear traces of the institutions of earlier times. One might expect the same to happen when the time came for privatisation.

What actually happened was that the inefficiency of the status quo became ever more commonly accepted as the post-war era progressed. Of course, that did not mean that it was established in any kind of formal way; after all, to do that one would have needed knowledge of a counterfactual. Instead, there was simply a mountain of complaints about lacking punctuality, the poor quality of the carriages, poor connections, escalating investment costs, budget overruns, etc. The complaints were generally vague as to what should be done, except of course that the government should get its act together. In most countries, the memory of genuinely private railways was rapidly fading, a circumstance that more likely to add to the unhappiness than to detract from it. After all, bad as the nationalised railways were, their private predecessors had been worse. That is why they had been nationalised in the first place. In other words, the apparent inefficiency may well have been more of a political problem than an economic one.

On the other hand, there were some objective reasons why decision makers could think that privatisation was more viable than it had been before. Starting in the 1960s, the theoretical work on Public Choice Theory and on franchise bidding¹²⁷⁹ began to develop, making it easier for the government to design contract-based governance structures that might actually work. At the same time, the improvements in communications and related technology meant that the information asymmetry problems that are inherent in any attempt to privatise became somewhat less of a factor. That means that today it is feasible for the government to observe directly whether the trains are running on time or not, and

¹²⁷⁸ N.B. both the costs and the benefits of switching have to be calculated in Present Value terms, of course. This makes the discount rate, which reflects the extent to which the relevant decision makers have a short-term or a long-term perspective, an important and often overlooked variable.

¹²⁷⁹ Cf. Demsetz (1968).

to write the concession accordingly. Similarly, it is feasibly today for different companies to share the use of a piece of track in a way that it never was before, because the capacity allocation systems are so much better than they were in the 19th century, not to speak of collision prevention devices such as ATP in the UK and ERTMS in continental Europe.

In this way, we can understand the path taken by the railway industry in the post-War era as one where the efficiency of the status quo did not change very much, but where the efficiency of the alternatives steadily increased. The big question remains, though, whether the economic inefficiency of the status quo was decisive in the end, or the political benefits associated with reform. One would surmise that the answer differs greatly according to time and place. After all, railway liberalisation was hardly an automatic vote-winner in every country under consideration.

When privatisation first appeared on the political agenda in the UK, it is quite obvious that there was little political pressure to proceed very far along these lines. Mrs. Thatcher was not making herself very popular with many of her other privatisation projects, and her decision to keep British Rail in public hands went unchallenged. The changes she did make concerning the manner in which BR operated were widely supported and were by all accounts motivated by a genuine belief in their desirability.

As we have seen, it was John Major's government that proceeded with the first widespread privatisation of the railway industry in Europe. While for some of the decision makers involved, this choice was motivated by an unwavering faith in its correctness, it is difficult to escape the impression that it was also considered important for him to be seen to be privatising something, anything, in order to placate the hardcore Tory electorate. While Mrs. Thatcher's position as the Messiah of Thatcherism afforded her the freedom to decline to privatise certain enterprises depending on her assessment of the economic costs and benefits of privatisation, Mr. Major was faced with an entirely different political calculation. For him, the political inefficiency of maintaining the status quo was much greater than it had been for his predecessor, creating a net benefit to switching where before there had been a net cost. Note that, to the extent that any of this is path

dependent, it would be first degree path dependence: whenever switching is optimal, it occurs.¹²⁸⁰

A similar analysis can be carried out for the European Institutions, and for the other countries under consideration, but it unnecessary to do so since it would tend to repeat what has already been said in section 3.4 and 4.2.4, above. The former recounted the story undiluted by political reality, at least to the extent that this was possible, while the latter gave the politicised version. In the former case, we would have to conclude that the switching costs involved in privatising the railway industry are not negligible, but also not of such magnitude as would significantly alter the calculation of costs and benefits. In the latter case, where the rent-seeking activities of various parties are taken into account as well, suddenly the efficiency gain that is to be expected is reduced, while the switching costs are increased. As a result, privatisation is postponed, possibly indefinitely. Casting this in path dependence language adds little to our understanding of the problem.

Finally, it is interesting to consider the issue of manner. As it happens, the observation earlier in this chapter that privatisation is almost always easier than nationalisation applies to manner as well as to timing. In terms of rent-seeking, nationalisation pits the government against the railway companies, while privatisation pits it against the railway unions. Given that the latter are generally weaker, and also cheaper to buy off¹²⁸¹, it follows that privatisation is easier than nationalisation, meaning that it will be accomplished with less delay, and with less path dependence in “manner”.

This is still true if we recognise the continuity that will exist in many countries, including the countries under consideration in this thesis. In some countries, the franchises that will be offered up for tendering will be identical to, or at least similar to, the regional divisions of the old state railway company. Almost everywhere the old railway

¹²⁸⁰ Or rather, we navigate our way through a minefield of unknown or unknowable variables by assuming that if switching occurred, it must have been optimal, or at least perceived as optimal, for someone somehow. Reasoning backwards in this manner, it follows from the fact that the key switch occurred under Mr. Major instead of under Mrs. Thatcher that the change in leadership changed the relevant political calculations, since it is impossible for the underlying economics to change this quickly.

¹²⁸¹ Privatisation will make the railway unions worry about the protected status their members would have enjoyed as civil servants. This problem can be resolved by extending some or all of these protections post-privatisation, something that will undermine the efficiency gains associated with privatisation, but will allow the project to go ahead.

executives will continue to play a leading role after the state's influence is reduced, either because they end up organising leveraged buyouts, as in the UK, or because they will be asked to continue in their old jobs. After all, running a railway requires extensive expertise that only few possess, making these executives a precious commodity regardless of the complaints about the way the old state railways operated. This is, however, a path dependence of details, not one of the main institutional setup.

All things considered, path dependence does not seem to fare particularly well in the age of privatisation. The timing of privatisation was too political to be feasibly modelled with an economic path dependence model, and the manner in which the railways were privatised was only path dependent in the details. In fact, one could sooner argue that the manner in which the railways were privatised was often not path dependent enough, in that it destroyed a lot of human capital by severing ties that perhaps ought not to have been severed, placing people who by all rights should be working together closely in different companies. Many a delayed train has been blamed on the severed ties between the infrastructure maintenance function, the infrastructure allocation function and the rail transport function, for example.

4.4.5 Conclusion

All things considered, a few things have become very clear about this path dependent history of the railways. First of all, and most importantly, this section has once again demonstrated the importance of thinking not only about the optimum, but also about how we are supposed to get there, the dynamics of the problem. This topic was already considered in section 2.6, and if the present discussion of path dependence has any value, it is as a different approach to alignment mechanisms.

Before, the road towards the optimum was implicitly imagined as a more or less continuous one, albeit aimed at a goal that was forever changing. Path dependence theory however, like Williamson (1991), imagines the problem as one of transition between the status quo and an alternative path that is arguably more efficient, at least in some sense of

the word. As a result, the question becomes why the allegedly superior alternative is not chosen, the most common answers being that the status quo displays too many economies of scale or that the transition away from the status quo would be too costly for other reasons.

As we have seen, the application of these ideas to institutions is far from obvious. Institutions are not typewriters, they are not typically associated with an immediately identifiable investment in physical assets. Instead, they affect people's investments in more indirect ways, as discussed in chapter 2. If a body of institutions performs this task suboptimally, it is not the institutions that are potentially path dependent, but the behaviour that the institutions seek to govern. Institutions guided the collaboration between the Big Six French railway companies of the 19th century and the government, resulting in the creation of a railway network that bore the traces of that governance structure. However, while the investments in real assets were caused in part by the contemporary institutional setup, the reverse is not true. Dividing the industry between six private companies caused them to build six recognisable networks, but this fact would not materially affect the choice of future governance structures; the number of companies operating on a given network can be increased or reduced, new networks can be added, the whole industry can be nationalised or consolidated into a single company, etc., without the original "Big Six"-approach playing any kind of special role in the consideration.

Given that institutions are not path dependent because of the investments they inspire, the only way that a path dependence framework can work here is if we identify one or more categories of costs that are necessitated by the creation or maintenance of an institution, costs that can create positive feedback loops directly. Unfortunately, while such costs do exist, they are so small that they tend to be dwarfed by the potential gains from abandoning the status quo. On the other hand, we could improve the explanatory power of the path dependence framework by extending the definition of costs and benefits to include political and other non-monetary costs and gains. However, that would be outside the scope of this thesis.

That leaves the question of whether it is nonetheless possible to improve our analysis of alignment mechanisms by incorporating some of the ideas of the current section. The answer is that, of course, to the extent that institutions involve significant economies of

scale or switching costs, a degree of path dependence can emerge. Unlike Public Choice Theory, the Path Dependence framework can be combined with NIE quite easily, when it is advantageous to do so, since the latter two have a similar perspective when it comes to the relevant variables and actors.

Beyond that, path dependence represents a useful concept when it comes to creating real connections with adjacent academic disciplines. Economic science may not be able to say very much about the timing issues discussed in this section, but other disciplines can. Having a section on path dependence and what I have called alignment mechanisms is a useful way of creating room for those disciplines. Even if the economist does not intent to fill that room himself, it is still useful to remind everyone that it is there, and that the outcomes predicted by economic science will not normally appear instantaneously.

4.5 Conclusion: Alternative Approaches

As we have seen in the previous sections, the two approaches discussed here, Public Choice Theory and Path Dependence, each have their strengths and weaknesses, varying between times and places. Moreover, on many occasions it seemed tempting to combine them into something of a synthesis, along the lines of Woerdman (2002), who applied a Path Dependence model of political barriers and institutional lock-in. However, at the moments in this chapter where this subject has come up, we have always concluded, for a variety of reasons, that this solution is less attractive than it appears at first sight. In this concluding section, it is appropriate to revisit our earlier conclusions in a more systematic manner, stating once and for all what the strengths and weaknesses are of PCT, of Path Dependence, and of the combination of the two.

It should be clear from section 4.2, above, that Public Choice-style models work best when there are relatively few actors or categories of actors involved, and when these actors each have clear incentives. The mirror image of those clear incentives is, of course, the absence of a clear social welfare maximising optimum that could get in the way of various people's rent-seeking efforts. Especially when the actors who have the power to take the ultimate decision do not themselves have a stake in the outcome of the decision making process, but are merely the relatively passive recipients of rent-seeking efforts, the social welfare maximising choice tends to carry significant weight, but only when it is sufficiently clear.

Of the four periods that we have distinguished, the period of consolidation fits this description the best. During that period, the relevant institutional choices were not necessarily system-wide, but were more often limited to only a part of the network, a specific connection or region, where consolidation was to be encouraged, discouraged, or allowed to happen as it would. During this period, in the second half of the 19th century, the railway companies were at the pinnacle of their financial success, and there were very few other actors who could – or would – stand in their way. This makes the Public Choice version of that history fairly straightforward: railway companies generally got what they

wanted – market power and government support – except to the extent that they were thwarted by other railway companies: an oligopsony of rents.

PCT also explains why the importance of rent-seeking declined in the subsequent period of nationalisation: because the railway industry came under increasing competition from other modes of transport, it was less able to hold the economy hostage than before, while its financial fortunes declined making it less able to finance its rent-seeking activities. All the while, the steady extension of the franchise in all countries under consideration brought a new kind of legislator to the fray, without any kind of predisposition against government intervention and often even explicitly socialist. Gone were the times when the railway companies could count on the potential sympathy of almost everyone in power.

Convenient as it may be for a theory to know its own limitations, it does leave a rather inconvenient question: what are we supposed to do when the outcome is only weakly influenced by rent-seeking? The answer is that we will obviously have to have another theory ready to fill in the gaps.

One such alternative is path dependence. After all, it is not unreasonable to speculate that path dependence will work particularly well in periods when the political stakes are low, when Public Choice Theory is at its weakest. However, for that to work, institutions would have to be in some sense similar to typewriters.¹²⁸²

As we have seen, this institutions-as-typewriters metaphor is rather unwieldy, at least if we insist on restricting it to economically cognisable variables. It is usually true that the financial costs of setting up a given governance structure, or the cost of switching from one structure to another, are dwarfed by the efficiency implications that are predicted for such a move. If the monetary outlay associated with setting up a system of privatised railway companies is several million euros or pounds worth of legislators', civil servants' and private actors' time and effort, as well as an office building or two, while the predicted efficiency gains run in the tens and hundreds of millions per year, how can the governance structures of the railway industry ever be path dependent?¹²⁸³ Clearly

¹²⁸² Cf. Woerdman (2002), p. 75.

¹²⁸³ Arguably, nationalisation and consolidation would be even cheaper, since they both involve a reduction in complexity.

such a purely economic perspective does not recognise the path dependencies that do seem to exist.

Adding a measure of politics brings only limited relief. What an economist would like to do, of course, is either to define some kind of “political costs and benefits” category, or to create a straightforward combination of Path Dependence and Public Choice Theory. The former is extremely difficult, since it involves an almost unbearable strain on the metaphor.¹²⁸⁴ The latter is certainly more viable, and would require a dynamic version of PCT: a model of changes – or the lack thereof – in power and rent-seeking over time. That is – of course – the appeal of Olson’s narrative of roving and stationary bandits¹²⁸⁵: It is a Public Choice model that involves the players explicitly making inter-temporal choices. However, the model is not entirely what we were looking for: the consequence of a bandit stealing too much is not a loss of power, but only a loss in wealth for his country. The power of the bandit at any given time is an exogenous variable. As such, it is insufficiently political for our present purposes.

From this, we might conclude that what we need is a model that formalises the intuition that one can trade money for power and power for money, i.e. a combination of Public Choice Theory and Path Dependence. Instead of resisting the railways to the utmost of their ability, the canal owners and their allies in 1830s Britain elected to accept the inevitable and trade their power for money in the short run. As a result, the rents paid by the railways to the landowners were extremely high, but the railways did get built. At the same time, many actors in powerful positions in Germany and France made the opposite choice: economic development would have increased the rents they could have obtained, but only at the expense of undermining their power. For that reason, the landed nobility of those countries initially traded money for power.

In subsequent periods, actors with power were faced with similar tradeoffs. Sometimes, the question was clear. A railway company should always oppose a reform that simultaneously increases the government’s involvement in the sector and decreases the scope for rent-seeking without some form of compensation. For this reason, there was no question that the French railway companies would not leave the initial proposals for the

¹²⁸⁴ Cf. Holterman (2007).

¹²⁸⁵ Cf. McGuire & Olson (1996) and Olson (2000).

1850s reform unopposed; they had to be watered down so that the companies' power to extract rents would be maintained.¹²⁸⁶ Only in 1937 did most of the major French railway companies choose cash over power.

Elected politicians, on the other hand, tend to worry less about obtaining rents, and more about the trade-off between the likelihood of obtaining power, and the magnitude of that power once obtained. That is how we could – if we were so inclined¹²⁸⁷ – understand privatisation and other forms of self-restraint: elected politicians voluntarily pledge to give up power, power that could potentially help them win votes in the future, in order to increase the likelihood of winning power in the first place.¹²⁸⁸

The only problem with this approach is, as we have seen in the preceding chapters, that power is too elusive a concept for an economist.¹²⁸⁹ Defining it by reference to the ability to affect the payoffs of others is too vague for actual use, while defining it as the ability to obtain rents is too narrow. After all, even though politicians and bureaucrats are severely limited in their ability to obtain rents for themselves, we would still say that they have significant power.¹²⁹⁰

In the end, we can probably not avoid the conclusion that economic science – strictly defined – can do no better than the alternatives presented in this chapter and its predecessor. While one could likely come up with other alternatives, theories and models as of yet unimagined, they too would have their strengths and weaknesses, and would therefore offer little additional explanatory power. Economics can only do so much, and we should resist the temptation to make it more interdisciplinary, a melting pot of different disciplines. Rather, we should tackle a problem like this in a multidisciplinary manner, with each discipline contributing what it can. For the railway industry, that probably means listening to political scientists first and foremost.

¹²⁸⁶ Cf. Caron (1997), p. 382-414.

¹²⁸⁷ The alternative would be, of course, to understand them as efforts to achieve an optimal governance structure, which is what we have done in chapter 3.

¹²⁸⁸ A variation of this explanation is the hypothesis that politicians favour privatisation because it obfuscates the true costs of the railways, as well as the responsibility for any failures. Cf. Gazendam & Homburg (1999), Groenendijk (1998), ch. 2 and Puviani (1903/1960).

¹²⁸⁹ And not just for an economist. Dahl (1957) and Dowding (1991) tackle the problem from a political science point of view, and Parsons (1963) is rooted in sociology.

¹²⁹⁰ Note that the same is not true for market power only. In microeconomics, the ability to extract rents is a characteristic of all market power. This is the basis of the Lerner index. Cf. Pindyck & Rubinfeld (2009), par. 10.2.

5. Conclusion

After the trials and tribulations of the preceding three chapters, it is finally time to return to our research questions: What is the value of the framework of the New Institutional Economics (NIE) for questions of government ownership and control of certain industries? In other words, to what extent can these decisions be explained and predicted as attempts to achieve an efficient governance structure for the industry in question? And can we identify efficiency in these circumstances with the recommendations of (some version of) the New Institutional Economics without more? In other words, how does NIE perform in dealing with such privatisation and nationalisation questions relative to other possible economic approaches?

5.1 Theoretical Findings

To answer those questions, an obvious first step is to examine the theory that is to be tested. This is something we did in chapter 2. Any economic theory that seeks to illuminate the workings of the law has to begin by making explicit what the law is, and which legal phenomena will and will not be considered.¹²⁹¹ Williamson, for example, has always quite enthusiastically endorsed the relational contract theory of Macaulay and MacNeil as his perspective on the law¹²⁹², a perspective that emphasises the limited value of lawyerly formalism in everyday life. That makes sense, because Williamson clearly had the ambition to model a wide range of institutions, both legal and other. In this thesis, however, we have seen a different approach. In principle, the goal was still to model all institutions – even though we might expect non-legal institutions to be less important in the context of public governance¹²⁹³ – but we have elected to describe only a subset of them as law. It is difficult to see what the cost has been of this gain in definitional clarity. While a gentlemen’s agreement is perfectly within the remit of

¹²⁹¹ Apparently, this is not as obvious as one might think. Cf. Hadfield & Weingast (2010) for a more detailed criticism of the literature.

¹²⁹² Cf. for example Williamson (1985), p. 68-76 and Williamson (1996), p. 370.

¹²⁹³ When the public sector is involved, values such as transparency and accountability will often preclude it from using the kinds of non-legal institutions described by Macaulay and MacNeil. Cf. Macaulay (1963) and MacNeil (1978).

institutional economics, be it old, new or neo, it is not law, and nothing is gained by acting as if it is.¹²⁹⁴

The next step was to take a look at the most universally accepted piece of wisdom that economists have applied to this type of problem: “In a world of zero transaction costs, the allocation of resources will be efficient, and invariant with respect to legal rules of liability, income effects aside.”¹²⁹⁵ As we have seen, this is about as baffling a statement as it is possible to make in so few words. What does it mean to speak of a world of zero transaction costs? What kind of efficiency is Coase talking about? What do we mean by “legal rules of liability”? How much rationality does this statement assume, and what happens if we would like to assume less of it? And finally, what if we do not want to leave “income effects aside”? Many of these questions are quite abstract and primarily of theoretical interest, meaning that they have few practical consequences for the manner in which we are supposed to think about the railways and similar industries until they have been further operationalised. Instead, the main takeaway is that contracting is costly, and that these costs consist mainly of the effort expended to search for an appropriate counterparty, to negotiate with them, and to monitor compliance with the contract.¹²⁹⁶ In situations where these costs are particularly high, rational actors will use other governance structures to avoid them, even when this comes at the expense of lower production efficiency.

This leads us to the next logical step, which is to attempt to describe and predict in which circumstances the costs of contracting will be “particularly high”. It is only at this stage that we encounter Williamson’s narrative of transaction-specific investments and hold-ups. While his discriminating alignment hypothesis is simply a version of the optimisation hypothesis that is a part of every economic model almost by definition¹²⁹⁷, the real meat of his work is in his description of the various transaction characteristics that lead to higher transaction costs. The only problem is that they are not the only ones, or even the ones that are necessarily most important in most situations.

¹²⁹⁴ On the contrary, as we have seen it can be quite useful to treat law as if it is nothing more than a gentlemen’s agreement, for example when enforcement is not realistic.

¹²⁹⁵ Zerbe (1980), p. 84.

¹²⁹⁶ To be precise, the transaction costs incurred after contracting consist of monitoring and bonding costs, since the counterparty also has an incentive to demonstrate its compliance, at least to some degree. Cf. Jensen & Meckling (1976).

¹²⁹⁷ Cf. the discussion of Robbins’ definition of economics in section 1.2.2, above.

Williamson's own case study of Oakland Cable Television¹²⁹⁸ illustrates this. It was not the cable company, the party making the transaction-specific investment, that was held up, but the city. Given the similarities between cable television and railways, it should come as no surprise that Williamson's case study proved more instructive for our present purposes than the model he presented later.

As an alternative to Williamson's TCE we have the various models by Grossman, Hart, and Moore, which were discussed in section 2.5. It is interesting to see that their manner of working is almost exactly the opposite of Williamson's. Where he uses words, they use calculus. Where he starts with the facts of the famous Fisher Body case study¹²⁹⁹ to work his way up, they start with abstract case positions to work their way down. The result is a series of propositions – or rather, several series of propositions – that are sometimes wonderful in their counterintuitiveness, sometimes regrettably impractical, but always beautifully illuminating. What these models have in common with Williamson's earlier effort is their emphasis on uncontractability and incompleteness: if it were feasible to write a complete contract, there would be no need to worry about hold-ups (Williamson) or misaligned incentives (GHM). According to Grossman, Hart, and Moore, it is only because certain investments cannot be contracted, that their amount will turn out too high or too low. This is, of course, where we find ourselves back with Coase's case studies.¹³⁰⁰ Although he does not discuss the issue in any great detail, it should be clear that in a world of non-zero transaction costs, his farmers might end up overinvesting or underinvesting in fencing, cattle or wheat.¹³⁰¹

In this way, we gathered together three different, and in no way competing, approaches to divining what the optimum governance structure might be. The only remaining problem was to consider why we would expect these results to be helpful in a positive model. Why would we expect to observe the optimum as opposed to any other possible governance structure? In a model with perfectly rational actors, the answer is obvious: perfectly rational actors will always choose the perfectly optimal alternative. However, given that NIE only assumes bounded rationality, the question becomes unavoidable.

¹²⁹⁸ Williamson (1976).

¹²⁹⁹ Cf. Klein, Crawford & Alchian (1978).

¹³⁰⁰ Cf. Coase (1960).

¹³⁰¹ Coase only briefly discussed the consequences of introducing transaction costs in section VI of Coase (1960), without revisiting his earlier case studies. It is left as an exercise for the reader to reflect on the exact circumstances in which each farmer would overinvest or underinvest in each of the goods.

The answer is that there are three mechanisms that tend to push the governance structure towards its optimal alignment: evolution, competition and rational design. As we have seen, all three will tend to be weaker when the government is one of the parties to the transaction than otherwise, and likewise all three will tend to be weaker in a small-n situation, where the number of transactions and the number of possible parties to any one transaction is small. In other words, even before we started with the detailed examination of the history of the railways, things were not looking good. The railways would seem to be the ultimate example of an industry where the outcome would tend towards the optimum only very slowly, if at all.

5.2 Historical Findings

In chapter 3, we endeavoured to apply these insights to the history of the railways. The difficulty is, of course, that even the most practical of economic theories, that is to say one that is cast entirely in observable variables, will still suffer from enormous measurement problems. For example, when Joskow first attempted to empirically test TCE, he looked at a very specific instance of asset specificity: the practice of building power plants next to coal mines.¹³⁰² That way, he could measure his exogenous variable with a dummy variable: some transactions involved asset specificity, others did not. As a result, he was able to estimate that mine-mouth power plants tended to have a contract duration that was 16 years longer than other power plants¹³⁰³, without having to say anything about the optimal contract duration or the efficiency of building a power plant next to a coal mine.

However, with the task we have set ourselves, we are not able to search for our keys under a lamp post in this way.¹³⁰⁴ So in order to preserve this wonderful ability to avoid saying anything too specific about the optimal governance structure in various circumstances, we were forced to state the theory in terms of changes: for example, whatever the optimum might be, an increase in the uncertainty of the transaction or a reduction of the observability and verifiability of performance should not be followed by a shift towards more hierarchy.¹³⁰⁵ Another way to examine the models without having to determine the theoretical optimum is to

¹³⁰² Cf. Joskow (1987).

¹³⁰³ The average for all observations was 12,75 years.

¹³⁰⁴ Cf. section 1.1, above.

¹³⁰⁵ N.B. the other two transaction characteristics that Williamson identified are less suitable for this kind of approach, since they remain more or less constant throughout the history of the railways. In other words, they would only come into play when making a comparison with other industries, for example an industry with lower investments in transaction-specific assets.

focus on perceived transaction costs; it would be interesting to know that certain decision makers, past or present, organised the railways a certain way because they were trying to avoid certain transaction costs, even if their choices they actually made were inefficient.

Looking at the actual historical narrative in search of changes in perceived transaction characteristics, the results were disappointing. The changes we observe are quite often in the opposite direction than what we would expect. For example, in the high-uncertainty environment of the first introduction of the technology, only Belgium elected to go public, and with astounding results. They built their core network more quickly, and less inefficiently, than any other country under consideration. In subsequent decades, as the level of uncertainty over this new technology decreased, and communication improved, all five countries – except Belgium, initially – moved towards greater state involvement in the industry, as well as more consolidation. Put simply, when they should have used the public option, they tended to go private, and when they should have privatised, they consolidated and regulated.

As the 19th century progressed, the governments' objective functions started to change, which again implied a shift in the theoretical optimum. Where initially they were happy to see the railways run in as many places as possible without any expense for the public purse, they increasingly came to value the military and socio-economic benefits of the railways. All of a sudden, railways had to be built in places where it was not otherwise profitable to do so, or they had to be run in a manner that was not profit-maximising. In order to get the railway companies to do what they would not do on their own, the government had to provide more consideration than the concession alone; they had to pay the companies somehow. Also, for the first time the government had to make sure that the companies were keeping up their end of the bargain. It is one thing to agree that there should be a railway service between two cities, but it is something entirely different to also agree that there should be at least so many trains, that some of them should be "penny trains", that they should adhere to certain minimum levels of safety and punctuality, for instance.

Grosso modo, the governance structures used followed the change in transaction characteristics in the manner predicted by the Coase and GHM models, albeit only with a delay of several decades, and with a somewhat roundabout way of arriving at the correct result. An increase in the number of demands the buyer has of the product will almost always make the contract more incomplete, because those added demands are rarely as observable or verifiable as the core issues of price and delivery. Greater contractual incompleteness, in turn, makes

contracting a less attractive governance choice, shifting the optimum away from the use of contracts and towards the use of hierarchy instead.

Instead of directly considering the implications of their increasingly complex wish list for the optimal governance structure of the railway industry in this way, government actors tended to attempt to impose said wish list in the least intrusive way possible. In other words, they tended to attempt to keep the old way of doing things intact as much as possible, much more than NIE would predict. As the various NIE theories would suggest, this caused enormous moral hazard problems, corresponding to bigger and bigger losses for the railway companies and bigger and bigger outlays for the state treasury. Complaints about the rising costs of the railways – and the lack of any tangible corresponding increase in quality – would then cause the various actors to start tweaking the governance structures. When such minor adjustments inevitably proved insufficient, a drastic change would follow, but only when the time was right, politically. It is not a coincidence that almost all major nationalisations discussed in this thesis were enacted in the wake of a World War. Even for Belgium and France, where this is not true, there was still the war of 1870-1871 and the Great Depression, respectively, that made nationalisation a more palatable option. The question of whether a given reform is “palatable” is not, however, one that NIE or economic science generally is equipped to deal with.¹³⁰⁶ The fact that we encounter it all the same is indicative of the importance of politics in the history of the railways.

It is not immediately obvious that the privatisation and liberalisation efforts of the last few decades can be explained in a similar manner, by pointing to changes in the government’s objective function. On the one hand, as public transportation is expected to offer an environmentally friendly and non-congested alternative to road transport, demands have increased still further since the time of nationalisation.¹³⁰⁷ On the other hand, the cost of contracting, specifically the costs of monitoring and bonding have decreased. For the first time, it is possible to actually measure which percentage of all trains arrives within 3 minutes of its scheduled arrival time.¹³⁰⁸ The net effect appears to be that this increase in the extent to which the gov-

¹³⁰⁶ Cf. Spiller & Tommasi (2003) for an attempt to model political transactions. In this context, public administration scholars often talk about “windows of opportunity”. Eg. De Hauteclouque & Perez (2011), p. 6.

¹³⁰⁷ Just look at the many things that are listed in the Passengers’ Rights Regulation, Regulation 1371/2007, OJ L 315, dd. 3.12.2007, p. 14.

¹³⁰⁸ According to the 2009 Annual Rapport of the Dutch former incumbent NS, in 2009 92,8% of their trains arrived within 5 minutes, compared to 93,0% in 2008. For a 3 minute cutoff, the numbers are 86,6% and 86,8%, respectively. (p. 42) In the Transportation plan 2010 (“Vervoerplan 2010”), the government and NS agreed on a target of 93% for the 5 minute cutoff, although there do not appear to be any direct consequences in case of a failure to achieve this norm.

ernment's wishes are observable and verifiable has had the greater effect, shifting the optimum away from hierarchy and towards contract-based governance structures, which is what we observed.

Unfortunately, there is once again a problem with this development. While all governments under examination unanimously dragged their feet on nationalisation, privatisation and liberalisation occurred sometimes too fast and sometimes too slowly. In Britain, the Tories put the ideological cart before the horse and privatised even transactions whose characteristics had changed little, such as track maintenance. While it may have become easier to measure punctuality, it is still as difficult as ever to observe and verify whether the maintenance company is maintaining the tracks as specified. The same goes for capacity allocation, a task that never existed before because the railway companies of the 19th century always owned their own track.¹³⁰⁹ Giving all these transactions the same governance structure is very much in conflict with any and all versions of NIE, something the British discovered the hard way.

In all other countries under consideration, the easiest way to sum up the problem is that there was a mismatch between the amount of liberalisation and the amount of privatisation. Liberalisation implies replacing hierarchical governance structures with governance structures more based on contracts and quasi-contracts. If this is done without privatisation, it puts the government in the awkward position of having to sue a company that is 100% publicly owned, in order to win damages or a fine that will come out of the pocket of the company's residual claimant, i.e. itself. Let us be frank, it takes a special kind of bureaucratic tunnel vision not to be embarrassed by the decision in September 2008 of the Dutch railway regulator – a Quango – to impose a total fine of € 876.000 on the Dutch rail infrastructure company ProRail – a limited liability company controlled by the state – for acting in violation of the relevant statutes in its capacity allocation decisions for 2007.¹³¹⁰ The only way that case could have been any more bizarre is if it had been based on a complaint by NS, the incumbent railway company which is a plc – a public limited company – also completely controlled by the

¹³⁰⁹ With the notable exception of the Netherlands between 1890-1917, when the SS and the HSM shared certain sections of track, with predictably disastrous consequences. Cf. section 3.3.3, above.

¹³¹⁰ Decision of the Dutch Competition Authority, Transport Division, dated 10 September 2008, No. 200069/57, in the matter of Capacity Allocation 2007 – Connexxion.

state.¹³¹¹ However, as it happens the original complaint came from Connexxion, a prominent Dutch bus company that also operates a few railway lines.¹³¹²

In this way, the current Dutch system – like the rest of the European continent – combines a partial liberalisation, which creates a system of bilateral monopoly on the main lines, with a partial privatisation, which means that the government is negotiating with a company it owns but does not fully control.¹³¹³ In such a regulatory environment, the outcomes are extremely difficult to predict, depending on the relative bargaining power of each side, which can differ wildly depending on the political posture of the disagreement and the identity of the individuals doing the bargaining. It makes all the difference whether the government is represented by ministers or parliamentarians working under – more or less serious – scrutiny from the media and the general public, or whether the problem is handled at the level of civil service bureaucrats working with a – more or less – clear mandate from their minister.

This is not just to say that the governance structure chosen is a hybrid characterised by occasional significant political involvement, the way the governance structure of the Belgian railways from the 1860s to the founding of the NMBS involved politicised decision making. Rather, the industry is politicised on a higher level as well; the very question of how vulnerable to politicisation the governance structure is, is a constant matter of political debate and manipulation. From one cabinet to the next, from one (winter) crisis to the next, the governance structures in place can change quite suddenly. More unsettlingly, the solution offered is not necessarily connected in any obvious way to the problem observed: underperformance of the railways can both lead to more and less regulation, depending on the ideological glasses through which the problem is analysed by the various actors, and the degree to which the government manages to successfully spin the story. Because the status quo is a contract/hierarchy hybrid, it is usually possible to make an argument for reform in either direction, at least as long as the audience is sufficiently ignorant of the details of the transaction.

¹³¹¹ NS have generally refrained from seeking fines or penalty payments from ProRail. Cf. for example the Decision of the Dutch Competition Authority, Transport Division, dated 2 September 2008, No. 200094/57, in the matter of NS v. ProRail – Capacity Allocation 2008/Network Statement 2007, specifically par. 192.

¹³¹² One third of the shares of Connexxion are in the hands of the Dutch Ministry of Finance, while the remainder belong to a consortium including the Dutch BNG – Bank of Dutch Municipalities – and the French transport company Transdev, which in turn is owned by the French government through the Caisse des dépôts et consignations. In other words, while Connexxion is less “public” than NS, there is still not a cent of private money involved in its equity.

¹³¹³ Cf. Book 2 Title 4 of the Dutch Civil Code, which deals with public limited companies, and specifically section 6 of Title 4, which reduces the power of the shareholders still further for plcs of a certain size.

5.3 Findings on Benchmarks

The findings outlined in the previous section would naturally lead one to suggest that public choice theory might offer a solution. After all, that is the most “political” branch of economics. However, the analysis of chapter 4 demonstrated that PCT offers only limited help. This is true for two main reasons: on the one hand, PCT has the same limitations as economic theory generally, because it is not based on the denial of any of economic theory’s main assumptions. It is still based on actors of near-perfect rationality, and the concomitant link between normative and positive statements. On the other hand, much of what public choice theory would normally have to offer is already included in NIE as it was discussed in chapter 2. For example, because NIE explicitly models government intervention in the economy as a transaction between it and one or more private parties – and one taking place in an environment of bilateral monopoly at that – many kinds of rent seeking are already included in the model. For example, it is perfectly fine to say that the railway companies were engaging in rent seeking when they tried to persuade the government to put the gates at a level crossing across the road instead of across the railroad, or when they were trying to talk the government into ever greater subsidies for their efforts. However, such observations can be modelled just as easily within NIE as negotiations about the terms of the contract between the railways and the state.

What we would need most from such a politically-oriented model is guidance as to the timing of reforms. After all, NIE is broadly correct in its predictions of the trends of reform, but the variation around that trend is enormous. Whether a given reform occurs in one decade or the next seems to depend less on the size of the inefficiency than on whether the political climate is ready, whatever that means. Translating this notion of political climate into public choice theory, however, proved difficult. It appears that there is little correlation between the political climate and the predictable self-interest of the relevant government actors, especially once democracy became fully established in all five countries. In the European countries that we have examined, there has never been a position on railway reform that was an automatic vote-winner. The ordinary citizenry generally cares more about other things.

Of course, the model that takes a look at timing more than any other is path dependence, the third approach we considered in this thesis. Path dependence explicitly models changes over time, making it a perfect framework for improving the value of NIE as a positive model. However, there, too, we quickly run into difficulties. Institutions are not typewriters, and to

the extent that it is possible to carry the analogy through, you end up with an analysis that rests almost entirely on variables that are difficult – or impossible – to measure. How is an economist supposed to understand the idea of investing political capital?¹³¹⁴ In the best of circumstances, we would still end up talking about investments in human capital a lot, since that is the only traditional form of capital at issue here. After all, creating institutions does not require any other kind of economically cognisable investment. And even such investments are problematic because the kind of human capital we tend to have in mind is usually gained by investing effort, not money. In other words, not even the ruse of valuing human capital by the amount of money invested in it¹³¹⁵ will allow us to apply the path dependence framework in a properly economic manner most of the time. Instead, the concept of path dependence will occasionally allow us to be more concrete in our discussion of alignment mechanisms, and otherwise it will offer a starting point for a more multidisciplinary approach to the analysis of railway regulation, something we have deliberately not offered in the present thesis.

Bringing it all together gives us an overwhelming number of buttons to push and dials to turn, yet the result is still far from satisfactory. Even when faced with the easiest of challenges – answering questions that we already know the answer to – the result is far from encouraging. Even the combination of all three models does not seem to be able to predict in which decade we might expect certain major reforms. As it happens, the French SNCF was created in the 1930s, but it could as easily have happened in the 1920s or the 1940s. To be sure, in the 1930s the private railway companies were particularly cash-strapped, but how are we meant to know how poor is too poor? At which level of underperformance do railway companies flee in the arms of the state? The answer is: we do not know. It depends. It depends on the circumstances of the case, circumstances that are entirely historical in nature. Idiosyncratic

For example, there is no question that it helped that – for reasons unconnected to the railways – the French parliamentary elections of 1936 had given the left a firm control of the Chamber of Deputies, making Leon Blum prime minister two times in the late 1930s. At the time, with the country suffering through the Great Depression, the people were in the mood for drastic left-wing government. For that reason, the Popular Front tried to give it to them, by enacting a shorter working week, two weeks paid vacation, promoting collective bargaining,

¹³¹⁴ Cf. Holterman (2007).

¹³¹⁵ Cf. Becker (1964).

as well as – as we have seen – nationalising the railways. How is any economic model supposed to predict all of this? Our theories allow us to estimate that this nationalisation was perhaps as much as 50 years overdue, but there is no way that we could have estimated the exact decade when government policy was going to catch up with the changes in transaction characteristics except in an exceedingly historicist manner.

To be specific, we can imagine improving our model by combining our public choice theoretic modelling with the parts of macroeconomics that predict and explain such phenomena as the Great Depression, in order to postulate that such a depression causes voters to put pressure on their governments to enact more interventionist policies, including nationalisation. Put simply: whenever the economy is bad, the government starts doing drastic things. Such an add-on would certainly help for the case of France, where several railway reforms coincided with economic crises.¹³¹⁶ In the other four countries, however, reforms generally correlated with major wars more than with economic crisis, and even in institutional economics war and peace is still the ultimate exogenous variable. Maybe there are other ad-hoc fixes we could think of in order to improve the explanatory power of our models with regard to the history so far. Introducing dummy variables would certainly help, in order to operationalise the unmodellable. That way we can decide that reforms following wars should be treated differently, or reforms carried out by newly established regimes or countries, or any other categories that we might care to dream up. None of that would help, though, if the goal is to answer questions that we do not yet know the answer to.

That said, it is clear that at least NIE does not underperform relative to the other approaches we have considered. They all have their own strengths and weaknesses, and their overall limitations represent the boundaries of economics as a science. Assuming that we insist on saying a little about a lot, instead of a lot about a little, we will always have to make do with only highly imprecise and vague answers to our questions. That is a problem that is shared by all fields of economics. The current economic crisis has only reminded us how difficult it is to predict an economic downturn and how even sophisticated risk modelling can fail spectacularly to protect a bank against bankruptcy. Even looking back, it is difficult enough to explain

¹³¹⁶ Not just in the 1930s, but also in the 1840s and 50s, when the newly created private railway companies were almost immediately bankrupted by the crisis of the 1840s, leading to a number of bailouts. Cf. section 3.2.3, above.

how we got into this mess. In that light, perhaps we should not be too hard on NIE as a positive model.

5.4 Positive and Normative Statements

In reviewing these findings, it is noteworthy that economics, as a science, encounters another problem in the area of semi-public institutions, besides the small-n problem and the politics-based decision making to which that leads. The link between positive and normative statements appears to be particularly weak, and that affects the usefulness of both.

Consider the behavioural assumptions of our work: While only Oliver Williamson is explicit in his assumption of opportunistic behaviour, all economic models by definition require a measure of rationality, defined as goal-oriented behaviour. This goal is defined as utility and largely treated as an exogenous variable, and the actor is expected to work towards optimising it. A naive application of this assumption to the Quango-phenomenon would lead the economist to predict that these organisations should run amok, restrained only by the possibility that their leaders might get fired or reassigned. Since many statutory arrangements make such remedies unavailable or at least unlikely, the prediction would be that the Quango does whatever it feels like.

Why should Dutch infrastructure owner ProRail care whether the trains run on time? They do not own the trains, or have any other kind of financial stake in their punctuality. To be sure, they have a contract with the government which says that they have to care about such things, but what is the government going to do if ProRail takes a few liberties here and there? It is not like there is a different railway infrastructure company that could step in and do the job. The board has a remuneration scheme that involves bonuses, but those bonuses are relatively small compared to the fixed salary.¹³¹⁷ And would the government, in its capacity as only shareholder, really use its power to summarily dismiss the board members?¹³¹⁸ A logical, rational economic analysis along the lines of decades of agency theory would suggest that the members of the board, and all those that work for the company, would cut corners just so much that they avoid getting caught. Their work is extremely difficult for the government to

¹³¹⁷ In 2009, the three board members had a total fixed salary of € 443.722, and a total amount in bonuses of € 49.495.

¹³¹⁸ Cf. art. 11.4 of the company bylaws, which are available on www.prorail.nl.

supervise, since most of the expertise rests with ProRail and not with the ministry.¹³¹⁹ Only the former know whether it is really necessary to carry out certain works during the day instead of during night-time, for instance.

In actual practice, however, such positive statements do not appear to describe the functioning of the railway industry very well. The board and staff of ProRail try much harder to make the trains run on time than economic theory would suggest. As discussed in chapter 4, above, the most likely explanation for this is probably something along the lines of bureaucratic work ethic; exactly because their employment relationship with the general public – their ultimate principal – is not based on mistrust and incentives, the railway employees feel obligated to do their best to fulfil society's expectations.¹³²⁰ This sense of obligation therefore enters into their utility function, together with such considerations as the desire for salary, for perking and for job security, to name but a few.¹³²¹

The problem with this phenomenon is that it throws off normative statements based on the positive claims mentioned earlier. If people do not respond to incentives in the manner predicted by (naive) economic models, then maybe the recommendation to introduce incentives where none existed should be re-evaluated. An obvious recommendation that can be drawn from this thesis, for example, is that infrastructure companies such as ProRail and RailTrack should not be privatised or liberalised, because there is no market for their services. Theirs is the ultimate natural monopoly, one that can only be broken if the state is willing to go back to the pre-nationalisation approach of having competing connections between cities. While that approach is already in use in the UK for the railway franchises, i.e. the companies that actually operate the train service, it seems infeasible for the track companies unless they are once again vertically integrated with the transport companies.¹³²² If there is no market to discipline the infrastructure companies, it follows by impeccable economic logic that they must be disciplined through hierarchy instead. Even performance contracts will probably not suffice.¹³²³

¹³¹⁹ Cf. Murray (2001).

¹³²⁰ For government employees generally, cf. Raad van Economisch Adviseurs (2007), p. 5.

¹³²¹ Cf. Wilson (1989).

¹³²² Apparently, PM Major inclined towards this solution for quite a long time before giving in to the judgement of his cabinet colleagues and civil servants. Cf. sections 3.4, 4.2 and 4.4, above.

¹³²³ In practice, governments tend to rely on performance-related bonuses, while being under constant political pressure to keep those bonuses as low as possible. Cf. Financial Times, April 2, 2009, 'Network Rail chief defends bonus policy'.

In actual fact, the practitioners who looked at this problem often had good reasons for suggesting that the infrastructure companies should be placed at some distance from the political decision makers. In the 1992 Wijffels rapport about the reform of the Dutch railway industry, for example, we read that a contract based approach helps to keep the costs of investments in new infrastructure clear and transparent, to make the correct actors accountable for budget overruns, and to help political decision makers supervise the efficiency of the sector.¹³²⁴ Needless to say, this is not something that would have immediately occurred to a theoretical economist, who would have assumed that all those things can be done within a single organisation as well. In fact, NIE generally assumes that using more hierarchy improves information flows, if for no other reason that lower powered incentives reduce moral hazard, giving actors less reason to hide information.

Unlike descriptive statements, which are useful even if they are in the little-about-a-lot form, normative statements are either helpful or they are not. If economic science has an unacceptably high probability of pointing us in the wrong direction, we are better off not having it at the table. The question of what constitutes an “unacceptably” high probability is better left for another day, but at least we can safely say that it would be preferable if the counterintuitiveness that we so value in our descriptive statements were treated with a little more mistrust in the normative sphere. Perhaps that is the best way to sum up this thesis: It would be better if major reforms were not carried out *only* on the say so of economic science. Fortunately, as a practical matter, there is little evidence that there ever has been.

5.5 NIE & the Public Sector: The Future

With these observations in mind, we can conclude this thesis by saying a few words about the future of NIE as applied to the public sector, as well as about semi-public institutions themselves. Given that Coase only barely survived the challenge posed by the history of the railways, the next step is to consider whether there are any obvious ways to improve the framework, improvements that we have not yet explored in chapters 2 and 3? Is it likely that the various versions of NIE will be applied by government actors more than they have been so far? And what will become of our beloved railways?

¹³²⁴ Cf. p. II-23.

First of all, what NIE certainly needs is an operationalisation of transaction costs that preserves the best of both current alternatives. It would need to be grounded in practical observations and suitable for practical application, as Williamson's TCE is, but at the same time it would have to be capable of application to a widespread range of transactions. Probably the best place to start is to think about the connection between mainstream NIE as it is and agency theory. In this thesis, we have devoted relatively little attention to examining how exactly these lines of theory fit together. While this is unquestionably true, a more detailed analysis of their confluence would probably yield valuable insights for non-agency transactions as well. After all, it is not unnatural to model an ordinary transaction as a combination of two agency relations: each party is the agent for the consideration they provide, and the principal for the consideration they receive. It should be possible to identify some useful – to stick with Williamson's parlance – transaction characteristics in this way. In so doing, there is no need to preserve Williamson's probity variable¹³²⁵, since it does not appear to have any added value in a model that explicitly takes moral hazard into account.

Secondly, NIE needs further reflection on the role of the law – and legal scholarship – in its analysis. While legal scholarship has been debating for decades about the merits of abandoning the formalism that has traditionally characterised their endeavours¹³²⁶, preciously few institutional economists seem interested in travelling in the opposite direction. This fact is aptly illustrated by the present thesis, which presents a wealth of legal detail, which more often than not gets abandoned when the time comes for serious economic analysis. And that is not merely an artefact of the relatively broad perspective chosen here: five countries over 185 years. Even if the case study had been chosen more narrowly – as in Holtermann (2009), which focused on just 10 years of the history of a single railway – the bulk of the legal nuance would still have been lost.

Consider what we got in return for adopting a more “legal” definition of a transaction.¹³²⁷ To be sure, given that we were considering a governance problem located in an environment with highly sophisticated actors and high stakes, adopting this definition did not cause any noticeable problems, but in what sense did it actually improve the analysis? Conceptually, it has the effect of focusing our attention away from futile attempts to pretend that transactions with a firm have something in common with transactions between firms, and to-

¹³²⁵ Cf. Williamson (1999).

¹³²⁶ This formalism is on full display in section 2.1, above.

¹³²⁷ Cf. section 2.3.

wards the more practical literature on management and (new) public management. Of course, that is what we would end up doing regardless, but at least now we are being honest about it.

Taking all things together, further work in this direction would likely produce a more useful theory of government intervention in the economy. Such a theory would be conscious of its limitations, both its limitations *ratione materiae* and its limitations in terms of the kinds of results that one can expect from an economic model, but within those limitations it would be able to help governments take decisions about privatisation and nationalisation in a way that has so far eluded us. This thesis represents a modest attempt at promoting this goal, but obviously there is much work yet to be done.

As for the railways, it would help a lot if government actors would stop hedging their bets and start being explicit about the thinking that underlies their work. We have long since stopped thinking of privatisation as the magic cure for all ailments, so if privatisation is really to be helpful, government actors should explain why, and then they should actually privatise the railways. If, on the other hand, the solution lies in some form of liberalisation without privatisation – which is what most countries have today – then the government should explain how this can work, and then they should design a clear and transparent system of governance structures with sufficient control rights left to the government itself. In that case, the biggest pitfall would probably be to design a system with too many elements derived from private law. Liberalisation without privatisation is not a private law problem, and it would be foolish in the extreme to let the rules of private law get in the way of designing sufficient, and sufficiently effective, control rights for the government.¹³²⁸

¹³²⁸ Cf. art. 5(4) of the Commission Proposal for a Directive establishing a Single European Railway Area (recast), COM(2010) 475F, which states that the government in its capacity of shareholder of a railway undertaking should not have more management and control rights than a normal shareholder in a normal company under the country's national company law.

Literature

1. Abel, A.B. & B.S. Bernanke (2005), *Macroeconomics*, international edition, Boston: Pearson.
2. Aderlini, L. & L. Felli (2006), 'Transaction Costs and the Robustness of the Coase Theorem', *Economic Journal*, Vol. 116, No. 1, p. 223-245.
3. Adler, B.E. (2008), 'Efficient Breach Theory through the Looking Glass', *New York University Law Review*, Vol. 83, No. 6, p. 1679-1725.
4. Akerlof, G.A. (1970), 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism', *Quarterly Journal of Economics*, Vol. 84, No. 3, p. 488-500.
5. Albert, W. (2007), *The Turnpike Road System in England: 1663-1840*, Cambridge: Cambridge University Press.
6. Alchian, A.A. & H. Demsetz (1972), 'Production, Information Costs, and Economic Organization', *American Economic Review*, Vol. 62, No. 5, p. 777-795.
7. Alchian, A.A. (1950), 'Uncertainty, Evolution, and Economic Theory', *Journal of Political Economy*, Vol. 58, No. 3, p. 211-221.
8. Alchourrón, C.E. & E. Bulygin (1971), *Normative Systems*, Vienna: Springer.
9. Allen, D.W. (1991), 'Book Review: The Firm, the Market, and the Law', *Canadian Journal of Economics*, Vol. 24, No. 3, p. 740-743.
10. Allen, D.W. (1999), 'Transaction Costs', in: B. Bouckaert & G. De Geest [eds.], *Encyclopedia of Law and Economics, Volume 1. The History and Methodology of Law and Economics*, Cheltenham: Edward Elgar.
11. Alston, L.J. (2008), 'The "Case" for Case Studies in New Institutional Economics', in: E. Brousseau & J-M. Glachant, *New Institutional Economics: A Guidebook*, Cambridge: Cambridge University Press, p. 103-121.
12. Arkel, E.G. van (2007), *A Just Cause for Dismissal in the United States and the Netherlands*, Ph.D. Thesis Erasmus University Rotterdam, The Hague: Boom Juridische Uitgevers.
13. Arkes, H.R. & C. Blumer (1985), 'The Psychology of Sunk Costs', *Organizational Behavior and Human Decision Processes*, Vol. 35, p. 124-140.
14. Arkes, H.R. & L. Hutzel (2000), 'The Role of Probability of Success Estimates in the Sunk Cost Effect', *Journal of Behavioral Decision Making*, Vol. 13, p. 295-306.
15. Arnold, A.J. & S. McCartney (2005), 'Rates of Return, Concentration Levels, and Strategic Change in the British Railway Industry, 1830-1912', *Journal of Transport History*, Vol. 26, No. 1, p. 41-60.
16. Arrow, K.J. (1963), *Social Choice and Individual Values* (rev. ed. based on the original 1951 ed.), New York: John Wiley & Sons.
17. Arrow, K.J. (1969), 'The organization of economic activity: Issues pertinent to the choice of market versus nonmarket allocation', in: *The Analysis and Evaluation of Public Expenditure: The PPB System*, Vol. 1, US Joint Economic Committee, 91st Congress, 1st Session, Washington DC: US Government Printing Office, p. 59-73.
18. Arthur, W.B. (1989), 'Competing Technologies, Increasing Returns, and Lock-In by Historical Events', *Economic Journal*, Vol. 99, No. 394, p. 116-131.
19. Arthur, W.B. (1994), *Increasing Returns and Path Dependence in the Economy*, Ann Arbor (MI): University of Michigan Press.
20. Athias, L. & A. Nuñez (2008), 'The More the Merrier? Number of Bidders, Information Dispersion, Renegotiation and Winner's Curse in Toll Road Concessions', available at <http://ssrn.com/abstract=1269630>.

21. Axelrod, R. (1981), 'The Emergence of Cooperation among Egoists', *American Political Science Review*, Vol. 75, No. 2, p. 306-318.
22. Ayres, I. & R. Gertner (1989), 'Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules', *Yale Law Journal*, Vol. 99, No. 1, p. 87-130.
23. Ayres, I. & R. Gertner (1992), 'Strategic Contractual Inefficiency and the Optimal Choice of Legal Rules', *Yale Law Journal*, Vol. 101, No. 4, p. 729-773.
24. Backhouse, R.E. & S.G. Medema (2009), 'Defining Economics: The Long Road to Acceptance of the Robbins Definition', *Economica*, Vol. 76, No. s1, p. 805-820.
25. Backhouse, R.E. (2003), 'The Stabilization of Price Theory, 1920-1955', in: W.J. Samuels, J.E. Biddle & J.B. Davis, *A Companion to the History of Economic Thought*, Oxford: Blackwell.
26. Bajari, P. & S. Tadelis (2001), 'Incentives versus Transaction Costs: A Theory of Procurement Contracts', *RAND Journal of Economics*, Vol. 32, No. 3, p. 387-407.
27. Baldwin, C.Y. (2008), 'Where do transactions come from? Modularity, transactions, and the boundaries of firms', *Industrial and Corporate Change*, Vol. 17, No. 1, p. 155-195.
28. Barea, J., D. Dizy & O. Ruiz (2007), 'The New Model of the Railway Industry in Spain within the European Framework', *Annals of Public and Cooperative Economics*, Vol. 78, No. 3, p. 353-380.
29. Baron, D. & R. Myerson (1982), 'Regulating a Monopolist with Unknown Costs', *Econometrica*, Vol. 50, No. 4, p. 911-930.
30. Barro, R.J. & R.M. McCleary (2011), 'Saints Marching In, 1590-2009', NBER Working Paper No. 16769.
31. Barro, R.J. (1974), 'Are Government Bonds Net Wealth?', *Journal of Political Economy*, Vol. 82, No. 6, p. 1095-1117.
32. Barry, N.P. (1979), *Hayek's Social and Economic Philosophy*, London: MacMillan.
33. Barzel, Y. (1982), 'Measurement Cost and the Organization of Markets', *Journal of Law and Economics*, Vol. 25, p. 27-48.
34. Baskerville, R.F. & W. O'Grady (2007), 'Does Darwin Belong in Business? The Danger and Comfort of the Evolutionary Metaphor', Victoria University of Wellington School of Accounting and Commercial Law, Working Paper no. 44. Available at: <http://www.victoria.ac.nz/sacl/CAGTR/workingpapers/WP44.pdf>.
35. Baskerville, R.F. (2007), 'Icons of Repute: The attribution of Lamarckian and Darwinian evolutionary mechanisms in economics', Victoria University of Wellington School of Accounting and Commercial Law, Working Paper no. 43. Available at: <http://www.victoria.ac.nz/sacl/CAGTR/workingpapers/WP43.pdf>.
36. Batina, R.G. & T. Ihori (2005), *Public Goods: Theory and Evidence*, Berlin: Springer.
37. Baumol, W.J., J.C. Panzar & R.D. Willig (1982), *Contestable Markets and the Theory of Industry Structure*, San Diego: Harcourt Brace Jovanovich.
38. Bazerman, M.H. & W.F. Samuelson (1983), 'I Won the Auction But Don't Want the Prize', *Journal of Conflict Resolution*, Vol. 27, No. 4, p. 618-634.
39. BCG (2007), 'The Relentless Advance of Private Label: Strategies to Compete', available at: <http://www.bcg.com/publications/files/TheRelentlessAdvanceofPrivateLabelStrategiestoCompeteApr07.pdf>.
40. Becker, F. & P. Hoekstra (1999), 'Marktwerking in het Openbaar Vervoer: het personenvervoer op het spoor', Verkennende Notitie voor de Commissie Van Thijn t.b.v. expertmeeting d.d. 16-6-1999.
41. Becker, G.S. & K.M. Murphy (1988), 'A Theory of Rational Addiction', *Journal of Political Economy*, Vol. 96, No. 4, p. 675-700.

42. Becker, G.S. & K.M. Murphy (1988), 'A Theory of Rational Addiction', *Journal of Political Economy*, Vol. 96, No. 4, p. 675-700.
43. Becker, G.S. & R. Posner (1993), 'Cross-Cultural Differences in Family and Sexual Life: An economic analysis', *Rationality and Society*, Vol. 5, No. 4, p. 421-431.
44. Becker, G.S. (1964), *Human Capital*, Chicago: University of Chicago Press.
45. Becker, G.S. (1981a), *A Treatise on the Family*, Cambridge: Harvard University Press.
46. Becker, G.S. (1981b), 'Altruism in the Family and Selfishness in the Market Place', *Economica*, Vol. 48, No. 189, p. 1-15.
47. Becker, G.S. (1993), 'Nobel Lecture: The Economic Way of Looking at Behavior', *Journal of Political Economy*, Vol. 101, No. 3, p. 385-409.
48. Ben Sassi, S. (2008), *Développement et organisation du secteur logistique: une analyse néo-institutionnelle*, Ph.D. thesis, Université Paris X, available at: <http://bdr.u-paris10.fr/theses/internet/2009PA100001.pdf>.
49. Berk, J. & P. DeMarzo (2007), *Corporate Finance*, Boston: Pearson.
50. Berle, A.A. & G.C. Means (1932), *The Modern Corporation and Private Property*, New Brunswick, NJ: Transaction Publishers.
51. Berlin, I. (1969), *Four Essays on Liberty*, Oxford: Oxford University Press.
52. Berlin, I. (2002), *Freedom and its Betrayal: Six Enemies of Human Liberty*, ed. By H. Hardy, London: Random House.
53. Berman, S.E. (2001), 'Modernization in Historical Perspective: The Case of Imperial Germany', *World Politics*, Vol. 53, p. 431-462.
54. Bernstein, L. (2001), 'Private Commercial Law in the Cotton Industry: Creating Cooperation Through Rules, Norms, and Institutions', *Michigan Law Review*, Vol. 99, p. 1724 c.a.
55. Besanko, D, D. Dranove & M. Shanley (2000), *Economics of Strategy* (2nd edition), New York: John Wiley & Sons.
56. Bick, A. (2008), 'Bernard Mandeville and the 'economy' of the Dutch', *Erasmus Journal of Philosophy and Economics*, Vol. 1, No. 1, p. 87-106.
57. Binmore, K. & P. Klemperer (2002), 'The Biggest Auction Ever: The Sale of the British 3G Telecom Licenses', *Economic Journal*, Vol. 112, p. C74-C96.
58. Blackstone, W. (1765), *Commentaries on the Laws of England, Vol. 1*.
59. Bleakley, H. & J. Lin (2010), 'Portage: Path Dependence and Increasing Returns in US History', NBER Working Paper No. 16314.
60. Bloembergen, A.R. et.al. (2001), *Rechtshandeling en Overeenkomst*, Deventer: Kluwer.
61. Bohlen, F.H. (1929), 'Misrepresentation as Deceit, Negligence, or Warranty', *Harvard Law Review*, Vol. 42, No. 6, p. 733-747.
62. Boland, L.A. (2002), *The Principles of Economics: Some lies my teachers told me*, London: Routledge.
63. Bond, J. (2007), 'Canal Construction in the Early Middle Ages: An Introductory Review', in: J. Blair [ed.], *Waterways and Canal-Building in Medieval England*, Oxford: Oxford University Press, p. 153-206.
64. Bordes, de (1886), 'De Spoorwegen in Italië', *De Economist*, Vol. 35, No. 1, p. 123-148.
65. Bork, R.H. (1966), 'Legislative Intent and the Policy of the Sherman Act', *Journal of Law and Economics*, Vol. 9, p. 7-48.
66. Breyer, S. (1970), 'The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs', *Harvard Law Review*, Vol. 84, p. 281-351.
67. Briotti, M.G. (2005), 'Economic Reactions to Public Finance Consolidation: A Survey of the Literature', European Central Bank Occasional Paper Series, No. 38.
68. Broeke, W. van den & E. van Nederveen Meerkerk (2001), 'Spoorlijnen en geldstromen. Een onderzoek naar de financiers van de Nederlandsche Rhijnspoorweg-Maatschappij 1845-1890', *Nederlands Economisch-Historisch Archief Jaarboek*, Vol. 64, p. 154-177.

69. Broeke, W. van den (1989), 'Het spoor terug gevolgd. De eerste honderd jaar (1839-1939).' in: J.A. Faber [ed.], *Het Spoor: 150 jaar spoorwegen in Nederland*, Amsterdam: Meulenhoff, p. 11-51.
70. Brooks, M.A. & B.J. Heijdra (1988), 'An Exploration of Rent-seeking', University of Tasmania Economics Discussion Paper No. 88-07.
71. Brousseau, E. (2001), 'Did the Common Law Bias the Economics of Contract... and May it Change?', in: B. Deffains & T. Kirat [eds.], *Law and Economics in Civil Law Countries*, Vol. 6 of N. Mercurio [ed.], *The Economics of Legal Relationships*, JAI Press.
72. Buchanan, J.M. & G. Tullock (1962), *The Calculus of Consent: Logical Foundations of Constitutional Democracy*, Ann Arbor (Mich): University of Michigan Press.
73. Buchanan, J.M. (1979), 'Politics without Romance: A Sketch of Positive Public Choice Theory and Its Normative Implications', Inaugural Lecture, Institute for Advanced Studies, Vienna, Austria, *Zeitschrift des Instituts für Höhere Studien*, Vol. 3, p. B1-B11.
74. Buchanan, J.M. (1983), 'Rent seeking, noncompensated transfers, and laws of succession', *Journal of Law and Economics*, Vol. XXVI, p. 71-85.
75. Buelens, F, J. van den Broeck & H. Willems (2008), 'British and French Investments in the Belgian Railway Sector During the Nineteenth Century', in: Roth, R. & G. Dinhl (2008), *Across the borders: Financing the World's Railways in the Nineteenth and Twentieth Centuries*, Aldershot: Ashgate, p. 95-108.
76. Cafaggi, F. & H.-W. Micklitz (2007), 'Administrative and Judicial Collective Enforcement of Consumer Law in the US and the European Community', EUI Working Paper LAW 2007/22.
77. Cafaggi, F. & H.-W. Micklitz [eds.] (2009), *New Frontiers of Consumer Protection: Interplay between Private and Public Enforcement*, Morsel (BE): Intersentia.
78. Calabresi, G. & A.D. Melamed (1972), 'Property Rules, Liability Rules, and Inalienability: One View of the Cathedral', *Harvard Law Review*, Vol. 85, No. 6, p. 1089-1128.
79. Cameron, R.E. (1991), 'Introduction', in: R.E. Cameron & V.I. Bovykin [eds.], *International Banking: 1870-1914*, Oxford: Oxford University Press, p. 3-24.
80. Carlson, R.E. (1969), *The Liverpool & Manchester Railway Project, 1821-1831*, Newton Abbot: David & Charles.
81. Carney, W.J. (1999), 'Limited Liability', in: B. Bouckaert & G. De Geest [eds.], *Encyclopedia of Law and Economics, Volume 3. The Regulation of Contracts*, Cheltenham: Edward Elgar.
82. Caron, F. (1997), *Histoire des Chemins de Fer en France, Tome Premier: 1740-1883*, Paris: Fayard.
83. Caron, F. (2005a), *Histoire des Chemins de Fer en France, Tome Second: 1883-1937*, Paris: Fayard.
84. Caron, F. (2005b), *Les grandes compagnies de chemen de fer en France: 1823-1937*, Archives économiques du Crédit Lyonnais, No. 5, Genève: Librairie DROZ.
85. Catalano, J.M. (1997), 'More Fiction than Fact: The Perceived Differences in the Application of Specific Performance under the United Nations Convention on Contracts for the International Sale of Goods', *Tulane Law Review*, Vol. 71, p. 1807-1834.
86. Centraal Planbureau (1995), 'Werkdocument no. 75: Economische effecten van de Betuweroute op basis van recente informatie', The Hague: Centraal Planbureau, April 1995.
87. Chadwick, E. (1859), 'Results of Different Principles of Legislation and Administration in Europe; of Competition for the Field, as Compared with Competition within the Field, of Service', *Journal of the Statistical Society of London*, Vol. 22, No. 3, p. 381-420.
88. Chandler, A.D. (1962), *Strategy and Structure*, Cambridge, Mass: MIT Press.
89. Chandler, A.D. (1965), *Railroads: The Nation's First Big Business*, New York: Columbia University Press.

90. Chatelier, M. Le (1845), *Chemins de Fer d'Allemagne: Description statistique, système d'exécution, trace, voie de fer, stations, materiel, frais d'établissement, exploitation, produit de l'exploitation*, Paris: L. Mathias.
91. Chesterman, S. (2008), "'We Can't Spy... If We Can't Buy!': The Privatization of Intelligence and the Limits of Outsourcing 'Inherently Governmental Functions'", *European Journal of International Law*, Vol. 19, No. 5, p. 1055-1074.
92. Cheung, S.N.S. (1969), *A Theory of Share Tenancy*, Chicago: University of Chicago Press.
93. Cheung, S.N.S. (1983), 'The Contractual Nature of the Firm', *Journal of Law and Economics*, Vol. 26, No. 1, p. 1-21.
94. Choi, J.J. et al. (2003), 'Optimal Defaults', *American Economic Review*, Vol. 93, No. 2, p. 180-185.
95. Christiansen, R. (2002), *The Victorian Visitors: Culture-Shock in Nineteenth-Century Britain*, Grove Press.
96. Church, J. & R. Ware (2000), *Industrial Organization: A Strategic Approach*, New York: McGraw-Hill.
97. Citters, E. van & J.C.A. van Roosendaal (1877-1944), *Verzameling van wetten, besluiten, enz. betreffende de spoorwegen in Nederland: 1832-1943*, 's Gravenhage: Van Cleef.
98. Clapp, E.J. (1911), *The Navigable Rhine*, Boston and New York: Houghton Mifflin.
99. Clark, R. & B. Clarke (2000), *Contract Cases and Materials*, Dublin: Gill & MacMillan.
100. Clark, R. (1998), *Contract Law in Ireland*, Dublin: Round Hall Sweet & Maxwell.
101. Coase, R.H. (1937), 'The Nature of the Firm', *Economica*, New Series, Vol. 4, No. 16, p. 386-405.
102. Coase, R.H. (1959), 'The Federal Communications Commission', *Journal of Law and Economics*, Vol. 2, p. 1-40.
103. Coase, R.H. (1960), 'The Problem of Social Costs', *Journal of Law and Economics*, Vol. 3, p. 1-44.
104. Coase, R.H. (1964), 'Discussion', *American Economic Review*, Vol. 54, No. 3, p. 194-197.
105. Coase, R.H. (1974), 'The Lighthouse in Economics', *Journal of Law and Economics*, Vol. 17, No. 2, p. 357-376.
106. Coase, R.H. (1988), *The Firm, the Market, and the Law*, Chicago: University of Chicago Press.
107. Coase, R.H. (1991), *The Institutional Structure of Production*, Lecture to the memory of Alfred Nobel, December 9, 1991.
108. Coase, R.H. (2000), 'The Acquisition of Fisher Body by General Motors', *Journal of Law & Economics*, Vol. 43, No. 1, p. 15-31.
109. Coase, R.H. (2006), 'The Conduct of Economics: The Example of Fisher Body and General Motors', *Journal of Economics & Management Strategy*, Vol. 15, No. 2, p. 255-278.
110. Coleman, P. (2005), 'Anyone seen the Invisible Hand?', *Rail Professional*, October 2005, p. 14-15.
111. Commons, J.R. (1924), *Legal Foundations of Capitalism*, New Brunswick, NJ: Transaction Publishers.
112. Commons, J.R. (1934), *Institutional Economics: Its Place in Political Economy*, New York: MacMillan.
113. Congelton, R.D. (1988), 'Evaluating rent-seeking losses: Do the welfare gains of lobbyists count?', *Public Choice*, Vol. 56, p. 181-184.
114. Coto-Millán, P. & V. Inglada [eds.] (2007), *Essays on Transport Economics*, Heidelberg: Physica-Verlag.

115. Cour des Comptes (2008), 'Le réseau ferroviare: Une réforme inachevée, une stratégie incertaine', Rapport Public Thématique, dd. 15 April 2008.
116. Coursey, D.L., J.L. Hovis & W.D. Schulze (1987), 'The Disparity Between Willingness to Accept and Willingness to Pay Measures of Value', *Quarterly Journal of Economics*, Vol. 102, No. 3, p. 679-690.'
117. Couwenberg, O. (2003), 'Incomplete contracten: een rechtseconomische benadering', inaugural address, Rijksuniversiteit Groningen, 18 February 2003.
118. Cowell, F.A. (1986), *Microeconomic Principles*, New York: Philip Allan.
119. Cowen, T. & A. Tabarrok (1999), 'The Opportunity Costs of Rent Seeking', *Journal of Public Finance and Public Choice*, Vol. XVII, No. 2-3, p. 121-127.
120. Crafts, N., T.C. Mills & A. Mulatu (2007), 'Total Factor Productivity Growth on Britain's Railways, 1852-1912: A reappraisal of the evidence', *Explorations in Economic History*, Vol. 44, p. 608-634.
121. Crosby, E.U. (1905), 'Fire Prevention', *Annals of the American Academy of Political and Social Science*, Vol. 26, p. 224-238.
122. Crozier, P. (2001), 'Why British Rail Privatisation has failed', *Economic Notes*, No. 91.
123. Cuncannon, F. (2004), 'The Case for Specific Performance as the Primary Remedy for Breach of Contract in New Zealand', *Victoria University of Wellington Law Review*, Vol. 35, p. 657 c.a.
124. Curran, C. (1999), 'The Endowment Effect', in: B. Bouckaert & G. De Geest [eds.], *Encyclopedia of Law and Economics, Volume 1. The History and Methodology of Law and Economics*, Cheltenham: Edward Elgar.
125. Cuyvers, L. & W. Meeusen (1976), 'The Structure of Personal Influence of the Belgian Holding Companies', *European Economic Review*, Vol. 8, No. 1, p. 51-69.
126. Daft, R.L. (1998), *Organization Theory and Design*, Cincinnati: South-Western College Publishing.
127. Dagnino-Pastore, J.M. & P.E. Farina (1999), 'Transaction Costs in Argentina', Paper presented at ISNIE 1999.
128. Dahl, R.A. (1957), 'The Concept of Power', *Behavioral Science*, Vol. 2, No. 3, p. 201-215.
129. Dahlman, C.J. (1979), 'The Problem of Externality', *Journal of Law and Economics*, Vol. 22, No. 1, p. 141-162.
130. Damart, S. & B. Roy (2009), 'The uses of cost-benefit analysis in transportation decision-making in France', *Transport Policy*, Vol. 16, No. 4, p. 200-212.
131. Darwin, C. (1859), *On the Origin of Species by Means of Natural Selection, or the Preservation of favoured Races in the Struggle for Life*, London: John Murray.
132. David, P.A. (1985), 'Clio and the Economics of QWERTY', *American Economic Review*, Vol. 75, No. 2, p. 332-337.
133. Davies, G.T. (2002), *European Union Internal Market Law*, London: Cavendish Publishing.
134. Davies, G.T. (2006), 'Subsidiarity as a Method of Policy Centralisation', Hebrew University of Jerusalem, Faculty of Law, Research Paper No. 11-06, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=921454.
135. Davies, G.T. (2006), 'The Legal Framework of Regulatory Competition', available at: <http://ssrn.com/abstract=903138>.
136. Davies, G.T. (2008), 'A Time to Mourn – How I Learned to Stop Worrying and Quite Like the European Union', Inaugural Lecture as Professor of European Law at the Faculty of Law of the VU University Amsterdam on 26th June 2008, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1314696.

137. Dawkins, R. (1976), *The Selfish Gene*, Oxford: Oxford University Press.
138. Dawkins, R. (1986), *The Blind Watchmaker*, New York: W.W. Norton.
139. Dawkins, R. (2007), *The God Delusion*, London: Bantam Press.
140. De Economist (1861), 'De Spoorwegen in Nederland, België en Pruisen', *De Economist*, Vol. 11, No. 1, p. 232-237.
141. De Economist (1862), 'De Belgische Spoorwegen', *De Economist*, Vol. 11, No. 1, p. 451-461.
142. De Economist (1874), 'Wat de Fransche Spoorwegen aan de Fransche Schatkist hebben gekost', *De Economist*, Vol. 23, No. 1, p. 243-257.
143. De Economist (1882), 'De Spoorweg-Enquête', *De Economist*, Vol. 31, No. 2, p. 1107-1120.
144. De Economist (1892), 'Geldelijke uitkomsten van het beheer der Staatsspoorwegen in Pruisen', *De Economist*, Vol. 41, No. 1, p. 175-176.
145. De Korte, R.W. (1990), 'Minder overheid, een betere economie', in: *Het sociaal-economisch beleid in de tweede helft van de twintigste eeuw: Opstellen aangeboden aan prof. dr. F.W. Rutten*, Groningen: Wolters-Noordhoff.
146. Debande, O. & E. Monami (1998), 'La réglementation et l'organisation du transport de personnes par chemins de fer dan l'union européenne: entre diversité et convergence', *Annals of Public and Cooperative Economics*, Vol. 69, No. 3, p. 347-371.
147. Dempsey, P.S. (2003), 'Transportation: A legal history', *Transport Law Journal*, Vol. 30, p. 235-366.
148. Demsetz, H. (1968), 'Why Regulate Utilities?', *Journal of Law and Economics*, Vol. 11, No. 1, p. 55-65.
149. Den Tex, J. (1962), *Oldenbarnevelt, Vol. II: Oorlog 1588-1609*, Haarlem: Tjeenk-Willink.
150. Dennison, H.S. (1924), 'Who Can Hire Management?', *Bulletin of the Taylor Society*, Vol. IX, p. 101-110.
151. Department for Transport (2010), 'The Future of Rail Franchising', available at: <http://www.dft.gov.uk/pgr/rail/passenger/franchises/futureoffranchising/>.
152. Devys, J. (1910), *Les Chemins de Fer de l'État Belge*, Paris: Arthur Rousseau.
153. DeWitt, B.S. (1970), 'Quantum Mechanics and Reality', *Physics Today*, Vol. 23, No. 9, p. 30-40.
154. Dicey, A.V. (1885), *Introduction to the Study of the Law of the Constitution*, London: MacMillan.
155. Distin, K. (2005), *The Selfish Meme: A Critical Reassessment*, Cambridge: Cambridge University Press.
156. Dixit, A.K. & M. Olson (2000), 'Does voluntary participation undermine the Coase Theorem?', *Journal of Public Economics*, Vol. 76, p. 309-335.
157. Dixit, A.K. (2004), *Lawlessness and Economics: Alternative Modes of Governance*, Princeton: Princeton University Press.
158. Doedens, A. & L. Mulder (1989), *Een Spoor van Verandering: Nederland en 150 jaar spoorwegen (1839-1989)*, Baarn: Bosch & Keuning.
159. Doeringer, P.B. & M.J. Piore (1971), *Internal Labor Markets and Manpower Analysis*, Lexington: D.C. Heath & Co.
160. Dollery, B. & W.H. Leong (1998), 'Measuring the Transaction Sector in the Australian Economy, 1911-1991', *Australian Economic History Review*, Vol. 38, No. 3, p. 207-231.
161. Dowding, K.M. (1991), *Rational Choice and Political Power*, Aldershot (UK): Edward Elgar.
162. Downs, A. (1957), *An Economic Theory of Democracy*, New York: Harper & Row.

163. Drucker, P. (1954), *The Practice of Management*, New York: HarperCollins.
164. Dunlavy, C.A. (1994), *Politics and Industrialization: Early Railroads in the United States and Prussia*, Princeton (NJ): Princeton University Press.
165. Easterbrook, F.H. (1984), 'The Limits of Antitrust', *University of Texas Law Review*, Vol. 63, p. 1 c.a.
166. Economides, N. & F. Flyer (1998), 'Compatibility and Market Structure for Network Goods', NYU Discussion Paper EC-98-02, available at: <http://www.stern.nyu.edu/networks/98-02.pdf>.
167. Economides, N. (1996), 'The Economics of Networks', *International Journal of Industrial Organization*, Vol. 14, p. 675-699, also available at: http://www.stern.nyu.edu/networks/Economides_Economics_of_Networks.pdf.
168. Economides, N. (2006), 'Competition Policy in Network Industries: An introduction', in: D. Jansen [ed.], *The New Economy and Beyond: Past, Present and Future*, Aldershot: Edward Elgar, also available at: http://www.stern.nyu.edu/networks/Economides_Competition_Policy.pdf.
169. Eggleston, K., E.A. Posner & R. Zeckhauser (2000), 'The Design and Interpretation of Contracts: Why Complexity Matters', *Northwestern University Law Review*, vol. 95, p. 91-132.
170. Ehrlich, I. & R.A. Posner (1974), 'An Economic Analysis of Legal Rulemaking', *Journal of Legal Studies*, Vol. 3, No. 1, p. 257-286.
171. Eijgelshoven, P.J., A. Nentjes, B.C.J. van Velthoven & L.J. van Gemerden (2000), *Markten en Overheid*, 3rd edition, Groningen: Wolters-Noordhoff.
172. Eisenhardt, K.M. (1985), 'Control: Organizational and Economic Approaches', *Management Science*, Vol. 31, No. 2, p. 134-149.
173. Eldredge, N. & S.J. Gould (1972), 'Punctuated equilibria: an alternative to phyletic gradualism', in: T.J.M. Schopf [ed.], *Models in Paleobiology*. San Francisco: Freeman Cooper. pp. 82-115, available online at: <http://www.blackwellpublishing.com/ridley/classic texts/eldredge.pdf>.
174. Ellison, G. & S. Fisher Ellison (2005), 'Lessons About Markets from the Internet', *Journal of Economic Perspectives*, Vol. 19, No. 2, p. 139-158.
175. Eloffson, J. (1996), 'The Dilemma of Changed Circumstances in Contract Law: An Economic Analysis of the Forseeability and Superior Risk Bearer Tests', *Columbia Journal of Law and Social Problems*, Vol. 30, p. 1-39.
176. Elster, J. (2007), *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*, Cambridge: Cambridge University Press.
177. Engartner, T. (2008), *Die Privatisierung der deutschen Bahn: Über die Implementierung marktorientierter Verkehrspolitik*, Wiesbaden: Verlag für Sozialwissenschaften.
178. Eppink, D.-J. (2007), *Life of a European Mandarin: Inside the Commission*, Tielt (Belgium): Lannoo.
179. Etzioni, A. (1967), 'Mixed-Scanning: A "Third" Approach to Decision-Making', *Public Administration Review*, Vol. 27, No. 5, p. 385-392.
180. Evans, A.W. (2007), 'Rail Safety and Rail Privatisation in Britain', *Accident Analysis and Prevention*, Vol. 39, No. 3, p. 510-523.
181. Evans, E.J. (1994), *The Great Reform Act of 1832* [2nd ed.], London: Routledge.
182. Everett, H. (1957), "'Relative State" Formulation of Quantum Mechanics', *Reviews of Modern Physics*, Vol. 29, No. 3, p.454-462.
183. Faber, J.A. [ed.] (1989), *Het Spoor: 150 jaar spoorwegen in Nederland*, Amsterdam: Meulenhoff.
184. Fama, E.F. & M.C. Jensen (1983), 'Separation of Ownership and Control', *Journal of Law and Economics*, Vol. 26, No. 2, p. 301-325.

185. Fehr, E., G. Kirchsteiger & A. Riedl (1993), 'Does Fairness Prevent Market Clearing? An Experimental Investigation', *Quarterly Journal of Economics*, Vol. 108, No. 2, p. 437-459.
186. Fisher, R., W. Ury & B. Patton (1991), *Getting to Yes: Negotiating an Agreement Without Giving In*, London: Random House.
187. Forde, M. (1999), *Company Law* (3rd edition), Dublin: Round Hall, Sweet & Maxwell.
188. Fox, J.W. (2003), 'Relational Contract Theory and Democratic Citizenship', *Case Western Reserve Law Review*, Vol. 54, No. 1, p. 1-67.
189. Francioni, F. (2008), 'Private Military Contractors and International Law: An Introduction', *European Journal of International Law*, Vol. 19, No. 5, p. 961-964.
190. Francis, J. (1851), *A History of the English Railway; Its Social Relations and Revelations, 1820-1845, Vol. II*, London: Longman, Brown, Green & Longmans.
191. Freeman, M., R. Pearson & J. Taylor (2007), 'Different and Better?' Scottish Joint-Stock Companies and the Law, c. 1720-1845', *English Historical Review*, Vol. CXXII, No. 495, p. 61-81.
192. Friebel, G., M. Ivaldi & C. Vibes (2010), 'Railway (De)Regulation: A European Efficiency Comparison', *Economica*, Vol. 77, No. 305, p. 77-91f.
193. Friedman, M. (1953), *Essays in Positive Economics*, Chicago: University of Chicago Press.
194. Friedman, M. (1999), 'Policy Forum: The Business Community's Suicidal Impulse', *Cato Policy Report*, Vol. 21, No. 2, p.6-7.
195. Fritschy, W. (1983), 'Spoorwegaanleg in Nederland van 1831 tot 1845 en de Rol van de Staat', *Economisch- en Sociaal-Historisch Jaarboek*, Vol. 46, chapter XII, p. 180-227.
196. Fudenberg, D. & E. Maskin (1990), 'Evolution and Cooperation in Noisy Repeated Games', *American Economic Review*, Vol. 80, No. 2, p. 274-279.
197. Fyffe, C.A. (1895), *History of Modern Europe: 1792-1878*, available at: www.gutenberg.net.
198. Gall, L. & M. Pohl [eds.] (1999), *Die Eisenbahnen in Deutschland: Von den Anfängen bis zur Gegenwart*, München: Beck.
199. Gazendam, H.W.M., & V.M.F.Homburg. (1999). 'Efficiëntie en verzelfstandiging: Economische en politieke efficiëntie als verklaring voor verzelfstandigingen', *Bestuurskunde*, Vol. 8, No. 1, p. 19-27.
200. Geerdink, G.C. & P.J. Stauvermann (2006), 'The Influence of Different Institutional Settings on Welfare in an Open Economy', paper presented at the MMGD conference in Faro, September 2006.
201. Geerdink, G.C. & P.J. Stauvermann (2007), 'Coase, Externalities, Property Rights and the Legal System', in: A. Prinz, A.E. Steenge & J. Schmidt [eds.], *The Rules of the Game: Institutions, Law and Economics*, Münster: LIT Verlag.
202. Geerdink, G.C. & P.J. Stauvermann (2009), 'Is Competition Between Regions Welfare Enhancing?', forthcoming.
203. Geest, G.G.A. de, 'Hoe maken we van de rechtswetenschap een volledige wetenschap', *Nederlands Juristenblad*, Vol. 2004, No. 2, p. 58-66.
204. Gemelli, C. & P. Royer (1864), *Révolution Belge de 1830 et Histoire de la Belgique jusqu'aujourd'hui*, Bruxelles & Oostende: Ferdinand Claassen.
205. Genugten, M.L. van (2008), *The Art of Alignment: Transaction Cost Economics and the Provision of Public Services at the Local Level*, Ph.D. Thesis Universiteit Twente, Enschede: Ipskamp.
206. Gepken-Jager, E., G. van Solinge & L. Timmerman [eds.] (2005), *VOC 1602-2002: 400 Years of Company Law*, The Hague: Kluwer.

207. Gibbon, E. (1776), *The Decline and Fall of the Roman Empire, Volume 1*, London: Random House.
208. Gindis, D. (2009), 'From fictions and aggregates to real entities in the theory of the firm', *Journal of Institutional Economics*, Vol. 5, No. 1, p. 25-46.
209. Gintis, H. (2000), *Game Theory Evolving: Reason, Dynamics, and Behavior*, Princeton: Princeton University Press.
210. Gintis, H. (2007), *Game Theory for Behavioral Scientists*, Princeton: Princeton University Press.
211. Glancey, J. (2005), 'The Beauty of Crewe: The home of steam-age locomotives can teach us something about how to build modern towns', *The Guardian*, 6 December 2005, available at: <http://www.guardian.co.uk/society/2005/dec/06/comment.columnists>.
212. Goldberg, V. (2007), 'Lawyers Asleep at the Wheel? The GM-Fisher Body Contract', Columbia University School of Law, Center for Law and Economic Studies Working Paper No. 316, available at http://papers.ssrn.com/paper.taf?abstract_id=1010982.
213. González-Benito, J., I. Suárez-González & M. Spring (2000), 'Complementarities between JIT purchasing practices: An economic analysis based on transaction costs', *International Journal of Production Economics*, Vol. 67, p. 279-293.
214. Gorbachev, M.S. (1987), *Perestrojka: Blauwdruk voor het gemeenschappelijke Europese Huis*, Utrecht: Het Spectrum. (*O perestrojke i novom mysjlenii*, translated by G.G. Baardman, et. al.)
215. Gordon, W.J. (1910), *Our Home Railways: How They Began and How They are Worked*, London: Frederick Warne & Co.
216. Goudriaan, J. (1939), *De toekomst der Nederlandsche Spoorwegen*, Rotterdam: Nijgh & Van Ditmar.
217. Goujon, S. (2004a), 'Le Transport Ferroviaire en Europe: Libéralisation et part modale du fer', Notes de synthèse du SES No. 151, Janvier-Février 2004.
218. Goujon, S. (2004b), 'La Réforme du Secteur Ferroviaire en Grande-Bretagne', Notes de synthèse du SES No. 154, Juillet-Août 2004.
219. Gravendaal, C.J.W. (1990), *De Rijnspoorweg (1838 – 1845) tussen oprichting en overdracht. Een analyse van het gevoerde beleid*, diss. Leiden.
220. Greif, A. (1993), 'Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition', *American Economic Review*, Vol. 83, No. 3, p. 525-548.
221. Greif, A. (2008), 'Coercion and Exchange: How did Markets Evolve?', available at: <http://ssrn.com/abstract=1304204>.
222. Greve, C., M. Flinders & S. van Thiel (1999), 'Quangos – What's in a Name? Defining Quangos from a Comparative Perspective', *Governance*, Vol. 12, No. 2, p. 129-146.
223. Groenendijk, N.S. (1998), *Overheidsinkomsten*, Ph.D. Thesis Universiteit Twente, Enschede: Twente University Press.
224. Grossman, S.J. & O.D. Hart (1986), 'The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration', *Journal of Political Economy*, Vol. 94, No. 4, p. 691-719.
225. Grothe, D. (1859), 'Spoorwegen in de Oude en in de Nieuwe Wereld', *De Economist*, Vol. 8, No. 1, p. 393-399.
226. Hadfield, G.K. & B.R. Weingast (2010), 'What is Law? A Coordination Model of the Characteristics of Legal Order', USC Center in Law, Economics and Organization Research Paper No. C10-17, available at: <http://ssrn.com/abstract=1707083>.
227. Hamilton Ellis, C. (1955), *The North British Railway*, London: I. Allen.

228. Hamilton Ellis, C. (1959), *British Railway History: An outline from the accession of William IV to the nationalisation of the railways, Vol. 2: 1877-1947*, London: George Allen & Unwin.
229. Hands, D.W. (2009), 'Effective Tension in Robbins' Economic Methodology', *Economica*, Vol. 76, No. s1, p. 831-844.
230. Hanson, S.E. (2009), 'The Founding of the French Third Republic', available at: <http://ssrn.com/abstract=1452236>.
231. Harris, R. (1994), 'The Bubble Act: Its Passage and Its Effects on Business Organization', *Journal of Economic History*, Vol. 54, No. 3, p. 610-627.
232. Hart, H.L.A. (1961), *The Concept of Law*, Oxford: Clarendon Press.
233. Hart, H.L.A. (1972), 'Bentham on Legal Powers', *Yale Law Journal*, Vol. 81, No. 5, p. 799-822.
234. Hart, O.D. & B. Holmstrom (2008), 'A Theory of Firm Scope', NBER Working Paper No. 14613, December 2008.
235. Hart, O.D. & J. Moore (1988), 'Incomplete Contracts and Renegotiation', *Econometrica*, Vol. 56, No. 4, p. 755-785.
236. Hart, O.D. & J. Moore (1990), 'Property Rights and the Nature of the Firm', *Journal of Political Economy*, Vol. 98, No. 6, p. 1119-1158.
237. Hart, O.D. & J. Moore (1999), 'Foundations of Incomplete Contracts', *Review of Economic Studies*, Vol. 66, p. 115-138.
238. Hart, O.D. (1995): *Firms, Contracts, and Financial Structure*, Oxford: Clarendon Press.
239. Hart, O.D., A. Shleifer & R.W. Vishny (1997), 'The Proper Scope of Government: Theory and an Application to Prisons', *Quarterly Journal of Economics*, Vol. 112, p. 1127-1161.
240. Harvey, D. (2005), *A Brief History of Neoliberalism*, Oxford: Oxford University Press.
241. Hauteclouque, A. de & Y. Perez (2011), 'Law & Economics Perspectives on Electricity Regulation', EUI Working Paper RSCAS 2011/21.
242. Hauwe, L. van den (1999), 'Public Choice, Constitutional Political Economy and Law and Economics', in: B. Bouckaert & G. De Geest [eds.], *Encyclopedia of Law and Economics, Volume 1. The History and Methodology of Law and Economics*, Cheltenham: Edward Elgar.
243. Hayek, F.A. (1944), *The Road to Serfdom*, London: Routledge.
244. Hayek, F.A. (1945), 'The Use of Knowledge in Society', *American Economic Review*, Vol. 35, No. 4, p. 519-530.
245. Hayek, F.A. (1960), *The Constitution of Liberty*, London: Routledge.
246. Head, F.B. (1861), *Stokers and Pokers: or the London and North-Western Railway, the Electric Telegraph, and the Railway Clearing House*, London: John Murray.
247. Heide, J.B. & G. John (1990), 'Alliances in Industrial Purchasing: The Determinants of Joint Action in Buyer-Supplier Relationships', *Journal of Marketing Research*, Vol. 27, No. 1, p. 24-36.
248. Heijdra, B.J. & R. van der Ploeg (2000), *Foundations of Modern Macroeconomics*, mimeo, Rijksuniversiteit Groningen.
249. Helfer, L.R. & G.B. Dinwoodie (2001), 'Designing non-national systems: The case of the uniform domain name dispute resolution policy', *William & Mary Law Review*, Vol. 43, p. 141-274.
250. Hendrikse, G. (2003), *Economics and Management of Organizations: Co-ordination, Motivation and Strategy*, New York: McGraw Hill.
251. Henket, M. (1992), 'On the Logical Analysis of Judicial Decisions', *International Journal for the Semiotics of Law*, Vol. V, No. 14, p. 153-164.

252. Hennessey, P. (1987), 'The Attlee Governments, 1945-1951', in: P. Hennessey & A. Seldon, *Ruling Performance: British Governments from Attlee to Thatcher*, Oxford: Blackwell, p. 28-62.
253. Herten, B. van der (2004), *België onder Stoom: Transport en Communicatie tijdens de 19de eeuw*, Leuven: Leuven University Press.
254. Hicks, J.R. (1939), 'The Foundations of Welfare Economics', *The Economic Journal*, Vol. 49, No. 196, p. 696-712.
255. Hill, L.B. [ed.] (1992), *The State of Public Bureaucracy*, M.E. Sharpe Series on Bureaucracies, Public Administration, and Public Policy, Armonk (NY): M.E. Sharpe.
256. Hilton, B. (2006), *A Mad, Bad, and Dangerous People? England 1783-1846*, Oxford: Oxford University Press.
257. Hirschleifer, J. (1988), 'The Analytics of Continuing Conflict', *Synthese*, Vol. 76, p. 201-233.
258. Hirschleifer, J. (1989), 'Conflict and rent-seeking success functions: Ratio vs. difference models of relative success', *Public Choice*, Vol. 63, p. 101-112.
259. Hirschman, A.O. (1970), *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*, Cambridge: Harvard University Press.
260. Hix, S. (2008), *What's Wrong with the European Union and How To Fix It*, Cambridge (UK): Polity Press.
261. Hobbes, T. (1651), *Leviathan*, Oxford: Oxford University Press.
262. Hobson, J.M. (1997), *The Wealth of States: A comparative sociology of international economic and political change*, Cambridge: Cambridge University Press.
263. Hodgson, G.M. (2002), 'The Legal Nature of the Firm and the Myth of the Firm-Market Hybrid', *International Journal of the Economics of Business*, Vol. 9, No. 1, p. 37-60.
264. Hoffmann, S. (1966), 'Obstinate or Obsolete? The Fate of the Nation-State and the Case of Western Europe', *Daedalus*, Vol. 95, No. 3, p. 862-915.
265. Hohfeld, W.N. (1913), 'Some Fundamental Legal Conceptions as Applied in Judicial Reasoning', *Yale Law Journal*, Vol. 23, No. 1, p. 16-59.
266. Holmström, B. (1979), 'Moral Hazard and Observability', *Bell Journal of Economics*, Vol. 10, No. 1, p. 74-91.
267. Holmstrom, B. (1999), 'The Firm as a Subeconomy', *Journal of Law, Economics & Organization*, Vol. 15, No. 1, p. 74-102.
268. Holterman, M.W. (2007), 'Transaction Cost Economics in International Relations: The Case of International Antitrust Enforcement', Master's Thesis, University of Groningen, available at ssrn.com/abstract=960428.
269. Holterman, M.W. (2009), 'The Introduction of Rail Transport in the Prussian Rhine Province: A Transaction Cost Economic History', in: A. Prinz, A.E. Steenge & N. Isegrei (Eds.), *New technologies, networks and governance structures* (Wirtschaft: Forschung und Wissenschaft, 24), Münster, Germany: LIT Verlag, p. 102-121, available at: <http://ssrn.com/abstract=1091151>.
270. Hood, C. (1991), 'A Public Management For All Seasons?', *Public Administration*, Vol. 69, No.1, p. 3-19.
271. Horn, N. (1979), 'Aktienrechtliche Unternehmensorganisation in der Hochindustrialisierung (1860-1920)', in: N. Horn & J. Kocka [eds.], *Recht und Entwicklung der Großunternehmen im 19. und frühen 20. Jahrhundert: Wirtschafts-, sozial- und rechtshistorische Untersuchungen zur Industrialisierung in Deutschland, Frankreich, England und den USA*, Göttingen: Vandenhoeck & Ruprecht.

272. Houtman-De Smedt, H. (1994), 'The banking system in Belgium through the centuries', in: M. Pohl & S. Freitag [eds.], *Handbook on the History of European Banks*, Aldershot: Edward Elgar.
273. Hovenkamp, H. (1992), 'Rationality in Law & Economics', *George Washington Law Review*, Vol. 60, p. 293-338.
274. Hsiung, B. (1999), 'Sailing Towards the Brave New World of Zero Transaction Costs', *European Journal of Law and Economics*, vol. 8, p. 153-169.
275. Hylton, S. (2007), *The Grand Experiment: The Birth of the Railway Age 1820-45*, Hersham, UK: Ian Allan.
276. Iliasu, A.A. (1971), 'The Cobden-Chevalier Commercial Treaty of 1860', *Historical Journal*, Vol. 14, No. 1, p. 67-98.
277. Imhof, L.A., D. Fudenberg & M.A. Nowak (2005), 'Evolutionary Cycles of Cooperation and Defection', *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 102, No. 31, p. 10797-10800.
278. Israel, J.I. [ed.] (1991), *The Anglo-Dutch Moment: Essays on the Glorious Revolution and its World Impact*, Cambridge: Cambridge University Press.
279. Jansen, H. (2005), 'E Contrario Reasoning: The Dilemma of the Silent Legislator', *Argumentation*, Vol. 19, p. 485-496.
280. Janssen, B. (2009), *Opportunisme op het spoor: Analyse van transacties tussen NS en ProRail*, Masterthesis University of Twente.
281. Jeanne, O., J.D. Ostry & J. Zettelmeyer (2008), 'A Theory of International Crisis Lending and IMF Conditionality', IMF Working Paper 08/236, available at: www.imf.org/external/pubs/ft/wp/2008/wp08236.pdf.
282. Jensen, M.C. & W.H. Meckling (1976), 'Theory of the firm: Managerial Behavior, Agency Costs and Ownership Structure', *Journal of Financial Economics*, Vol. 3, p. 305-360.
283. Jonckers Nieboer, J.H. (1907), *Geschiedenis der Nederlansche Spoorwegen*, Haarlem: H.D. Tjeenk Willink & Zoon.
284. Jonckers Nieboer, J.H. (1938), *Geschiedenis der Nederlansche Spoorwegen: 1832-1938*, Rotterdam: Van Ditmar.
285. Jones, A. & B. Sufrin (2008), *EC Competition Law: Text, Cases, and Materials (3rd edition)*, Oxford: Oxford University Press.
286. Joskow, P. (1985), 'Vertical Integration and Long-Term Contracts: The Case of Coal-Burning Electric Generating Plants', *Journal of Law, Economics, & Organization*, Vol. 1, No. 1, p. 33-80.
287. Joskow, P. (1987), 'Contract Duration and Relationship-Specific Investments: Empirical Evidence from Coal Markets', *American Economic Review*, Vol. 77, No. 1, p. 168-185.
288. Joskow, P. (1988), 'Asset Specificity and the Structure of Vertical Relationships: Empirical Evidence', *Journal of Law, Economics, & Organization*, Vol. 4, No. 1, p. 95-117.
289. Judt, T. (2009), 'What is Living and What is Dead in Social Democracy?', *New York Review of Books*, Vol. 56, No. 20, December 17, 2009.
290. Jupe, R. (2009), 'New Labour, Public-Private Partnerships and Rail Transport Policy', *Economic Affairs*, Vol. 29, No. 1, p. 20-25.
291. Kahneman, D. & A. Tversky (1979), 'Prospect Theory: An Analysis of Decision under Risk', *Econometrica*, Vol. 47, p. 263-291.
292. Kaldor, N. (1939), 'Welfare Propositions of Economics and Interpersonal Comparisons of Utility', *The Economic Journal*, Vol. 49, No. 195, p. 549-552.

293. Kalmbacher, J. (1998), 'Le rôle de l'état dans le développement et l'achèvement du réseau ferré', *Revue générale des chemins de fer*, 1998, No. 7-8, p. 17-24.
294. Kant, I. (1781), *Kritik der reinen Vernunft*.
295. Kaplan, R.D. (2007), 'Outsourcing Conflict', *The Atlantic Monthly*, September 2007.
296. Kaptein, H. (1993), 'E Contrario Arguments in Law: From Interpretation to Implicit Premises', *International Journal for the Semiotics of Law*, Vol. VI, No. 18, p. 315-324.
297. Kapteyn, P.J.G. & P. VerLoren van Themaat (2008), *The Law of the European Union and the European Communities*, 4th Revised Edition, edited by P.J.G. Kapteyn, A.M. McDonnell, K.J.M. Mortelmans & C.W.A. Timmermans, Alphen a/d Rijn (NL): Wolters Kluwer.
298. Karni, E., 'Subjective Expected Utility Theory without States of the World', on-line working paper, <http://www.econ.jhu.edu/people/karni/seurr.pdf>. (Last accessed on May 7, 2008.)
299. Kennedy, E. (1979), "'Ideology" from Destutt De Tracy to Marx', *Journal of the History of Ideas*, Vol. 40, No. 3, p. 353-368.
300. Kennedy, P.M. (1987), *The Rise and Fall of the Great Powers: Economic Change and Military Conflict From 1500 to 2000*, New York: Random House.
301. Kerkwijk, J.J. van (1878), 'Een en ander over onze Staatsspoorwegen', *De Economist*, Vol. 27, No. 1, p. 546-562.
302. Kerkwijk, J.J. van (1886), 'Een woord over de exploitatie der Staatsspoorwegen', *De Economist*, Vol. 35, No. 1, p. 381-390.
303. Keynes, J.M. (1936), *The General Theory of Employment, Interest, and Money*, Basingstoke (UK): Palgrave Macmillan.
304. Klein, B. (1983), 'Contracting Costs and Residual Claims: The Separation of Ownership and Control', *Journal of Law and Economics*, Vol. 26, No. 2, p. 367-374.
305. Klein, B. (1988), 'Vertical Integration as Organizational Ownership: The Fisher Body-General Motors Relationship Revisited', *Journal of Law, Economics, & Organization*, Vol. 4, No. 1, p. 199-213.
306. Klein, B., A. Crawford & A.A. Alchian (1978), 'Vertical integration, appropriable rents, and the competitive contracting process', *Journal of Law and Economics*, Vol. 21, p. 297-326.
307. Klein, P.G. (1999), 'New Institutional Economics', in: B. Bouckaert & G. De Geest [eds.], *Encyclopedia of Law and Economics, Volume 1. The History and Methodology of Law and Economics*, Cheltenham: Edward Elgar.
308. Kleinbaum, D.G. et. al. (2008), *Applied Regression Analysis and Other Multivariable Methods*, Belmont, CA: Thomson.
309. Knight, F.H. (1921), *Risk, Uncertainty, and Profit*, Chicago: University of Chicago Press.
310. Knight, F.H. (1940), "'What is Truth" in Economics?', *Journal of Political Economy*, Vol. 48, No. 1, p. 1-32.
311. Knill, C. & D. Lehmkuhl (1998), 'An Alternative Route of Legal Integration: The Community's Railways Policy', *European Integration Online Papers*, Vol. 2, No. 3.
312. Knowles, L.C.A. (2006), *Economic Development in the Nineteenth Century: France, Germany, Russia and the United States*, London: Routledge.
313. KPMG (2010), 'Rail Franchising Policy: Analysis of Historic Data', available at <http://www.dft.gov.uk/pgr/rail/passenger/franchises/historicaldata/>.
314. Krockow, C. Graf von (1990), *Die Deutschen in ihrem Jahrhundert: 1890-1990*, Reinbek bei Hamburg: Rowohlt.

315. Kronman, A.T. (1983), 'Paternalism and the Law of Contracts', *Yale Law Journal*, Vol. 100, No. 5, p. 763-798.
316. Krueger, A.O. (1974), 'The Political Economy of the Rent-Seeking Society', *American Economic Review*, Vol. 64, No. 3, p. 291-303.
317. Krugman, P. (1994), 'Competitiveness: A Dangerous Obsession', *Foreign Affairs*, Vol. 73, No. 2, p. 28-44.
318. Krugman, P. (1996), *Pop Internationalism*, Cambridge, Mass: MIT Press.
319. Künneke, R.W., J. Groenewegen & J.-F. Auger [eds.] (2009), *The Governance of Network Industries: Institutions, Technology and Policy in Reregulated Infrastructures*, Cheltenham (UK): Edward Elgar.
320. Lamarck, J.B. (1809), *Philosophie Zoologique*.
321. Lamming, C. (1998), 'L'Année 1878: Importante pour le Chemin de Fer Français', *Revue générale des chemins de fer*, 1998, No. 7-8, p. 9-15.
322. Landes, W.M. & R.A. Posner (1976), 'Legal Precedent: A Theoretical and Empirical Analysis', *Journal of Law & Economics*, Vol. 19, No. 2, p. 249-307.
323. Landes, W.M. & R.A. Posner (1981), 'Market Power in Antitrust Cases', *Harvard Law Review*, Vol. 94, p. 937-996.
324. Laperrouza, M. & M. Finger (2009), 'Regulating Europe's single railway market: Integrating performance and governance', Second Annual Conference on Competition and Regulation in Network Industries, 20 November 2009.
325. Laperrouza, M. (2008), 'Does liberalisation increase systemic risk in the railway sector?', First Annual Conference on Competition and Regulation in Network Industries, 28 November 2008.
326. Lardner, D. (1850), *Railway Economy: A Treatise on the New Art of Transport, its Management, Prospects, and Relations, Commercial, Financial, and Social, with an exposition of the practical results of the railways in operation in the United Kingdom, on the continent, and in America*, London: Taylor, Walton, and Maberly.
327. Laveleye, A. de (1862), 'Histoire des vingt-cinq premières années des chemins de fer belges', Brussels: A. Decq, reproduced in: Linters, A. [ed.] (1985), *Spoorwegen in België/Chemins de Fer en Belgique/Railways in Belgium*, Gent: Vlaamse Vereniging voor Industriële Archeologie, p. 79 c.a.
328. Lazear, E.P. (2000), 'Economic Imperialism', *Quarterly Journal of Economics*, Vol. 115, No. 1, p. 99-146.
329. Lehmann, H. (1998), *Eisenbahnen in Preußen*, Düsseldorf: Alba.
330. Lehmann, S.H. (2009), 'The German elections in the 1870s: why Germany turned from liberalism to protectionism', Preprints of the Max Planck Institute for Research on Collective Goods, Bonn, No. 2009/34, also available at: <http://ssrn.com/abstract=1500308>.
331. Lehning, J.R. (2001), *To Be a Citizen: The Political Culture of the Early French Third Republic*, Ithaca (NY): Cornell University Press.
332. Leontief, W. (1971), 'Theoretical Assumptions and Nonobserved Facts', *American Economic Review*, Vol. 61, No. 1, p. 1-7.
333. LeRoy Miller, R. & D.D. VanHoose (2004), *Macroeconomics: Theories, Policies, and International Applications*, Belmont, CA: Thomson.
334. Lessig, L. (2001), 'Copyright's First Amendment', *UCLA Law Review*, Vol. 48, p. 1057-1073.
335. Lester, R.A. (1946), 'Shortcomings of Marginal Analysis for Wage-Employment Problems', *American Economic Review*, Vol. 36, No. 1, p. 63-82.
336. Liberman, M. (2006), 'Noclone', blog post dated March 05, 2006, on the Language Log blog, <http://itre.cis.upenn.edu/~myl/languageblog/archives/002898.html>, last visited on July 16, 2008.

337. Liebowitz, S.J. & S.E. Margolis (1990), 'The Fable of the Keys', *Journal of Law and Economics*, Vol. 33, No. 1, p. 1-25.
338. Liebowitz, S.J. & S.E. Margolis (1995), 'Path Dependence, Lock-in and History', *Journal of Law, Economics, and Organization*, Vol. 11, No. 1, p. 205-226.
339. Liebowitz, S.J. & S.E. Margolis (1999), 'Path Dependence', in: B. Bouckaert & G. De Geest [eds.], *Encyclopedia of Law and Economics, Volume 1. The History and Methodology of Law and Economics*, Cheltenham: Edward Elgar.
340. Liebowitz, S.J. & S.E. Margolis (2000), 'Path Dependence', in: B. Bouckaert & J. Ottoson [eds.], *Encyclopedia of Law and Economics*, Cheltenham: Edward Elgar, p. 981-998.
341. Liebowitz, S.J. & S.E. Margolis (2010), 'How the Lock-In Movement Went off the Tracks', available at: <http://ssrn.com/abstract=1698486>.
342. Linters, A. [ed.] (1985), *Spoorwegen in België/Chemins de Fer en Belgique/Railways in Belgium*, Gent: Vlaamse Vereniging voor Industriële Archeologie.
343. Lipsey, R.G. (2001), 'Successes and failures in the transformation of economics', *Journal of Economic Methodology*, Vol. 8, No. 2, p. 169-201.
344. Lipsey, R.G. (2009), 'Some Legacies of Robbins' *An Essay on the Nature and Significance of Economic Science*', *Economica*, Vol. 76, No. s1, p. 845-856.
345. Locke, J. (1689), *Second Treatise on Government*, Cambridge: Cambridge University Press.
346. Lodge, M. (2002), 'The Wrong Type of Regulation? Regulatory Failure and the Railways in Britain and Germany', *Journal of Public Policy*, Vol. 22, No. 3, p. 271-297.
347. Lodge, M. (2003), 'Institutional Choice and Policy Transfer: Reforming British and German Railway Regulation', *Governance: An International Journal of Policy, Administration, and Institutions*, Vol. 16, No. 2, p. 159-178.
348. Louis, F. & A. Vallery (2004), 'Ferring Revisited: the Altmark Case and State Financing of Public Service Obligations', *World Competition*, Vol. 27, No. 1, p. 53-74.
349. Lowe, J. (2007), 'What Would Grotius Do? Methods and Implications of Incorporating the Contract Law Doctrine of Illusory Promises into the Law of Treaty Interpretation', *Washington University Global Studies Law Review*, Vol. 6, p. 703-722.
350. Luger, K. (2008), *Chinese Railways: Reform and Efficiency Improvement Opportunities*, Heidelberg: Physica-Verlag.
351. Lülfsmann, C. (2007), 'On the virtues of privatization when government is benevolent', *Journal of Economic Behavior & Organization*, Vol. 64, No. 1, p. 17-34.
352. Macaulay, S. (1963), 'Non-contractual relations in business', *American Sociological Review*, Vol. 28, No. 1, p. 55-67.
353. MacNeil, I.R. (1978), 'Contracts: Adjustments of Long-Term Economic Relations under Classical, Neoclassical, and Relational Contract Law', *Northwestern University Law Review*, Vol. 72, p. 854-906.
354. Mahaim, E. (1906), 'Cheap Railway Tickets for Workmen in Belgium', *The Economic Journal*, Vol. 16, No. 64, p. 536-545.
355. Mahoney, P.G. & C.W. Sanchirico (2005), 'General and Specific Legal Rules', *Journal of Institutional and Theoretical Economics*, Vol. 161, p. 329-346.
356. Manne, H.G. (1965), 'Mergers and the Market for Corporate Control', *Journal of Political Economy*, Vol. 73, No. 2, p. 110-120.
357. Marsden, B. (2002), *Watt's Perfect Engine: Steam & the Age of Invention*, Royston (UK): Icon Books.
358. Marshall, A. (1890), *Principles of Economics*, London: MacMillan.
359. Martin, S. (1993), *Industrial Economics: Economic Analysis and Public Policy*, Upper Saddle River, NJ: Prentice Hall.

360. Marx, K. (1859), *Zur Kritik der Politischen Ökonomie*.
361. Marx, K. (1871), *Der Bürgerkrieg in Frankreich*.
362. Marx, K.H. (1867), *Das Kapital, Teil I, Der Produktionsprozess des Kapitals*, Berlin: Gustav Kiepenheuer.
363. Masini, F. (2009), 'Economics and Political Economy in Lionel Robbins's Writings', *Journal of the History of Economic Thought*, forthcoming.
364. Maskin, E. & J. Moore (1999), 'Implementation and Renegotiation', *Review of Economic Studies*, Vol. 66, No. 1, p. 39-56.
365. Maskin, E. & J. Tirole (1999), 'Unforeseen Contingencies and Incomplete Contracts', *Review of Economic Studies*, Vol. 66, No. 1, p. 83-114.
366. Mason, R. (2000), 'Network externalities and the Coase conjecture', *European Economic Review*, Vol. 44, No. 10, p. 1981-1992.
367. Masten, S.E., J.W. Meehan & E.A. Snyder (1991), 'The Costs of Organization', *Journal of Law, Economics, & Organization*, Vol. 7, No. 1, p. 1-25.
368. Mauer, J.G. (1996), 'Common Contractual Risk Allocations in International Power Projects', *Columbia Business Law Review*, Vol. 1996, p. 37-59.
369. May, T.E. (1863), *The Constitutional History of England since the Accession of George the Third* (2nd edition), London: Longman.
370. Mayor, A. (1998), 'Views of the Gardens at Marly: Louis XIV: Royal Gardener', *Magazine Antiques*, November 1998.
371. McGuire, M.C. & M. Olson (1996), 'The Economics of Autocracy and Majority Rule: The Invisible Hand and the Use of Force', *Journal of Economic Literature*, Vol. XXXIV, No. 1, p. 72-96.
372. McMahon, B. & W. Binchy (2000), *Law of Torts*, Dublin: LexisNexis.
373. Medema, S.G. & R.O. Zerbe (1999), 'The Coase Theorem', in: B. Bouckaert & G. De Geest [eds.], *Encyclopedia of Law and Economics, Volume 1. The History and Methodology of Law and Economics*, Cheltenham: Edward Elgar.
374. Medema, S.G. (2009), 'From Dismal to Dominance? Law and Economics, and the Values of Imperial Science', in: A. Hatzis [ed.], *Norms and Values in Law and Economics*, London: Routledge.
375. Ménard, C. & S. Saussier (2002), 'Contractual Choice and Performance: The Case of Water Supply in France', in: Brousseau, E. & J. Glachant [eds.], *The Economics of Contracts: Theories and Applications*, Cambridge: Cambridge University Press, p. 440-462.
376. Ménard, C. (1995), 'Markets as institutions versus organizations as markets? Disentangling some fundamental concepts', *Journal of Economic Behavior and Organization*, Vol. 28, p. 161-182.
377. Ménard, C. (1996), 'On Clusters, Hybrids, and Other Strange Forms: The Case of the French Poultry Industry', *Journal of Institutional and Theoretical Economics*, Vol. 152, No. 1, p. 154-196.
378. Menger, C. (1871), *Grundsätze der Volkswirtschaftslehre*, Leipzig: Freytag.
379. Merola, M. & C. Medina (2003), 'De l'arrêt Ferring à l'arrêt Altmark: Continuité ou revirement dans l'approche du financement des services publics', *Cahier de Droit Européen*, p. 639-694.
380. Milgrom, P. & J. Roberts (1992), *Economics, Organization, and Management*, Upper Saddle River, New Jersey: Prentice Hall.
381. Mises, L. von (1949), *Human Action: A Treatise on Economics*, 4th revised edition, 1996, San Francisco: Fox & Wilkes.
382. Mizutani, F. & K. Nakamura (2001), 'The Japanese Experience with Railway Restructuring', in: F. Mizutani & K. Nakamura, *Governance, Regulation, and Privatisation in the Asia-Pacific Region*, NBER East Asia Seminar on Economics, Vol. 12, p. 305-342.

383. Moravcsik, A. (1998), *The Choice for Europe: Social Purpose and State Power From Messina to Maastricht*, London: Routledge.
384. Mueller, D.C. (1979) *Public Choice*, Cambridge: Cambridge University Press.
385. Mueller, D.C. (2003), *Public Choice III*, Cambridge: Cambridge University Press.
386. Muirhead, J.P. (1858), *The Life of James Watt, with selections from his correspondence*, London: John Murray.
387. Murray, A. (2001), *Off the Rails: The Crisis on Britain's Railways*, London: Verso.
388. Myerson, R.B. (1983), 'Mechanism Design by an Informed Principal', *Econometrica*, Vol. 51, No. 6, p. 1767-1797.
389. Myerson, R.B. (1979), 'Incentive Compatibility and the Bargaining Problem', *Econometrica*, Vol. 47, p. 61-73.
390. Myerson, R.B. (1981), 'Optimal Auction Design', *Mathematics of Operations Research*, Vol. 6, No. 1, p. 58-73.
391. Myerson, R.B. (1982), 'Optimal Coordination Mechanisms in Generalized Principal-Agent Problems', *Journal of Mathematical Economics*, Vol. 11, No. 1, p. 67-81.
392. Myerson, R.B. (1991), *Game Theory: Analysis of Conflict*, Cambridge: Harvard University Press.
393. Na, H.-S. (2002), *Staatsvorstellung, Mentalität und ökonomisches Denken des deutschen Wirtschaftsbürgertums im 19. Jahrhundert – am Beispiel des Rheinisch-Westfälischen (Handels- und) Gewerbevereins*, Inaugural-Dissertation zur Erlangung des Grades des Dr.phil., Universität Siegen.
394. Nash, C.A. & C. Rivera-Trujillo (2004), 'Rail Regulatory Reform in Europe – principles and practice', Paper Presented at the STELLA Focus Group 5 synthesis meeting, Athens, June 2004.
395. Nash, C.A. (1991), 'Rail Policy in the European Community', Institute for Transport Studies Working Paper 324, January 1991.
396. Nash, C.A. (1999), 'Developments in Rail Policy in the EU', Paper Presented at the 6th international conference on competition and ownership in land passenger transport, Cape Town, South Africa, 19-23 September 1999.
397. Nash, C.A., B. Matthews & J. Shires (2004), 'The EU Transport Policy White Paper: An assessment of progress', Institute for Transport Studies, July 2004.
398. Nash, J.F. (1950), 'The Bargaining Problem', *Econometrica*, Vol. 18, No. 2, p. 155-162.
399. Nash, J.R. (2007), 'Economic Efficiency versus Public Choice: The Case of Property Rights in Road Traffic Management', University of Chicago Law School, John M. Olin Program in Law and Economics Working Paper No. 374, available at <http://ssrn.com/abstract=1071186>.
400. Naylor, T.H. (1988), *The Gorbachev Strategy: Opening the Closed Society*, Lexington, Mass: Lexington Books.
401. Neff, S.C. (1990), *Friends But No Allies: Economic Liberalism and the Law of Nations*, New York: Columbia University Press.
402. Neiertz, N. (1991), 'L'État et son réseau', *Revue d'histoire des chemins de fer*, Vol. 4, No. 1, p. 57-68.
403. Nentjes, A. (1993), *Elementaire Rechtseconomie*, Groningen: Stenfert Kroese.
404. Neste, F. van (1990), *Beginselen van Belgisch Privaatrecht V: Zakenrecht: Boek 1: Goederen, Bezit en Eigendom*, Brussel: Story-Scientia.
405. Neumann, J. von & O. Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.
406. Niblett, A., R.A. Posner & A. Shleifer (2008), 'The Evolution of a Legal Rule', NBER Working Paper No. 13856 (March 2008).

407. Nieuwerburgh, S. van, F. Buelens & L. Cuyvers (2005), 'Stock Market Development and Economic Growth in Belgium', NYU Working Paper No. FIN-05-024, available at: <http://ssrn.com/abstract=1294160>.
408. Niskanen, W.A. (1971), *Bureaucracy and representative government*, Chicago: Aldine-Atherton.
409. NMBS (2001), *75 jaar op het spoor van de toekomst: de nationale maatschappij der Belgische spoorwegen 1926-2001*, Brussel: NMBS.
410. North, D.C. & B.R. Weingast (1989), 'Constitutions and Commitment: The Evolution of Institutional Governing Public Choice in Seventeenth-Century England', *Journal of Economic History*, Vol. 49, No. 4, p. 803-832.
411. North, D.C. & R.P. Thomas (1973), *The Rise of the Western World: A New Economic History*, Cambridge: Cambridge University Press.
412. North, D.C. (1990), *Institutions, Institutional Change and Economic Performance*, Cambridge: Cambridge University Press.
413. North, D.C. (2005), *Understanding the Process of Economic Change*, Princeton: Princeton University Press.
414. Nugent, N. (2006), *The Government and Politics of the European Union* [6th edition], Basingstoke (Hampshire): Palgrave MacMillan.
415. OECD (1997), *Liberalisation and Structural Reform in the Freight Transport Sector in Europe*.
416. OECD (1998), *Railways: Structure, Regulation and Competition Policy*, DAF/CLP(98)1.
417. OECD (2005), *Structural Reform in the Rail Industry*, DAF/COMP(2005)46.
418. Olson, M. (1965), *The Logic of Collective Action: Public Goods and the Theory of Groups*, Cambridge (MA): Harvard University Press.
419. Olson, M. (2000), *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*, New York: Basic Books.
420. Oman, N.B. (2009), 'Specific Performance and the Thirteenth Amendment', *Minnesota Law Review*, Vol. 93, p. 2020 c.a., also available at: <http://ssrn.com/abstract=1462728>.
421. Ooststroom, H. van (2008), 'Markt op het spoor', *Economisch-Statistische Berichten*, Vol. 93, No. 4533, p. 242-245.
422. Aghion, P. & P. Howitt (1992), 'A Model of Growth Through Creative Destruction', *Econometrica*, Vol. 60, No. 2, p. 323-351.
423. Page, S.E. (2006), 'Essay: Path Dependence', *Quarterly Journal of Political Science*, Vol. 1, p. 87-115.
424. Page, W.H. (2008), 'The Ideological Origins and Evolution of US Antitrust Law', *Issues in Competition Law and Policy*, Vol. 1, No. 1, p. 1-17.
425. Parkin, M. (2008), *Microeconomics*, Boston: Pearson.
426. Parsons, T. (1963), 'On the Concept of Political Power', *Proceedings of the American Philosophical Society*, Vol. 107, No. 3, p. 232-262.
427. Pease, N. (1999), 'Chapter 1 – Historical Setting', in: G.E. Curtis [ed.], *Poland: A Country Study*, Washington: Library of Congress.
428. Pelkmans, J. (2006), *European Integration: Methods and Economic Analysis* (3rd edition), Harlow: Pearson.
429. Pérard, E. (2009), 'Water Supply: Public or private? An approach based on cost of funds, transaction costs, efficiency and political costs', *Policy and Society*, Vol. 27, p. 193-219.
430. Peritz, R.J.R. (1996), *Competition Policy in America, 1888-1992: History, Rhetoric, Law*, Oxford: Oxford University Press.

431. Pessali, H.F. (2006), 'The rhetoric of Oliver Williamson's transaction cost economics', *Journal of Institutional Economics*, Vol. 2, No. 1, p. 45-65.
432. Pierson, P. (2000), 'Increasing Returns, Path Dependence, and the Study of Politics', *American Political Science Review*, Vol. 94, No. 2, p. 251-267.
433. Pierson, P. (2004), *Politics in Time: History, Institutions, and Social Analysis*, Princeton (NJ): Princeton University Press.
434. Pietrantonio, L. Di & J. Pelkmans (2004), 'The Economics of EU Railway Reform', Bruges European Economic Policy briefing no. 8.
435. Pindyck, R.S. & D.L. Rubinfeld (2009), *Microeconomics* (7th edition), Upper Saddle River (NJ): Pearson.
436. Pittman, R. (2004), 'Russian Railways Reform and the Problem of Non-Discriminatory Access to Infrastructure', *Annals of Public and Cooperative Economics*, Vol. 75, No. 2, p. 167-192.
437. Pollock, F. & F.W. Maitland (1898), *The History of English Law before the time of Edward I*, Cambridge: Cambridge University Press.
438. Pollock, F. (1913), 'The Genius of the Common Law VII: Perils of the Market Place', *Columbia Law Review*, Vol. 13, No. 1, p. 1-12.
439. Posner, R.A. (1969), 'Oligopoly and the Antitrust Laws: A Suggested Approach', *Stanford Law Review*, Vol. 21, p. 1562-1606.
440. Posner, R.A. (1972), 'The Appropriate Scope of Regulation in the Cable Television Industry', *Bell Journal of Economics and Management Science*, Vol. 3, No. 1, p. 98-129.
441. Posner, R.A. (1973), *Economic Analysis of the Law*, Boston: Little, Brown & Co, 4th edition, 1992.
442. Posner, R.A. (1974), 'Exclusionary Practices and the Antitrust Laws', *University of Chicago Law Review*, Vol. 41, No. 3, p. 506-535.
443. Posner, R.A. (1979), 'The Chicago School of Antitrust Analysis', *University of Pennsylvania Law Review*, Vol. 127, p. 925-948.
444. Posner, R.A. (1981), 'the Next Step in the Antitrust Treatment of Restricted Distribution: Per Se Legality', *University of Chicago Law Review*, Vol. 48, No. 1, p. 6-26.
445. Posner, R.A. (2001), 'Antitrust in the New Economy', *Antitrust Law Journal*, Vol. 68, p. 925-943.
446. Posner, R.A. (2005), 'Vertical Restraints and Antitrust Policy', *University of Chicago Law Review*, Vol. 72, No. 1, p. 229-241.
447. Pot, C.W. van der & A.M. Donner (2001), *Handboek van het Nederlandse Staatsrecht*, edited by L. Prakke, J.L. De Reede & G.J.M. van Wissen, Deventer: W.E.J. Tjeenk Wilink.
448. Puviani, A. (1903/1960), *Die Illusionen in der öffentlichen Finanzwirtschaft*, Berlin: Duncker & Humblot, translated from: *Teoria dell'Illusione Finanziaria* (1903).
449. Raad van Economische Adviseurs (2007), 'De lof der eenvoud', KST 2006-2007, 30 942, nr. 2.
450. Raaijmakers, M.J.G.C. (2000), *Pitlo: Het Nederlands burgerlijk recht, Deel 2: Vennootschaps- en rechtspersonenrecht* [4th ed.], Deventer: Gouda Quint.
451. Ramsay, W. (1875), 'Viae', in: W. Smith [ed.], *A Dictionary of Greek and Roman Antiquities*, London: John Murray.
452. Ramuschkat, D. (2002), *Die Eisenbahnlinie Wittenberge-Buchholz: Eine Fallstudie über Entscheidungs- und Handlungsabläufe bei der Planung einer Privatbahn in Preußen und dem zum Bau erforderlichen Grunderwerb*, unpublished Ph.D. thesis, University of Hamburg.
453. Rapport, M. (2008), *1848: Year of Revolution*, New York: Basic Books.

454. Rasmussen, J. (1997), ‘What Kind of Vision is That?’ British Public Attitudes towards the European Community during the Thatcher Era’, *British Journal of Political Science*, Vol. 27, No. 1, p. 111-118.
455. Rawls, J. (1971), *A Theory of Justice*, Oxford: Clarendon Press.
456. Raz, J. (1972), ‘Legal Principles and the Limits of the Law’, *Yale Law Journal*, Vol. 81, No. 5, p. 823-854.
457. Reitsma, S.A. (1940), ‘Van Particuliere naar Staatsexploitatie der Nederlandsche Spoorwegen’, *De Economist*, Vol. 89, No. 1, p. 759-796.
458. Renzulli, A. (2008), ‘Services of General Economic Interest: The Post-Altmark Scenario’, *European Public Law*, Vol. 14, No. 3, p. 399-432.
459. Reynebeau, M. (2003), *Een Geschiedenis van België*, Tielt: Lannoo.
460. Riddell, P. (1989), *The Thatcher Era and its Legacy*, Oxford: Blackwell.
461. Robbins, L. (1932), *An Essay on the Nature & Significance of Economic Science*, 2nd edition, revised and extended, 1952, London: MacMillan & Co.
462. Robinson, J. (1933), *The Economics of Imperfect Competition*, London: MacMillan.
463. Roe, M.J. (1996), ‘Chaos and Evolution in Law and Economics’, *Harvard Law Review*, Vol. 109, p. 641-668.
464. Rohou, J.-L. (1998), ‘Une Voie Française pour le Renouveau du Chemin de Fer: La creation de réseau ferré de France’, *Revue générale des chemins de fer*, Vol. 1998, No. 7, p. 31-38.
465. Roschmann, C. (1993), ‘Comfort Letters under German and Austrian Law’, *SPG International Law Practicum*, Vol. 6, p. 21-27.
466. Ross, M. (2000), ‘Art. 16 EC and Services of General Interest: From Derogation to Obligation’, *European Law Review*, Vol. 25, No. 1, p. 22-38.
467. Rousseau, J.J. (1762), *Du Contrat Social*, Paris: Béziat.
468. Rowley, C.K. & F. Schneider [eds.] (2004), *The Encyclopedia of Public Choice*, Dordrecht (NL): Kluwer.
469. Roy, R.N. & K.K. Guin (1999), ‘A proposed model of JIT purchasing in an integrated steel plant’, *International Journal of Production Economics*, Vol. 59, p. 179-187.
470. Rubinstein, A. (1979), ‘Equilibrium in Supergames with the Overtaking Criterion’, *Journal of Economic Theory*, Vol. 21, No. 1, p. 1-9.
471. Rubinstein, A. (1998), *Modeling Bounded Rationality*, Cambridge: MIT Press.
472. Ruiter, D.W.P. (1993), *Institutional Legal Facts: Legal Powers and their Effects*, Dordrecht: Kluwer Academic Publishers.
473. Ruiter, D.W.P. (1994), ‘Economic and Legal Institutionalism: What can they learn from each other?’, *Constitutional Political Economy*, Vol. 5, No. 1, p. 99-115.
474. Ruiter, D.W.P. (2005), ‘Is Transaction Cost Economics Applicable to Public Governance?’, *European Journal of Law and Economics*, Vol. 20, p. 287-303.
475. Ruiter, D.W.P. (2008), ‘Unital and Multital Legal Relations’, in: M. Van Genugten & M. Harmsen [eds.], *De vorm behouden: Verslag van een levenswerk door Dick W.P. Ruiter*, Enschede: Universiteit Twente, p. 87-129.
476. Ruiter, D.W.P. (2008a), ‘De vormen van het recht’, in: *De vorm behouden: Verslag van een levenswerk door Dick W.P. Ruiter*, M. van Genugten & M. Harmsen [eds.], Enschede: University of Twente, p. 5-60.
477. Ruiter, D.W.P. (2008b), ‘Calabresi and Melamed’s entitlements: a Hohfeldian approach to ‘The Cathedral’’, in: *De vorm behouden: Verslag van een levenswerk door Dick W.P. Ruiter*, M. van Genugten & M. Harmsen [eds.], Enschede: University of Twente, p. 61-86.
478. Sacasas, R. & D. Wiesner (1987), ‘Comfort Letters: The Legal and Business Implications’, *Banking Law Journal*, Vol. 104, No. 4, p. 313 c.a.

479. Samuelson, P.A. (1958), 'An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money', *Journal of Political Economy*, Vol. 66, No. 6, p. 467-482.
480. Samuelson, W. & R. Zeckhauser (1988), 'Status Quo Bias in Decision Making', *Journal of Risk and Uncertainty*, Vol. 1, No. 1, p. 7-59.
481. Savage, L.J. (1954), *The Foundations of Statistics*, New York: John Wiley and Sons.
482. Schelling, T.C. (1956), 'An Essay on Bargaining', *American Economic Review*, Vol. 46, No. 3, p. 281-306.
483. Schmitthoff, C.M. & J.H. Thompson (1968), *Palmer's Company Law* [21st ed.], London: Stevens & Son.
484. Schugart, W.F. & L. Razzolini [eds.] (2001), *The Elgar Companion to Public Choice*, Cheltenham (UK): Edward Elgar.
485. Schumaker, P. (2008), *From Ideologies to Public Philosophies: An Introduction to Political Theory*, Oxford: Blackwell.
486. Schumpeter, J.A. (1949), 'Science and Ideology', *American Economic Review*, Vol. 39, No. 2, p. 346-359.
487. Schwartz, A. (1979), 'The Case for Specific Performance', *Yale Law Journal*, Vol. 89, p. 271-306.
488. Seita, A.Y. (1984), 'Uncertainty and Contract Law', *University of Pittsburgh Law Review*, Vol. 46, p. 75-148.
489. Shapiro, F.R. (1996), 'The Most-Cited Law Review Articles Revisited', *Chicago-Kent Law Review*, Vol. 71, p. 751-766.
490. Sherman, J.G. (2008), 'Can Religious Influence Ever Be "Undue" Influence?', *Brooklyn Law Review*, Vol. 73, No. 2, p. 579-644.
491. Siegfried, T. (2006), *A Beautiful Math: John Nash, Game Theory, and the Modern Quest for a Code of Nature*, Washington D.C.: Joseph Henry Press.
492. Simon, H.A. (1947), *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations*, 4th edition, New York: The Free Press.
493. Simon, H.A. (1955), 'A Behavioral Model of Rational Choice', *Quarterly Journal of Economics*, Vol. 69, No. 1, p. 99-118.
494. Simon, H.A. (1978), 'Rationality as Process and as Product of Thought', *American Economic Review*, Vol. 68, No. 2, p. 1-16.
495. Simons, P. & G. De Ridder (1839), *Le Chemin de Fer Belge, ou recueil des memoires et devis pour l'établissement du chemin de fer d'Anvers et Ostende a Cologne, avec embranchement de Bruxelles et de Gand aux frontieres de France* (3ième edition), Bruxelles: Hauman.
496. Singer, D. (1996), 'Mitterand's Legacy', *The Nation*, January 29, 1996.
497. Sloot, L.M. & M. van Aalst (2005), 'Prijsoorlog en prijsbeleving', *Economisch-Statistische Berichten*, Vol. 90, No. 4466, p. 321.
498. Smelser, N.J. & R. Swedberg (2005) [eds.], *The Handbook of Economic Sociology*, Princeton: Princeton University Press.
499. Smiles, S. (1859), *The Story of the Life of George Stephenson, Railway Engineer*, London: John Murray.
500. Smiles, S. (1868), *The Life of George Stephenson and of his Son Robert Stephenson, comprising also a history of the invention and introduction of the railway locomotive*, New York: Harper & Brothers.
501. Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*, Oxford: Clarendon Press.
502. Smith, J. (1900), 'Liability for Negligent Language', *Harvard Law Review*, Vol. 14, No. 3, p. 184-199.

503. Solsten, E. (1999), 'Chapter 1 – Historical Setting', in: E. Solsten [ed.], *Germany: A Country Study*, Washington, D.C.: Library of Congress.
504. Spencer, H. (1851), *Social Statics: or, The Conditions essential to Happiness specified, and the First of them Developed*, London: John Chapman.
505. Spencer, H. (1864), *The Principles of Biology*, New York: Appleton & Co.
506. Spenkuch, H. [ed.] (1999), *Die Protokolle des Preussischen Staatsministeriums 1817-1934/38, Band 7: 8. Januar 1879 bis 19. März 1890*, Berlin-Brandenburgischen Akademie der Wissenschaften, Hildesheim: Olms-Weidmann.
507. Spenkuch, H. [ed.] (2003), *Die Protokolle des Preussischen Staatsministeriums 1817-1934/38, Band 8-I und 8-II: 21. März 1890 bis 9. Oktober 1900*, Berlin-Brandenburgischen Akademie der Wissenschaften, Hildesheim: Olms-Weidmann.
508. Spiller, P.T. & M. Tommasi (2003), 'The Institutional Foundations of Public Policy: A Transactions Approach with Application to Argentina', *Journal of Law, Economics, and Organization*, Vol. 19, No. 2, p. 281-306.
509. Starling, G. (2002), *Managing the Public Sector* [6th edition], Belmont (CA): Thomson Learning.
510. Stasavage, D. (2007), 'Cities, Constitutions, and Sovereign Borrowing in Europe, 1274-1785', *International Organization*, Vol 61, p. 489-525.
511. Stauvermann, P.J. (2002), 'Why is there so much peace?', *Defence and Peace Economics*, Vol. 13, No. 1, p. 61-75.
512. Steenge, A.E. (2004), 'The Coase Theorem, Economic Lineage, and the Small Numbers Problem', in: A. Prinz, A.E. Steenge & J. Schmidt [eds.], *Institutions in Legal and Economic Analysis*, Münster: LIT Verlag.
513. Steenge, A.E. (2008), 'Structuur en hartstocht: gedachten over een magisch vierkant', Inaugural Lecture held at the occasion of accepting the position of professor of economics at the Rijksuniversiteit Groningen, on september 30, 2008.
514. Steger, M.B. & R.K. Roy (2010), *Neoliberalism: A Very Brief Introduction*, Oxford: Oxford University Press.
515. Steiner, G.A. & J.F. Steiner (2003), *Business, Government, and Society: A Managerial Perspective* (10th edition), New York: McGraw-Hill.
516. Steitz, W. (1974), *Die Entstehung der Köln-Mindener Eisenbahn-gesellschaft: Ein Beitrag zur Frühgeschichte der Deutschen Eisenbahnen und des Preussischen Aktienwesens*, Köln: Rheinisch-Westfälischen Wirtschaftsarchiv.
517. Sternø, S. (2004), *Solidarity in Europe: The History of an Idea*, Cambridge: Cambridge University Press.
518. Stigler, G. J. (1962), 'Information in the Labor Market', *Journal of Political Economy*, Vol. 70, No. 5, part 2, p. 94-105.
519. Stigler, G.J. (1961), 'The Economics of Information', *Journal of Political Economy*, Vol. 69, No. 3, p. 213-225.
520. Stigler, G.J. (1966), 'The Economic Effects of the Antitrust Laws', *Journal of Law and Economics*, Vol. 9, p. 225-258.
521. Stigler, G.J. (1968), *The Organization of Industry*, Chicago: University of Chicago Press.
522. Stigler, G.J. (1972), 'The Law and Economics of Public Policy: A Plea to the Scholars', *Journal of Legal Studies*, Vol. 1, No. 1, p. 1-12.
523. Stolker, C. (2002), 'Ja, geléerd zijn jullie wel!', *Nederlands Juristenblad*, Vol. 2002, No. 15, p. 766 c.a.
524. Suematsu, C. & M. Makabenta-Ikeda (2006), 'Interface from Transaction Cost Approach', Working Paper No. 84, Graduate School of Economics, Kyoto University, available at: <http://ssrn.com/abstract=917079>.

525. Sugden, R. (2009), 'Can Economics be founded on 'Indisputable Facts of Experience'? Lionel Robbins and the Pioneers of Neoclassical Economics', *Economica*, Vol. 76, No. s1, p. 857-872.
526. Sunstein, C.R. & R.A. Thaler (2003), 'Libertarian Paternalism is Not an Oxymoron', *University of Chicago Law Review*, Vol. 70, No. 4, p. 1159-1202.
527. Sunstein, C.R. & R.A. Thaler (2006), 'Preferences, Paternalism, and Liberty', *Royal Institute of Philosophy: Supplement*, Vol. 59, p. 233-264.
528. Surrey, J. [ed.] (1996), *The British Electricity Experiment: Privatization: the Record, the Issues, the Lessons*, London: Earthscan Publications.
529. Szladits, C. (1955), 'The Concept of Specific Performance in Civil Law', *American Journal of Comparative Law*, Vol. 4, No. 2, p. 208-234.
530. Taylor, A.J.P. (1945), *The course of German history: A survey of the development of German history since 1815*, London: Routledge.
531. Tetlock, P.E. & A. Belkin [eds.] (1996), *Counterfactual Thought Experiments in World Politics: Logical, Methodological, and Psychological Perspectives*, Princeton, NJ: Princeton University Press.
532. Teulings, C.N., A.L. Bovenberg & H.P. van Dalen (2003), *De Calculus van het Publieke Belang*, The Hague (NL): Kenniscentrum voor Ordeningsvraagstukken.
533. Thaler, R.A. & C.R. Sunstein (2003), 'Behavioral economics, Public Policy, and Paternalism – Libertarian Paternalism', *American Economic Review*, Vol. 93, No. 2, p. 175-179.
534. Thaler, R.A. & C.R. Sunstein (2008a), *Nudge: Improving Decisions about Health, Wealth and Happiness*, New Haven: Yale University Press.
535. Thaler, R.A. & C.R. Sunstein (2008b), 'Easy Does It – How to make lazy people do the right thing', *The New Republic*, issue of April 9, 2008, p. 20-22.
536. Thatcher, M.H. (1995), *The Path to Power*, New York: HarperCollins.
537. Thirlway, H. (2001), 'Reflections on *Lex Ferenda*', *Netherlands Yearbook of International Law*, Vol. XXXII, p. 3-26.
538. Tiebout, C.M. (1956), 'A Pure Theory of Local Expenditures', *Journal of Political Economy*, Vol. 64, No. 5, p. 416-424.
539. Tilburg, C. van (2005), *Romeins Verkeer: Weggebruik en verkeersdrukte in het Romeinse Rijk*, Amsterdam: Amsterdam University Press.
540. Tomlinson, W.W. (1915), *The North Eastern Railway: its rise and development*, London: Longmans.
541. Trebilcock, M.J. & R. Howse (1995), *The Regulation of International Trade*, London: Routledge.
542. Tsu, F. (2007), 'Analyzing Comfort Letters: The Brazilian Legal Perspective', *Law and Business Review of the Americas*, Vol. 13, p. 167-179.
543. Tsui, K.K. (2010), 'More Oil, Less Democracy: Evidence from Worldwide Crude Oil Discoveries', *Economic Journal*, forthcoming.
544. Tullock, G. (1967), 'The Welfare Costs of Tariffs, Monopolies and Theft', *Western Economic Journal/Economic Inquiry*, Vol. 5, p. 224-232.
545. Tullock, G. (1970), 'A Simple Algebraic Logrolling Model', *American Economic Review*, Vol. 60, No. 3, p. 419-426.
546. Tullock, G. (2002), 'The Theory of Public Choice', in: G.A. Tullock, A. Seldon & G.L. Brady (2002), *Government Failure: A Primer in Public Choice*, Washington DC: Cato Institute, p. 1-79.
547. Tullock, G. (2005), *Public Goods, Redistribution and Rent Seeking*, Cheltenham (UK): Edward Elgar.

548. Tullock, G., A. Seldon & G.L. Brady (2002), *Government Failure: A Primer in Public Choice*, Washington DC: Cato Institute.
549. Tyrrall, D. (2003), 'The UK railway industry: a failed experiment in transaction cost economics', *European Business Journal*, Vol. 15, No. 1, p. 38-48.
550. Tyrrall, D. (2004), 'The UK Railway Privatisation: Failing to Succeed?', *Economic Affairs*, Vol. 24, No. 3, p. 32-38.
551. Usher, D. (1998), 'The Coase theorem is tautological, incoherent or wrong', *Economics Letters*, Vol. 61, p. 3-11.
552. Van Dale (1992), *Groot Woordenboek der Nederlandse Taal*, edited by G. Geerts and H. Heestermans, Utrecht: Van Dale.
553. Veblen, T. (1915), *Imperial Germany and the Industrial Revolution*, New York: MacMillan.
554. Veenendaal, A.J. (1998), *De IJzeren Weg in een Land vol Water: Beknopte geschiedenis van de spoorwegen in Nederland, 1834-1958*, Amsterdam: De Bataafsche Leeuw.
555. Veenendaal, A.J. (2004), *Spoorwegen in Nederland: Van 1834 tot nu*, Amsterdam: Boom.
556. Vincent, J. (1987), 'The Thatcher Governments, 1979-1987', in: P. Hennessy & A. Seldon, *Ruling Performance: British Governments from Attlee to Thatcher*, Oxford: Blackwell, p. 274-300.
557. Vogel, S. (1996), *Freer Markets, More Rules: Regulatory Reform in the Advanced Industrial Countries*, Ithaca (NY): Cornell University Press.
558. Volokh, A. (2007), 'Privatization and the Effectiveness of Monitoring Agencies', Georgetown Law and Economics Research Paper No. 982146, available at <http://ssrn.com/abstract=982146>.
559. Vromans, M.J.C.M. (2005), *Reliability of Railway Systems*, Ph.D. Thesis Erasmus University Rotterdam.
560. Vromen, J. (1995), *Economic Evolution: An Inquiry into the Foundations of New Institutional Economics*, London: Routledge.
561. Wall, S. (2008), *A Stranger in Europe: Britain and the EU from Thatcher to Blair*, Oxford: Oxford University Press.
562. Wall-Bake, Van den (1907a), 'De Zwitsersche Spoorwegen', *De Economist*, Vol. 56, No. 1, p. 1-24.
563. Wall-Bake, Van den (1907b), 'De Italiaansche Spoorwegen', *De Economist*, Vol. 56, No. 1, p. 85-108.
564. Wallis, J.J. & D.C. North (1986), 'Measuring the Transaction Sector in the American Economy, 1870 to 1970', in: S.L. Engerman & R.E. Gallman [eds.], *Long Term Trends in the American Economy*, Chicago: University of Chicago Press.
565. Wallis, J.J. & D.C. North (1988), 'Should Transaction Costs be Subtracted from Gross National Product?', *Journal of Economic History*, Vol. 48, No. 3, p. 651-654.
566. Wallström, M. (2008), 'Communicating Europe Together', *EUObserver* of 22.10.2008, available at: <http://euobserver.com/843/26956>.
567. Weatherhead, P.J. (1979), 'Do Savannah Sparrows commit the Concorde Fallacy?', *Behavioral Ecology and Sociobiology*, Vol. 5, p. 373-381.
568. Weber, M. (1920), *Wirtschaft und Gesellschaft*, Tübingen: Mohr.
569. Wegberg, M. van & A. van Witteloostuijn (1992), 'Credible Entry Threats into Contestable Markets: A Symmetric Multi-Market Model of Contestability', *Economica*, Vol. 59, No. 236, p. 437-452.
570. Wegberg, M. van, A. van Witteloostuijn & M. Roscam Abbing (1994), 'Multimarket and Multiproject Collusion: Why European Integration May Reduce Intra-Community Competition', *De Economist*, Vol. 142, No. 3, p. 253-285.

571. Weidmann, U. (2008), 'Open access to railway networks: Hidden discrimination potential in an integrated railway organisation', First Annual Conference on Competition and Regulation in Network Industries, 28 November 2008.
572. Wels, C.B. (2001), *Vrije Doorvaart Schelde in Historisch Perspectief: 160 jaar Permanente Commissie*, pamphlet available online at: <http://www.scheldenet.nl/nl/scheldebekleid/overlegstructuren/permcom/brochure?lng=nl>.
573. Werner, W.G. & R.A. Wessel (2010), *Internationaal en Europees Recht: Een verkenning van de grondslagen en kenmerken sinds het Verdrag van Lissabon*, 2nd revised edition, Groningen: Europa Law Publishing.
574. Werner, W.G. (1995), *Het recht geworden woord: Over de geschiedenis van het rechtspositivisme en de mogelijke betekenis van het pragmatisme voor de toekomst daarvan*, Enschede: Universiteit Twente.
575. Wessel, R.H. (1967), 'A Note on Economic Rent', *American Economic Review*, Vol. 57, No. 5, p. 1221-1226.
576. Whitehead, P. (1987), 'The Labour Governments, 1974-1979', in: P. Hennessy & A. Seldon, *Ruling Performance: British Governments from Attlee to Thatcher*, Oxford: Blackwell, p. 241-273.
577. Whitehouse, L. (2003), 'Railtrack is Dead: Long Live Network Rail? Nationalization under the Third Way', *Journal of Law and Society*, Vol. 30, No. 2, p. 217-235.
578. Whitfield, B. (2001), *The Extension of the Franchise: 1832-1931*, Oxford: Heinemann.
579. Wickevoort Crommelin, I. van (1883), 'De Spoorweg-Enquête en het Personenvervoer', *De Economist*, Vol. 32, No. 1, p. 303-338.
580. Wijffels, H.H.F., R.J. in 't Veld & J.F.A. de Soet (1992), *Sporen voor Straks: Advies over de toekomstige relatie tussen overheid en Nederlandse Spoorwegen*, Advies uitgebracht aan de Minister van Verkeer & Waterstaat op 5 Juni 1992.
581. Williamson, J. (1990), 'What Washington Means by Policy Reform', in: J. Williamson [ed.], *Latin American Adjustment: How Much Has Happened?*, Washington D.C., Institute for International Economics.
582. Williamson, J. (2002), 'Did the Washington Consensus Fail?', <http://www.petersoninstitute.org/publications/papers/paper.cfm?ResearchID=488>.
583. Williamson, O.E. (1967), *The Economics of Discretionary Behaviour: Managerial Objectives in a Theory of the Firm*, London: Kershaw Publishing.
584. Williamson, O.E. (1971), 'The Vertical Integration of Production: Market Failure Considerations', *American Economic Review*, Vol. 61, No. 2, p. 112-123.
585. Williamson, O.E. (1975), *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: The Free Press.
586. Williamson, O.E. (1976), 'Franchise bidding for natural monopolies – in general and with respect to CATV', *Bell Journal of Economics*, Vol. 7, No. 1, p. 73-104.
587. Williamson, O.E. (1976), 'Franchise bidding for natural monopolies – in general and with respect to CATV', *Bell Journal of Economics*, Vol. 7, No. 1, p. 73-104.
588. Williamson, O.E. (1977), 'Predatory Pricing: A Strategic and Welfare Analysis', *Yale Law Journal*, Vol. 87, p. 284-340.
589. Williamson, O.E. (1981), 'The Economics of Organization: The Transaction Cost Approach', *American Journal of Sociology*, Vol. 87, No. 3, p. 548-577.
590. Williamson, O.E. (1982), 'Dominant Firms and the Monopoly Problem: Market Failure Considerations', *Harvard Law Review*, Vol. 85, p. 1512-1531.
591. Williamson, O.E. (1983), 'Credible Commitments: Using Hostages to Support Exchange', *American Economic Review*, Vol. 73, No. 4, p. 519-540.

592. Williamson, O.E. (1983), 'Organization Form, Residual Claimants, and Corporate Control', *Journal of Law and Economics*, Vol. 26, No. 2, p. 351-366.
593. Williamson, O.E. (1985), *The Economic Institutions of Capitalism*, New York: The Free Press.
594. Williamson, O.E. (1987), *Antitrust Economics: Mergers, Contracting, and Strategic Behavior*, Oxford: Blackwell.
595. Williamson, O.E. (1988), 'Corporate Finance and Corporate Governance', *Journal of Finance*, Vol. 43, No. 3, p. 567-591.
596. Williamson, O.E. (1991), 'Comparative Economic Organization: The Analysis of Discrete Structural Alternatives', *Administrative Science Quarterly*, Vol. 36, No. 2, p. 269-296.
597. Williamson, O.E. (1993a), 'Calculativeness, Trust, and Economic Organization', *Journal of Law and Economics*, Vol. 36, No. 1, Part. 2, p. 453-486.
598. Williamson, O.E. (1993b), 'Transaction Cost Economics and Organization Theory', *Industrial and Corporate Change*, Vol. 2, p. 107-156.
599. Williamson, O.E. (1996), 'Transaction cost economics and the Carnegie connection', *Journal of Economic Behavior & Organization*, Vol. 31, p. 149-155.
600. Williamson, O.E. (1998), 'Transaction Cost Economics: How it works; Where it is headed', *De Economist*, Vol. 146, No. 1, p. 23-58.
601. Williamson, O.E. (1999), 'Public and Private Bureaucracies: A Transaction Cost Economics Perspective', *Journal of Law, Economics, and Organization*, Vol. 15, No. 1, p. 306-342.
602. Williamson, O.E., M.L. Wachter & J.E. Harris (1975), 'Understanding the Employment Relation: The Analysis of Idiosyncratic Exchange', *Bell Journal of Economics*, Vol. 6, No. 1, p. 250-278.
603. Williston, S. (1911), 'Liability for Honest Misrepresentation', *Harvard Law Review*, Vol. 24, No. 6, p. 415-440.
604. Wilske, S. & T. Schiller (1997), 'International Jurisdiction in Cyberspace: Which States may Regulate the Internet?', *Federal Communications Law Journal*, Vol. 50, p. 117-178.
605. Wilson, J.Q. (1989) *Bureaucracy: What Government Agencies Do and Why They Do It*, Basic Books.
606. Wilson, R. (1969), 'An Axiomatic Model of Logrolling', *American Economic Review*, Vol. 59, No. 3, p. 331-341.
607. Winden, F. van (2004), 'Interest Group Behavior and Influence', in: C.K. Rowley & F. Schneider [eds.] (2004), *The Encyclopedia of Public Choice*, Dordrecht (NL): Kluwer, p. 118-128.
608. Winston, C. (2006), *Government Failure versus Market Failure: Microeconomics Policy Research and Government Performance*, Washington, D.C.: AEI-Brookings Joint Center for Regulatory Studies.
609. Witte, K. de & D.S. Saal (2008), 'Is a Little Sunshine All We Need? On the Impact of Sunshine Regulation on profits, productivity and prices in the Dutch drinking water sector.', available at <http://ssrn.com/abstract=1290919>.
610. Witteloostuijn, A. van & M. van Wegberg (1992), 'Multimarket competition: Theory and Evidence', *Journal of Economic Behavior and Organization*, Vol. 18, p. 273-282.
611. Wittman, D. (1999), 'General Structure of the Law', in: B. Bouckaert & G. De Geest [eds.], *Encyclopedia of Law and Economics, Volume 1. The History and Methodology of Law and Economics*, Cheltenham: Edward Elgar.
612. Wittsiepe, R. (2008), *IFRS for Small and Medium-Sized Enterprises: Structuring the Transition Process*, Wiesbaden: Gabler.

613. Woerdman, E. (2002), *Implementing the Kyoto Mechanisms: Political Barriers and Path Dependence*, Dissertation University of Groningen.
614. Wolmar, C. (2007), *Fire & Steam: How the Railways Transformed Britain*, London: Atlantic Books.
615. WRR (2000), *Het Borgen van Publiek Belang*, The Hague: Sdu Uitgevers.
616. Wyckoff, P.G. (1990), 'The simple analytics of slack-maximizing bureaucracy', *Public Choice*, Vol. 67, No. 1, p. 35-47.
617. Yale Law Journal (1938), 'Liability of Newspaper for Negligent Misstatements', *Yale Law Journal*, Vol. 47, No. 3, p. 461-464.
618. Yvrande-Billon, A. & C. Ménard (2005), 'Institutional Constraints and Organizational Changes: The Case of the British Rail Reform', *Journal of Economic Behavior & Organization*, Vol. 56, p. 675-699.
619. Yvrande-Billon, A. (1999), 'Is the New British Railways Structure Stable? A Transaction Cost Economics Analysis', Danish Research Unit for Industrial Dynamics Conference Winter 2000.
620. Yvrande-Billon, A. (2000), 'The New British Railways Structure: A Transaction Cost Economics Analysis', Danish Research Unit for Industrial Dynamics Working Paper 00-5.
621. Yvrande-Billon, A. (2006), 'The Attribution Process of Delegation Contracts in the French Urban Public Transport Sector: Why Competitive Tendering is a Myth', *Annals of Public and Cooperative Economics*, Vol. 77, No. 4, p. 453-478.
622. Zanmarone, G. (2008), 'Fiat without Authority under Vertical Integration', available at <http://ssrn.com/abstract=1285960>.
623. Zerbe, R.O. (1980), 'The Problem of Social Cost in Retrospect', *Research in Law and Economics*, Vol. 2, p. 83-102.
624. Zeumer, S. (2003), *Die Kölner Privatbanken und die Industriefinanzierung im frühen 19. Jahrhundert*, Diplomarbeit Universität zu Köln.
625. Ziegler, D. (1996), *Eisenbahnen und Staat im Zeitalter der Industrialisierung: Die Eisenbahnpolitik der Deutschen Staaten im Vergleich*, Stuttgart: Franz Steiner.
626. Zimmerling, R. (2005), *Influence and Power: Variations on a Messy Theme*, Dordrecht (NL): Springer.

Nederlandstalige Samenvatting

De Transactiekosten van semi-publieke instituties: Spoorweggeschiedenis als horde voor Coase

1. Inleiding

Cru gezegd kun je met een economisch model en een bijbehorend stuk werkelijkheid twee dingen doen: je kunt de werkelijkheid gebruiken om het model beter te begrijpen, en eventueel aan te passen, of je kunt het model gebruiken om de werkelijkheid beter te begrijpen (en eventueel aan te passen). In dit proefschrift behandel ik een aantal modellen die voortbouwen op het werk van Coase over transactiekosten, en pas ik die modellen toe op de geschiedenis van de spoorwegen. Het doel is om te onderzoeken of deze modellen zinvol op vragen van regulering en nationalisatie kunnen worden toegepast.

De transactiekostentheorie van Coase heeft zijn wortels in de private sector; het is een model dat verklaart waarom bedrijven fuseren en splitsen, waarom ze sommige dingen binnenshuis regelen terwijl andere zaken via een externe partij opgelost worden. Het moge duidelijk zijn dat nationalisering naar analogie van een fusie behandeld kan worden, terwijl privatisering gewoon het tegenovergestelde is.

De spoorwegen zijn hiervoor een interessante casus. Het is een sector met een lange geschiedenis, waarin zowel periodes van puur vrije marktdenken als publieke controle voorkomen. Tegelijkertijd is het een sector waar we redenen hebben om aan te nemen dat de optimale oplossing niet altijd zal prevaleren. Het ontbreekt immers vaak aan voldoende prikkels die de staat dwingen de optimale instituties in te voeren, en bovendien ontbreekt het aan voldoende vergelijkingsmateriaal om de staat in staat te stellen te ontdekken wat het optimum überhaupt is.

Dit betekent wel dat we vastzitten aan een sector met weinig “data-points”. In de Franse watersector heeft elk *departement* zijn eigen institutionele oplossingen. Hetzelfde geldt voor gevangenissen in verschillende Amerikaanse staten. Maar de spoorwegen zijn meestal op nationaal niveau georganiseerd, zodat een kwantitatieve analyse vrijwel onmogelijk is. In

plaats daarvan beschouwen we – op een kwalitatieve manier – de momenten waarop het institutionele kader van de spoorwegen in het verleden hervormd is. Door te kijken of deze hervormingen de rol van de overheid vergrootten of verkleinden op een manier die vanuit de theorie te verklaren valt, kunnen we hopelijk iets zeggen over de verklarende waarde van de verschillende modellen.

2. Theorie

Het beginpunt voor alle Institutionele Economie is het inzicht dat alle contracten, behalve misschien spot-contracten, onvermijdbaar incompleet zijn. Omdat het schrijven van een contract kosten met zich meebrengt, zullen partijen nooit alle denkbare situaties uitonderhandelen, en zullen ze veel gevallen op één hoop gooien ook als een flexibeler contract misschien efficiënter zou zijn geweest. Daar komt nog bij dat zelfs de dingen die partijen wel in het contract schrijven niet altijd voor de rechter afdwingbaar zijn, bij voorbeeld omdat ze onmogelijk te controleren zijn.

In 1933 trok Ronald Coase hieruit de – achteraf gezien – enig juiste conclusie: ondernemingen bestaan omdat het gebruik van de markt kostbaar is. De vraag of een bepaalde transactie via de markt of binnen de onderneming wordt opgelost hangt ervan af of de transactiekosten die voortvloeien uit het gebruik van de markt worden gecompenseerd door de efficiëntievoordelen van de markt. Het Coase Theorema is bijna een spiegelbeeld van dit inzicht: In een wereld zonder transactiekosten is het onbelangrijk hoe de eigendomsrechten oorspronkelijk verdeeld zijn, want uiteindelijk zal de markt er altijd voor zorgen dat ze op een welvaartsmaximaliserende manier verdeeld worden.

De vraag is dus waaruit die transactiekosten bestaan, en waardoor ze veroorzaakt worden. Over de eerste kunnen we kort zijn: transactiekosten bestaan uit de kosten van het zoeken van een tegenpartij, het onderhandelen met de tegenpartij en het handhaven van het contract. Dit kan in de praktijk weliswaar leiden tot de conclusie dat meer dan 60% van de economie uit transactiekosten bestaat, maar dat is niet echt een bezwaar. De vraag is immers niet hoe hoog de transactiekosten in absolute zin zijn, maar hoe ze beïnvloed worden door verschillende *governance* oplossingen, d.w.z. door de keuze voor markt, hiërarchie of iets ertussenin.

Dus resteert de vraag naar de bron van transactiekosten. In welke omstandigheden zullen ze bijzonder hoog zijn? Volgens Oliver Williamson wordt het met name moeilijk als een van de partijen geacht wordt transactie-specifieke investeringen te doen, d.w.z. investeringen die geen of weinig waarde hebben behalve voor deze specifieke transactie. In dat geval, be- toogt hij, zal de partij die deze investering niet heeft proberen misbruik te maken van de situa- tie door het contract zoveel mogelijk in zijn voordeel te “heronderhandelen”. Dergelijk gedrag is slechts met moeite – met kosten – te vermijden omdat alle contracten incompleet zijn.

De modellen van Hart en Moore en hun medeauteurs benadrukken eerder de manier waarop de prikkels van partijen niet op een lijn zitten. Het contract laat immers elke partij tot op zekere hoogte vrij om zijn eigen belang te dienen, en de manier waarop hij dat doet hangt af van de persoonlijke kosten en baten die aan zijn handelen verbonden zijn, niet van de totale kosten en baten. Een fusie beïnvloedt niet zozeer de vrijheid van handelen van de overgeno- men partij, maar wel de prikkels die hij heeft om keuzes te maken die het algemeen belang niet dienen.

Resteert nog de vraag waarom het optimum dat beschreven wordt door de theorie ook daadwerkelijk de verwachte uitkomst zou moeten zijn. Deze stap wordt door economen vaak overgeslagen, maar als het gaat om semi-publieke sectoren zoals de spoorwegen is dit wel degelijke een vraag die enige aandacht behoeft. Of er voor de spoorwegen, via evolutie, con- currentie of bewuste analyse, een waarneembare tendens is richting het optimum is een empi- rische vraag, die in hoofdstuk 3 aan de orde komt.

3. Spoorweggeschiedenis

In dit hoofdstuk bespreek ik de spoorweggeschiedenis van vijf landen: Nederland, België, Engeland, Frankrijk en Duitsland. Daarbij worden vier periodes onderscheiden: Innovatie, Consolidatie, Nationalisering en Privatisering.

Tijdens de eerste periode wisten de regeringen eigenlijk niet goed wat ze met de nieu- we technologie aan moesten. Er bestond veel onbegrip, en met vallen en opstaan wordt een werkend systeem op poten gezet. Vrijwel altijd komt het initiatief hiervoor van private partij- en, omdat directe overheidsinterventie – letterlijk – ondenkbaar was. Alleen in België kwam het eerste initiatief van de overheid, vlak na de onafhankelijkheid. In Nederland is er dan nog het geval van de Rijnspoorweg, die door persoonlijke bemoeienis van Koning Willem I als

privaat bedrijf werd opgezet, terwijl in enkele Duitse staten ook publieke spoorwegen bestonden. Frankrijk is een vreemd geval. In dat land werd een wet aangenomen om, naar Belgisch voorbeeld, een kernnet op overheidskosten aan te leggen, maar zodra die wet er was werd onmiddellijk gebruik gemaakt van de mogelijkheid, die de wet ook schiep, om voor een private oplossing te kiezen. Blijkbaar ontbrak het vooral aan vertrouwen in voldoende ondersteuning van overheidskant.

Tijdens de consolidatieperiode begint het netwerk steeds meer vorm te krijgen. Als gevolg hiervan begonnen actoren zich steeds meer bewust te worden van de economies of scale op het spoor. Grotere bedrijven kunnen efficiënter met hun materieel omgaan, en kunnen een betere afstemming bieden tussen verschillende lijnen. Daar komt nog bij dat er in de 19^e eeuw uiteraard nog geen sprake was van mededingingsrecht, zodat ook om die reden grote ondernemingen een voordeel hadden.

Het resultaat was dat in drie van de vijf landen onder beschouwing een kleine groep grote spelers overbleef: HSM, NRS en SS in Nederland, de “grote zes” in Frankrijk en een soortgelijk aantal in het VK. In België werd de markt weliswaar gecontroleerd door twee grote banken, de Generale Maatschappij en de Bank van België, maar dat betekende niet er ook maar twee spoorwegondernemingen waren. Alleen in Duitsland was de situatie anders: zolang het land nog versnipperd was in een groot aantal onafhankelijke staten en staatjes was van consolidatie in de meeste staten (uiteraard) geen sprake. Tegelijk met de toenemende invloed van Pruisen in Duitsland voltrok zich binnen Pruisen een proces van gestage nationalisatie. Dit model werd in de rest van Duitsland overgenomen, alhoewel de andere Bondsstaten zich krachtig verzetten tegen pogingen om de spoorwegen op Rijksniveau te beheren.

In de andere vier landen werd nationalisatie tegen het einde van de 19^e eeuw ook steeds meer een thema. Vaak kwam dat door militaire overwegingen, bijvoorbeeld in Frankrijk na het debacle van 1870, maar nog vaker doordat de grote spoorwegondernemingen ondanks hun marktmacht toch niet bijzonder winstgevend bleken. In België moesten de Generale en de BB hun spoorwegbelangen om financiële redenen van de hand doen, in Frankrijk lagen de grote zes – na 1870 gereduceerd tot vijf – vrijwel permanent aan het overheidsinfuus en zelfs in het VK en in Nederland was zo nu en dan overheidsinterventie nodig om financiële moeilijkheden te voorkomen.

Voor zover bij het aanbreken van de Eerste Wereldoorlog de spoorwegen in Europa nog niet genationaliseerd waren, gebeurde dat tijdens en na de oorlog wel. In Nederland werd

de NS gevormd als joint venture tussen de twee overgebleven spoorwegmaatschappijen, in Duitsland werd uit de ruïnes van de oorlog de Deutsche Reichsbahn gevormd, en in Engeland werd het spoor tussen vier gereguleerde bedrijven verdeeld, die uiteindelijk na de Tweede Wereldoorlog genationaliseerd werden. Alleen in Frankrijk bleef verandering uit. Daar werd de SNCF pas in 1937 gevormd, toen de vier overgebleven grote spoorwegmaatschappijen hun verliezen door de Depressie nog verder zagen oplopen dan voordien.

Tot de jaren '80 bleef het overheidsmonopolie in alle vijf landen ongewijzigd. Gedurende die tijd werd er weliswaar flink bezuinigd en gereorganiseerd, en werd een significant deel van het netwerk gesloten, maar men ging er toch vanuit dat de spoorwegen onvermijdelijk door de overheid gecontroleerd moesten worden. Pas bij het aantreden van premier Thatcher in 1979 werd dat idee voor het eerst weer in twijfel getrokken, al schrok ook zij terug voor daadwerkelijke privatisering. Dus moest de privatisering van British Rail wachten tot het kabinet-Major in de jaren '90, een voorbeeld dat door velen is bewonderd maar door niemand is nagedaan.

In plaats daarvan kwam de Europese wetgever vanaf 2001 met een serie spoorwegpakketten, waarin een verregaande hervorming van de sector werd opgelegd, maar zonder het Britse model zelfs maar bij benadering verplicht te stellen. Het gevolg is dat in de meeste Europese landen de voormalige monopolist nog steeds monopolist dan wel bijna-monopolist is, tenminste in de passagierssector, terwijl zowel de dominante vervoerder als ook de netwerkbeheerder nog steeds eigendom zijn van de staat. Alleen in het Verenigd Koninkrijk is British Rail bij de privatisering compleet opgeheven, terwijl de netwerkbeheerder na een korte periode als beursgenoteerde onderneming tegenwoordig als stichting ("company limited by guarantee") is georganiseerd.

Resteert de vraag of deze ontwikkelingen te begrijpen zijn vanuit de Institutionele Economie. Het antwoordt luidt dat de Institutionele Economie zo goed werkt als van een economisch model, dat vele categorieën verklaringen *a priori* uitsluit, verwacht mag worden. Zonder twijfel is het slechts met grote moeite mogelijk om de politieke aspecten van de spoorweggeschiedenis in de mal van de Institutionele Economie te dwingen. De geschiedenis bevat echter ook veel gevallen van agency-problemen, die prima binnen het oorspronkelijke model van Coase en de latere modellen van Hart en Moore passen. Alleen de Transactiekostentheorie van Williamson valt tegen; hoewel de spoorwegmaatschappijen investeren in transactiespecifieke activa, zijn zij het juist die misbruik maken van de overheid, niet omgekeerd.

4. Andere Benaderingen

Niettegenstaande de conclusies van hoofdstuk drie, kunnen we pas echt iets zeggen over de waarde van het werk van Coase voor semi-publieke sectoren als we iets hebben om het mee te vergelijken. De twee economische benaderingen die daarvoor het meest in aanmerking komen zijn Public Choice Theory en Padafhankelijkheid.

De kern van Public Choice Theory is het inzicht dat ook de mensen die de overheid vormen maar gewone mensen zijn, die net als alle andere individuen hun nut maximaliseren. Als zodanig baseren ze hun keuzes op hun persoonlijke belangen, voor zover ze niet door hun superieuren, kiezers of volk worden gedwongen een ander belang ook mee te wegen. Voeg daar nog aan toe de rol van de belangengroeperingen en andere lobbyisten, en er ontstaat een behoorlijk complex beeld van het bestuur, veel complexer dan de benadering van hoofdstuk 2 en 3 die de regering als één geheel beschouwt.

Voor de spoorwegen betekent dit aan de ene kant meer aandacht voor de mogelijkheden die de sector heeft om de overheid te bewegen in haar voordeel – en in het nadeel van de bevolking als geheel – van het algemeen belang af te wijken. De spoorsector bestaat immers maar uit een beperkt aantal actoren, die relatief makkelijk samen kunnen lobbyen. Het belang van de kiezer/belastingbetaler is daarentegen veel minder goed vertegenwoordigd bij de regering.

Aan de andere kant blijven nog vele vragen over: Waarom zou een machtige vorst als de koning van Pruisen een revolutionaire technologie als de spoorwegen willen bevorderen? Zoals de hertog van Wellington al zei, de spoorwegen “[only] encourage the lower classes to travel about”. In alle landen die in dit proefschrift beschouwd zijn, hebben de oude machthebbers het algemeen belang veel meer laten prevaleren dan een strikt “egoïstisch” model van menselijk gedrag zou laten vermoeden. Hoewel de economische theorie van het menselijk gedrag op geen enkele manier belet om aan het geluk van anderen gewicht toe te kennen, betekent dit in de praktijk een nogal onhandige manier van werken. Occam’s razor suggereert dan dat het handiger is om het domein van de economische wetenschap te verlaten en gewoon aan te nemen dat mensen daadwerkelijk altruïstisch kunnen zijn.

Daar komt nog bij dat Public Choice Theory maar weinig waarde heeft als voorspelend model. In vrijwel alle gevallen is het mogelijk om de historische hervormingen vanuit

Public Choice perspectief te verklaren, maar die verklaringen zijn vaak nogal *ad hoc*. In het ene geval ligt de nadruk op de rol van de bureaucratie, in een ander geval wordt juist de rol van lobbygroepen benadrukt, etc. Dit maakt het niet makkelijker om naar de toekomst gerichte positieve en normatieve uitspraken te doen.

Padafhankelijkheid complementeert het werk van Coase op een heel andere manier. In tegenstelling tot de meeste economische theorie op het gebied van instituties focust het expliciet op de manier waarop instituties veranderen in de tijd. Of, beter gezegd, de manier waarop ze niet veranderen. Een puur economische benadering zou immers doen veronderstellen dat de “regels van het spel” constant meeveranderen met de variabelen die worden gebruikt om ze te verklaren. De modellen die onder de noemer padafhankelijkheid vallen bieden een raamwerk om te verklaren waarom instituties in werkelijkheid maar zelden veranderen.

Hoewel het padafhankelijkheidsraamwerk deze functie uitstekend vervult als het gaat om de spoorwegen, is dat toch onvoldoende om te concluderen dat het meer verklarende kracht heeft dan de benaderingen die gebaseerd zijn op het werk van Coase. Het probleem is namelijk dat padafhankelijkheid idealiter ook zou moeten verklaren waarom dingen soms juist wel veranderen. En als we dat proberen te doen, dan stuiten we onvermijdelijk weer op het feit dat de timing van vrijwel alle belangrijke hervormingen zijn wortels had in politieke omstandigheden, en niet in economische.

5. Conclusie

De conclusie luidt dan ook dat de theorieën van Coase en zijn navolgers, met uitzondering van Williamson, vrij aardig werken. De kosten van het gebruik van de markt, zoals Coase ze in 1933 beschreef, verklaren een significant deel van de ontwikkelingen in de spoorwegsector. Hetzelfde geldt voor de oncontracteerbare investeringen van Hart en Moore *cum suis*. Het probleem is alleen dat deze modellen nooit een grotere verklarende kracht kunnen hebben dan de economische wetenschap in het algemeen. Gegeven de grenzen die we in het inleidende hoofdstuk aan de economische wetenschap hebben gesteld, met name de aanname dat mensen op wat voor manier dan ook rationeel handelen om hun nut te maximaliseren, zullen er altijd fenomenen zijn die niet of slechts met moeite economisch verklaard kunnen worden. Semi-

publieke ondernemingen zijn van nature deels politiek en deels economisch, en zullen dus altijd maar deels economisch begrepen kunnen worden.

Curriculum Vitae

Martin Holterman was born in Emmen, the Netherlands, on November 16, 1981. He attended VWO (pre-university education) at the Gemeentelijke Scholengemeenschap in Emmen from 1992 to 1999. After graduating, he studied economics at the University of Groningen from 1999 until 2007, specialising in International Economics & Business and in Finance, as well as, from 2000 until 2004, International and European Law. After taking his degrees with thesis work on implied powers in European Community Law, on executive remuneration in US large-cap companies and on the transaction costs of a WTO-level competition law regime in 2004, 2005 and 2007, respectively, Martin started work as a Ph.D. fellow in the department of Legal and Economic Governance Studies of the Faculty of Management and Governance of the University of Twente in November 2005, which resulted in the present thesis. In the meantime, he also spent five months as a *stagiaire* at the Council of Ministers of the European Union in early 2006. Currently, he is a Research Assistant at the Transport department of the Florence School of Regulation, which is part of the Robert Schuman Centre for Advanced Studies at the European University Institute in Florence, Italy.

The Transaction Costs of Semi-Public Institutions

Railway History as a Challenge for Coase

In order to examine usefulness of the New Institutional Economics (NIE) for semi-public institutions such as public utilities and public transport, this thesis confronts the various “children of Coase” with a case study of the history of the railways in five European countries.

ISBN 978-90-365-3254-9