UNIVERSITE PARIS I- PANTHEON SORBONNE
Institut d'Administration des Entreprises de Paris

École doctorale de Management Panthéon-Sorbonne - ED 559
Equipe de Recherche GREGOR - EA 2474

Restructuring Railways in Europe: Regulation to Supplement Market Mechanisms

THESE
présentée et soutenue publiquement le **30 novembre 2016**
en vue de l'obtention du
DOCTORAT EN SCIENCES DE GESTION
par

Jean-Christophe Thiebaud

JURY

Directeurs de recherche :	**Mme Carine Staropoli** Maître de Conférences HDR, Université de Paris I Panthéon-Sorbonne
	M. Stéphane Saussier Professeur, IAE de Paris
Rapporteurs :	**M. Carlo Cambini** Professeur, Politecnico di Torino
	M Fréderic Marty Chargé de recherche au CNRS, Université de Nice-Sophia Antipolis
Suffragants :	**M Andrew Smith** Maître de Conférences, Institute for Transport Studies
	Mme Anne Yvrande-Billon Vice-présidente de l'Autorité de Régulation des Activités Ferroviaires et Routières

L'Université Paris 1 Panthéon-Sorbonne n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions doivent être considérées comme propres à leurs auteurs.

Cette thèse a été écrite dans le cadre d'un contrat CIFRE en partenariat avec l'Autorité de Régulation des Activités Ferroviaires et Routières. Les opinions exprimées sont propres aux auteurs et indépendantes de l'ARAFER.

Acknowledgements

Writing a Ph.D. dissertation is no easy task. Still, thanks to the people along the way giving you advice and help and support, it becomes such a fulfilling experience. For this experience alone I could be grateful, but I would then fail to acknowledge also how all those people pushed me to the finish. On both accounts, I want to thank them.

First of all, I would like to thank my Ph.D. advisors, Stéphane Saussier and Carine Staropoli for giving me the opportunity of doing a Ph.D. under their guidance. They were always here to give me suggestions or fruitful criticisms that kept me on the move whenever I was stalling. I also benefited a lot from the energy they put in creating an excellent and dynamic work environment for Ph.D. students at the Chaire EPPP, so that the ducklings can reach the shore.

I am equally grateful to Miguel Amaral for being my advisor at the ARAFER and being my co-author. He really helped me understand the requirements of academic research. Miguel provided me with crucial reviews and directions. He also made it easy to combine my work at ARAFER with my Ph.D.. On this note, I would like to thank the Authority for its financial support and the working conditions. I would also like to thank especially my managers at the ARAF(ER), Michel Vermeulen and Pierre Ravier who at the time accepted that I do a Ph.D. within their institution. I also benefited greatly from their

knowledge of railways and from the time they took to review my work.

I would also like to thank Fréderic Marty, Carlo Cambini Anne Yvrande-Billon and Andrew Smith for accepting to be members of my jury. But also for the comments they provided while I was writing my dissertation, way ahead of my Ph.D. defense.

I want to thank as well the other experienced researchers who kindly took time out of their busy schedules to listen and comment on my work during my Ph.D. For this, I am hugely in debt to Ricard Gil, Andrei Shleifer, Philippe Gagnepain, Yannick Perez and Matej Drev. I am grateful to the whole ESNIE team and, in particular, Eric Brousseau who made many of these meetings possible. I also would like to thank every scholar who discussed or commented on my presentations at any conference I attended.

I mentioned earlier the people that help you reach the finish line and want to thank especially John Moore and Zoé Le Squeren. I am very appreciative of all of John's teachings in Stata and econometrics, but also the reviews he made to my work. Hopefully I won't be constantly interrupting him in his work any more to ask questions - for that I am both thankful and sorry. But most importantly he has been a great person to be around at the workplace (be it at the Chaire, at the ARAFER, or where his phone thinks he works).

And I want to acknowledge the special role of my PhD-mate, with whom I've shared since day one many of the experiences of writing a dissertation or travelling to conferences. Zoé's cheerful presence has been a stupendous source of motivation, in particular towards the end when it all gets so bleak.

More generally it has been a great chance to do my Ph.D. with the all the members of the chaire EPPP on many levels. The welcome I received made me feel at home right away, because of the geek talks with Eshien Chong

and Michael Klien who instantly convey their passion for economics, of Jean Beuve's jokes during a break and the overall camaraderie maintained by everyone in the office: Anissa Boulemia, Lisa Chever, Aude Le Lannier, Alexandre Mayol, Tra Phuong Tran and Louise Vidal. Learning about their research during seminars has been very enriching, as well as their comments when I was presenting. Thank you especially to Eshien, Anissa and Aude for their feedbacks on this dissertation. I would like to mention as well all members of the chaire: Laure Athias, Paula Berdugo, Julie de Brux, Claudine Desrieux, Guillaume Fonouni-Farde, Simon Porcher, Emmanuel Raynaud and Maria Salvetti. I also wish best of luck to the current Ph.D. students: Charles Bizien, Oliver Cristofini, Marion Fernet, Ouiam Kaddouri and Yoan Robin.

I have also greatly benefited from my time at ARAFER, learning what is a train path, accounting separation and what it is to regulate. It was an endless source of inspiration for my Ph.D. I also thank Mathias Laffont who took the time to read and comment my dissertation. On a more personal note, I've met there a lot of great people, and it was a nice experience to share with them the commute to le Mans and its occasional *wagon-bar*. And since I have mentioned this ordeal, I need also to mention my fellow *piou-pious*: Olga Bohorquez-Suarez, Amir Chemli, Sibylle Couffignac, Nicolas Lluch, Jérémie Pieté and Marion Sallandre, for raising the mood at work.

My thoughts also go to my friends in the *civilian* life for their support during those three years and for asking me if my dissertation *avance*. It did help me come to the realisation that it won't be writing itself. I would like to thank especially my flatmate during (part of) my Ph.D. Paul Meunier who is probably the one I talked to the most about trains. And for his insights on the regulation of the energy market. During that time, I appreciated a lot the efforts of Marine to make us eat healthy, and the numerous occasions to relax with Thomas, Jean-Baptiste, Pierre-Louis and Paul. Thank you also to Clarisse, Lucile, Morgane, Constance, Julie, Anne Laure, Boris, Pauline,

Alexandre, Clémentine, Coline and Aurélien. I would like to thank as well my friends from Dauphine: Valentin, Emmanuel, Clement and Lina who were the first to enable my fondness for economics and for the much-appreciated annual ski trips.

Last but not least I would like to thank my family, my parents, brother, sister and the new members that joined during my dissertation. Thank you for for the proof reading, the meals on Sundays and being there even if you are not in Paris! This work is dedicated to you.

Jean-Christophe

Foreword

This Ph.D. dissertation is made up of three chapters. The links between them and the underlying logic of the whole dissertation is explained in the General Introduction. Nevertheless, since each chapter corresponds to an independent study, they can be read separately. This implies the presence of redundant information.

Abstract

Restructuring Railways in Europe: Regulation to supplement market mechanisms

In an objective to revitalize the sector, railways in Europe have gone through changes in order to introduce competitive forces in an industry previously dominated by state monopolies. Not unlike what occurred in other network industries, the upstream management of infrastructure -considered a natural monopoly- was separated from the downstream market of train services which was deemed contestable. But the experience from previous reforms shows that the markets stemming from this new organisation are made and don't necessarily just happen.

This dissertation preys on those changing times to analyse how such transition can be achieved. More specifically, it uses the framework developed by the Theory of the Firm to investigate some of the challenges that have risen from this new organisation due to vertical separation and the need to develop coordination. Then using the tools developed by Agency Theory, the dissertation investigates the introduction of tendering in public procurement for rail services. Throughout the analysis, an emphasis is put on finding how regula-

tion can provide a remedy to the identified challenges.

In chapter 1, we review how the reforms were designed in Europe. It highlights that there was very various applications across countries on several dimensions such as the vertical structure, regulation or downstream competition, triggering debates on the relative merits of each option. Furthermore this plurality questions the motivations behind these choices. And although those motivations are brought up in the literature, the previous empirical studies do not take them into account when comparing the relative performance of one form against the other. In such a case, endogeneity might come and biased the results of econometric regressions. Having constructed a database covering four years between 2009 and 2012 in 25 European countries we use a two-stage-least-square model to obtain an unbiased estimate of the effects of the above mentioned dimensions. In conclusion, we find that endogeneity does create a bias, in particular when measuring the merits of full separation.

Chapter 2 builds on the French rail sector's example to shed the light on the crucial and understudied impact of coordination costs, one of the drawbacks arising with separation in the sector. We develop a model explaining why inefficient outcomes may arise in the railway sector when vertically separated firms have to commit ex ante on quantities. Our results indicate that credible and effective price regulation can overcome the limits of separation on the infrastructure side. On the other hand, if the market is not flexible enough, it may become harder, as the downstream market is becoming more competitive, for train operators to make optimistic production forecasts.

In chapter 3, we focus on the downstream market and, what should become the norm in Europe: the tendering of public service contracts in railways. Drivers of future performance are the expertise of the public buyer -in particular its ability to monitor the contract- and contractual design. To assert the effect of these two drivers, we rely on a stochastic cost frontier model and analyse the

efficiency of PSO contracts for railway regional transport in France between 2009 and 2012. The empirical results highlight that the incentive properties of contracts do not have the expected effects: fixed price schemes do not increase the efficiency of procurement and performance decreases during the contract execution with a ratchet effect occurring at the end of the contract. Beyond those results, we suggest that a regulator centralizing information on behalf of public buyers would increase the efficiency of the process, in the spirit of yardstick competition.

Key words: Vertical separation, Regulation, Public Procurement, Railroads, network Industries, Econometrics

Résumé

Réformes des chemins de fer en Europe : la régulation en complément des mécanismes de marché

Afin de trouver un nouveau dynamisme, l'Europe des chemins de fer a entrepris d'importants changements visant in fine à introduire une pression concurrentielle dans un secteur précédemment dominé par des monopoles étatiques. A linstar des changements ayant eu lieu dans les autres industries de réseaux, la gestion en amont des infrastructures - considérée comme un monopole naturel - a été séparée du marché aval des services ferroviaires, considéré lui comme contestable, c'est-à-dire propice à l'introduction de la concurrence. Néanmoins l'expérience de réformes similaires nous rappelle que les marchés découlant d'une telle réorganisation doivent être manufacturés et ne se forment pas spontanément.

Dans cette thèse, nous capitalisons sur cette période de changements afin d'analyser comment une telle transition se réalise. Plus spécifiquement, en appliquant le cadre développé par la Théorie de la Firme, les défis liés à la séparation verticale et au besoin de coordination sont examinés. Puis en s'appuyant sur les outils de la Théorie de l'Agence, nous étudions les implications dé-

coulant de l'introduction d'appels d'offres dans la délégation de service public pour le transport régional de voyageur. Au long de la thèse, l'accent est mis sur le rôle de la régulation et sur les remèdes qu'elle peut offrir aux écueils rencontrés.

Le chapitre 1 analyse comment les réformes ont été conçues à travers l'Europe. Car en pratique la mise en oeuvre de cette nouvelle stratégie ferroviaire varie grandement d'un pays à l'autre à plusieurs niveaux, tels que la structure verticale, la régulation ou encore la libéralisation du secteur aval, engendrant d'importants débats sur les mérites respectifs de chacune de ces options. Cette pluralité interroge également sur les motivations sous-jacentes à chacun de ces choix. Et bien que la littérature économique aborde ces motivations, les études empiriques précédentes ne les prennent pas en compte quand il s'agit de comparer les différentes formes verticales. De ce fait, l'endogénéité peut venir biaiser les résultats des estimations économétriques. Ayant construit une base de données couvrant les quatre années entre 2009 et 2012 pour 25 pays européens, nous appliquons la méthode des doubles moindres carrées (2SLS) pour obtenir une estimation non-biaisée des effets de chacune des dimensions précédemment évoquées. En particulier, nous trouvons que l'endogénéité crée effectivement un biais qui péjore les effets mesurés de la séparation verticale.

Le chapitre 2 sappuie sur le cas du secteur ferroviaire français pour illustrer une des répercussions souvent sous-estimées de la séparation verticale, celle des coûts de coordination. Nous développons un modèle expliquant pourquoi la coordination peut aboutir à des résultats inefficients dans un secteur ferroviaire où les firmes doivent sengager *ex ante* sur les quantités. Nos résultats indiquent qu'une régulation crédible peut permettre de surmonter les limites de la séparation verticale pour le gestionnaire d'infrastructure. Cependant dans un marché qui n'est pas suffisamment flexible, il devient plus dur au fur et à mesure que le marché aval s'ouvre à la concurrence de soutenir une production importante.

Enfin dans le chapitre 3 nous nous intéressons plus particulièrement au marché aval et ce qui devrait devenir la norme avec l'introduction de la concurrence: la mise en place d'appel d'offres pour les services régionaux de transport de voyageurs. Parmi les moteurs d'efficacité pour une autorité organisatrice se trouve l'expertise qu'elle peut mettre en oeuvre et la conception des contrats. Pour en mesurer les effets, nous avons constitué une base de données sur les contrats TER en France entre 2009 et 2012. Les résultats obtenus suite à l'estimation d'une frontière stochastique montrent que la clause mettant en place un prix plafond ne conduit pas à une réduction du prix acquitté par lautorité organisatrice mais entraîne au contraire la création d'un effet cliquet en fin de contrat. Au-delà de ces résultats, la mise en place d'un régulateur centralisant l'information et apportant son expertise sur les coûts aux acheteurs publics devrait améliorer l'efficacité du processus, dans un esprit de concurrence par comparaison.

Mots clés : Séparation verticale, Régulation, Marchés publics, Chemins de fer, Industries de réseaux, Econométrie

Contents

General Introduction

Despite being considered a source of sustainable mobility, Europe witnessed a decline of its railways since the 1970's. Between 1970 and 1990 alone, the modal share of passenger services declined from 10.3% to 6.9% while the modal share of freight services decreased even faster, shrinking from 31.7% to 18.9%[1]. The European Commission attributed this downfall to the discontentment of consumers due to prices and quality and proposed a solution to revitalize railways across Europe. This strategy[2] consisted in introducing market forces to replace state monopolies that where the norm in the sector. A requirement was the split between infrastructure management and the train operations in order to introduce competition in the downstream market of train services. This reform, which shares similarity with those that occurred in other network industries, started with directive 1991/440/EEC and is still ongoing. Currently under discussion, the fourth railway package should be the last component of the legislation and introduce competition in national passenger services in the 2020's.

[1] Source: *A Strategy for Revitalising the Community's Railways*, COM(96) 421 final, Brussels, 30.07.1996.

[2] Made explicit in a white paper published in 1996. See *A Strategy for Revitalising the Community's Railways*, COM(96) 421 final, Brussels, 30.07.1996.

The emergence of a new model

Enforcing vertical separation has triggered many debates on the organisation of railways over the recent years. But it is worth noting that it is not a new topic for economists. On the contrary, similar discussions took place back at creation and development of railways in the middle on the 19^{th} century. The first item to be debated was to know if there could or should be competition in the railway market. Interestingly the debate occurred on both side of the Atlantic and reached the same intermediary conclusion: competition was not possible between railways. To be more precise, the question was to know if there could be competition between integrated railways. Competition on the tracks had been quickly ruled out, due to technical reasons (rather than economics arguments) as summed up by Walras et al. [1897] (cited in Perennes [2014]). The implications of this impossibility are going to be different, and according to Giocoli [2014], influence economists' view on competition policy, in different ways.

In the case of the USA, Giocoli [2014] offers the following story: the emergence of railways created a shift from the point of view of the *Classics*. Indeed going back to A. Smith, it was considered that as long as freedom of trade and freedom of contract were respected, then any monopoly could only be temporary and higher profits in a market could not be sustainable in the long run. Yet in the case of railways, due to the presence of sunk investments as described by Hadley [1897], that is "*a large permanent investment, which can be used for one narrowly defined purpose, and for no other. The capital, once invested, must remain. It is worth little for any other purpose*", competition cannot be viable in the long run and will lead to some form of market power, such as ruinous competition, predatory pricing or the creation of a cartel. The response made by the legislator was to create the Interstate Commerce Commission (ICC) in 1887 in order to enforce just and reasonable shipping rates. Moreover, according to Giocoli [2014] the emergence of naturally big firm led,

under the impulsion of J.B. Clark, to the creation of modern antitrust law in the United States. Namely the creation of a jurisdiction, composed not only of jurists but also economists that could ban practices. This led to the emergence of a type institution that will take time to cross the Atlantic: the independent regulator[3].

In France, Perennes [2014] notes that the debate was first brought on by Dupuit [1853] who insisted that railways are *de facto* a monopoly due to the importance of investment needed to create a railway while Walras et al. [1897] stressed the issue of scale economies. The idea that railways are a "natural monopoly" is going to spread around the 1870's. The notion of natural monopoly would encompass several criteria: barriers to entry, economies of scale and among politicians, Perennes [2014] points out that the determining criterion was that it was not socially beneficial to duplicate railway infrastructure. Thus, it was assess that railways where essential facilities. This assessment, along with the idea developed by Walras et al. [1897] that railways had a "moral monopoly", that is they contribute to the economic development, was going to lead to the nationalisation of railways in France[4] after it was taken over for instance by Leon Blum in 1919, who considered that railways were a *public service*.

In Europe, state owned monopoly became the norm in railways, with this organization carrying on until the 1980's. From this moment on, it was gradually reconsidered at the same time as the concept of natural monopoly was being re-examined in the economic literature. In particular we can stress the influence of the theory of contestable market on regulation of railways and more generally network industries. Baumol et al. [1982] developed a framework to analyse what the efficient industry structure should be given the cost structure in a multiproduct context. Contrary to perfect competition, this framework takes into account scale economies. In a nutshell, as long as hit-

[3] Although the ICC was chosen by Stigler [1971] to illustrate the issue of regulatory capture, making it ultimately a poor example of regulatory independence.

[4] 1937 in the case of France.

and-run strategies are possible, the industry structure will be efficient even in the absence of regulation and in the presence of scale economies. The possibility of hit-and-run strategies make the market contestable, but such strategy can be hindered either by the need to make large irreversible investments to enter the market, making exit costly, or by regulation which defers entry.

As far as railways goes, the assessment remains so far the same as a century ago. The presence of large sunk cost to enter the market make any hit-and-run strategy impossible. Therefore *a priori* the market remains not contestable. Yet Bailey and Baumol [1983] suggest that a possible approach "*is to isolate the sunk investments, leaving a relatively contestable part of the industry's operations to be controlled by market forces, while the portion with substantial amounts of sunk capital is regulated or even operated by the public sector*". This approach was not *per se* new. After all Bailey and Baumol [1983] illustrate the proposition they make with the reform that had just taken place in the US in the telecom sector. The local networks necessitating fewer sunk investments had been opened to competition, while the long distance network remained under the monopoly of AT&T. But the contestable market theory laid the theoretical foundation for unbundling network industries.

In concrete terms, the railway monopolies in Europe could be split in (at least) two firms as depicted in figure 1. Upstream is the infrastructure manager: a firm in charge of providing access to the rail network, and the reform driven by the European Union was not to affect its natural monopoly status. Downstream, are the train operators (or *railway undertakings* in European lingo) which operate commercial services on the network and that could be (progressively) opened to competition. Respectively in 2003 and 2007 international and national freight services were opened to competition, while international passenger services opened *de jure* to competition in 2010. It is worth noting that both for freight and passenger services, international services were opened to competition first in order to fast track the creation of a common market

for rail services. And as of this day, the liberalization of national passenger services have not been made mandatory in European law, although some countries anticipated it.

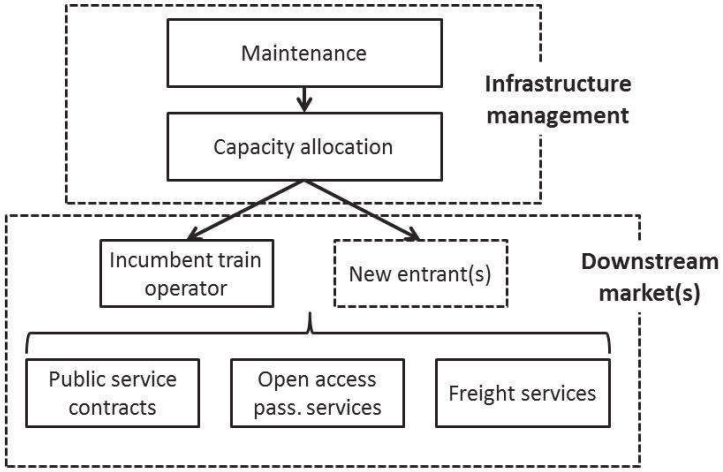

Figure 1: Organisation of the railway sector

The challenges

This dissertation is dedicated to studying some of the implications of such a reform. Yet, at this point it is worth reminding that Walras et al. [1897] considered it impossible to separate infrastructure from downstream operations. Although technological progress might explain how such a separation is now possible, Bailey and Baumol [1983] point out that one of the condition for this reform to be successful is that it does not prevent *the realization of economies of scope* pinpointing that separation may have an overall cost. But economies of scope might be too broad a notion to describe the issues raised by the unbundling of network industries. At least from a point of view of

an economist, the black box that was the firm has been opened since Walras et al. [1897] made their claim. In particular, a theory of the firm was developed following the seminal article by Coase [1937] who addressed the make-or-buy question, that is to know why some transactions are more efficiently made in a firm than in a market depending on organizational constraints. Gibbons [2005] gives a very good overview of the four main theories of the firm that have since been developed and what vertical separation may trigger. They are succinctly summed up bellow.

A first approach is to focus on **rent-seeking** behaviours. Williamson [1971, 1975, 1985] and Klein et al. [1978] point out that a transaction leads to the creation of an appropriable quasi-rent which can be described as the opportunity cost of switching partners once specific investments have been made. Although it is not socially productive, in the absence of integration, conflicts will arise between the parties involved which will try and appropriate this rent. Despite some differences, the notion of specific investment echoes the one of irreversible investments developed by Baumol et al. [1982] in the case of network industries and could lead to rent-seeking behaviours from the firms.

A second approach was based on the distribution of **property rights** by Grossman and Hart [1986] and Hart and Moore [1990]. This theory takes its foundation in the fact that there exists non-contractible specific investments. The distribution of ownerships rights determines the share of the surplus one of the two parties may demand once the transaction has happened and by that the incentive to invest *ex ante*. In the absence of integration, investments may be hindered if it is important to maximize one party's investment in particular. In the case of network industries, quality, when it is not easily measured, can be considered as a non-contractible investments. Therefore unbundling could lead to quality issues if such investments rely heavily on one of the two parties.

It is also possible to view the firm as an **incentive system**. Following the

lead of Holmstrom and Milgrom [1991] and Holmstrom and Tirole [1991] it can be argued that with separation, the value of the asset at the end of the transaction is going to be a major factor in the decision made by an agent, in addition to the payment it receives for the task undertaken. When various means can be used to achieve a task, the agent will look to increase, or damage less the value of its asset, some time at the expense of the principal. Similar concerns can be found in network industries in the trade-off between light maintenance which can lead to a slight decrease in performance of the network and renewing it, which is more costly and more time consuming but keeps the network at higher performance level.

Finally, a corollary of rent-seeking behaviours is the concern for **adaptation**. The theory developed by Williamson [1975, 1985] and Klein and Murphy [1988] suggests that the fear of rent-seeking behaviours might hinder decision making in the absence of integration. On the contrary, integration allows more possibility of adaptation in an uncertain environment. Within a firm, a manager can use its authority to redefine the task as the state of nature is observed instead of having to lead a costly renegotiation of the contract. Coordination becomes therefore more flexible. For instance, we will argue later in this dissertation that when demand is uncertain, vertical separation entails new challenges to match demand and supply in network industries.

All four theories of the firm give us the insights to understand why it is not possible to replicate the same mechanisms, once the industry has been unbundled, and therefore help us comprehend what might be the drawbacks associated with vertical separation. Such a questioning has occurred in the other network industries where similar reforms have been applied. Besides the example of telecoms with a separation between long distance and local network, the electricity sector has been similarly opened to competition. To do so, there was a split in four activities: generation, transmission, distribution and retail. This separation has allowed to isolate monopolist activities -transmission

and distribution- from competitive ones - generation and retail. But given the complex nature of the electricity market (Joskow and Schmalensee [1983]), restructuring it and introducing competition did not lead to less regulation. And as pointed out by Hogan [2002]: "*power markets are made, they don't just happen.*"

In the case of electricity, this complexity lies mostly in the nature of the good which cannot be stored which imposes to balance supply and demand permanently to avoid local failures that may entail great consequences (see Stoft [2002] for instance). This leads necessarily to what Hogan [2002] calls coordination for competition. In that sense we can draw a parallel between electricity and railways. In both cases, one of the key aspect of an unbundled network lies in the management of capacity constraints. Although such challenges do not become trivial within a firm, they might be exacerbated with vertical separation and requires the development of complex market mechanism. On the one hand, the network must be able to sustain a high level of production during the peak-hour but on the other hand superfluous capacity has very low value when it is not null. With no mechanism to coordinate the capacity of production, investing in a competitive electricity market would be very risky as pointed out by Newbery [2002].

Although railways do not suffer from the stringent constraint of a massive black-out, it does face similar capacity management issues and a gap between peak and off-peak capacity. Since a train cannot overtake another one in the absence of dedicated facilities, the speed of the train is going to impact on the capacity it consumes and a slow train can consume double the capacity of a fast train. Coordination is then needed to satisfy the rival uses of the network between freight and passenger services for instance.

In drawing the parallel between both industries, we notice that finding the balance between regulation and liberalisation has been a long process. Two

accidents at the beginning of the 2000's, respectively the Hatfield accident and the California crisis were in that sense turning points and marked a slow down in the on-going reform process. The Hatfield accident occurred on October, 17^{th} 2000. The derailment was caused by the fracturing of a rail, itself attributed to rolling contact fatigue. The privatised infrastructure manager, Rail Track, was pointed out, and given the safety concerns in railways, the British government back pedalled and put in charge a not-for-dividend company, Network Rail, to manage the network. And in Europe, infrastructure manager remained state-owned companies. Regarding the California electricity crisis, it unravelled between 2000 and 2001 and can be attributed according to Hogan [2002] to "*the product of a volatile combination of bad economic theory and worse political economy practice.*" As a consequence, the rise of spot market prices and fixed retail rates drove major electricity wholesalers to bankruptcy while another wholesaler became notorious for gaming the market. The crisis called for state intervention and highlighted the consequences of a bad market design which consisted in "*boundless faith in the ability of markets to solve all problems*" (Hogan [2002]).

If one should not have a boundless faith in the market, Staropoli and Yvrande-Billon [2009] show that for both industries in Great Britain, when the attributes of the transaction and of the governance structure were not aligned, the characteristics of the assets were changed when possible by the players in the industry. This is in line with the predictions of Riordan and Williamson [1985]. And it is the reason why, according to the authors, the electricity generators invested in natural gas thermal power station which are more flexible. Or why there was a big increase in rolling stock investments and in particular in a more standardized rolling stock which could be easily redeployed. In other words, after separation occurred, innovation made it possible to avoid some of the associated drawbacks. Also according to Staropoli and Yvrande-Billon [2009], the reforms that took place in the UK can be considered radical. In the case of electricity, the most extreme option was systematically adopted and

the de-integration took place in less than six years. In the case of railways, the authors point out that the vertical monopoly was shattered into short term bilateral contracts as no other country has done since then. This radical change might have allowed the emergence later on of a more efficient industry structure. As it happens the stakeholders moved toward more integrated structure.

All in all, the literature does anticipate drawbacks to unbundling that are more precise than the broad notion of economies of scope. And if markets can in some cases adapt, it is not necessarily the effective solution as pointed out by previous cases of unbundling. The time period which is studied in this dissertation is very interesting as the railway industry, particularly in France, the transition from vertical monopolies to unbundled markets is not over. The necessary market design is still largely imperfect and the shift is not over. Beyond the possibility to witness first hand cases of misalignment, it is also an era which is suitable for policy implications from the point of view of an economist. And if the technical implications of each network industry differ, it can be argued that a common trend exists in the way reforms are adapted in order to reach a more steady state.

Independent regulation in railways

The importance of market design was stressed but so far no mention has been made of the implementation of a new player that emerged during the reform: the regulator, designed as an independent regulatory agency (IRA). Since Leibenstein [1966] the idea that in the absence of competitive pressure, a monopoly will develop inefficiencies has been deeply rooted in economists' minds. It has since fell upon the regulator to set in place a cost reimbursement scheme that gives the right incentives to the monopoly as illustrated by the seminal work of Laffont and Tirole [1993]. A first task of the regulator is

therefore to promote efficiency. But the regulators in railways where usually not given the role of enforcing efficiency as described later in the first chapter of this dissertation[5].

In the case of railways, more emphasis has been put on a second task attributed to the regulator, that is creating a level playing field for competition. This task can be found in article 56 of directive 2012/34/EU stating that the regulator "*shall, in particular, check whether the network statement contains discriminatory clauses or creates discretionary powers for the infrastructure manager that may be used to discriminate against applicants.*" In their survey, Benedetto et al. [2015] find that the promptness on tackling problems related to non-discrimination issues is satisfactory to the firms on the downstream market, suggesting that regulators are indeed tackling the issue at hand.

One of the key characteristic the regulator brings to the table is its independence. Since Levy and Spiller [1994], the idea that commitments made by an independent agency are more credible, has been widely accepted. The independence therefore becomes an important feature to give firms the right incentives for investments in a context of privatization and liberalization. In the case of access regulation as in railway, the position of the regulator can be interpreted as twofold as described by Stern [1997] when it comes to investments. It has to be credible in the eyes of new entrants on the downstream sector, especially when the incumbent firm is owned by the government, and convince them that fair competition will be promoted. It also has to reassure the infrastructure manager that, in a context of sunk investments, a fair return on capital expenses will be awarded. On this matter, Cambini and Jiang [2009] review the various implications regulation has on investment's decision of firms and the possible adverse effects of the different regulatory regimes.

[5] As pointed out by Benedetto et al. [2015] only in few countries does the regulator play an active role in controlling efficiency. This is confirmed by Cherchye et al. [2015] which points out that only the regulator in Great Britain uses efficiency as a tool to regulate the infrastructure manager. Further details on the process that has to be put in place can be found in Stern [2013] in its review of the Office of Rail and Road (ORR) process.

But independence was not established from the start. As pointed out by Crozet et al. [2012] three models of regulation had emerged at the beginning. In Great Britain, the ORR has had the characteristic of an independent regulator from the start, in charge of reviewing tariffs and capacity allocation with a specialization in transports. In some countries such as Germany, regulation was entrusted to an agency already in charge of supervising other network industries. Finally in France and Italy, before the creation of an IRA, regulation started under the supervision of the ministry of transport, a common trend for all sectors according to Thatcher [2002]. The fact that some regulators are fairly young is not neutral. For instance in the case of electricity, Newbery [2002] points out that countries in mainland Europe had little time to prepare for such a shift, making it "*improbable that all member countries can painlessly accommodate their institutions and concerns to the ideals of a politically independent regulatory system.*" According to the author, this is a substantial difference compared to the US where liberalisation took place within a well defined regulatory framework.

The fast shift in institutions, and more generally the little experience of the regulator comes as a challenge to its independence and its credibility. In his review on regulatory capture, Dal Bó [2006] offers many cases of capture that might affect particularly a young regulator. A regulator might be, as the author puts it, out-consulted by a regulated firm and be convinced by (false) arguments. Also, in Leaver [2009], the fear that a firm may publicly complain about a mistake made by the regulator will buy regulated firm some regulatory slack. Therefore, in the first year of its creation a regulator might be cautious in the decisions it issues. A new regulator might also be particularly subject to revolving doors. A regulator that needs to gain quickly technical knowledge on the sector might need to recruit some former employees of the regulated firm. And even in the case of well meaning individuals, it might create a bias to worry too much about the concerns of their former firm, which they are

more familiar with and may have more empathy.

The role of the regulator goes beyond being a safeguard for the firms against political decisions. The various theories of the firm we have briefly described before all point to the fact that the drawbacks of vertical separation are linked to *ex post* commitment issues making coordination more difficult. A regulator in its role of appeal body can contribute to reduce opportunistic behaviours and enforce more complex contracts. On the contrary to a non specialized jurisdiction, a regulator should be able to provide its expertise in solving sector specific market failures and do its bit toward an efficient market design.

Outline of the dissertation

The objective in this dissertation is to addresses the question of vertical separation in railways, but the approach chosen is not to try and determine the overall relevance of this organisation, rather to identify the vulnerabilities that stem from the shift to a vertically separated structure. In order to do so, we make use of the theoretical framework of the Theory of the Firm to understand what are the changes implied and how they may deter the market outcomes. The second step is to find some solutions. In other words, can the unbundled railway markets be made both contestable and efficient ? We find, as in other network industries that an outside intervention may be needed and in this dissertation a special emphasis on the regulatory mechanisms that should be enforced.

This analysis appears as opportune on several accounts. First on the timing: despite the fact that the reform was introduced more than two decades ago, railways in Europe remain riddled with cases of (temporary) misalignments between the characteristic of the transaction, and the governance structure.

Therefore the vulnerabilities mentioned earlier have not been addressed. Yet, as competition should soon be introduced in the common market, correcting those market failures becomes a more pressing concern. Secondly, Europe is an interesting case to study given the heterogeneity both in the pre-existing setting and in the manner the reforms where conducted. It enables to understand the arbitrage behind the choice of a structure, but also the effects it entails. Last, this dissertation was written along while working for the French regulatory body (ARAFER), granting us the opportunity to have a direct access with the stakeholders and their concerns which had an influence on the topics tackled.

Given the complexity of railways and of the interactions that take place, our approach cannot be exhaustive. Nevertheless, the dissertation deals with different stages of the production chain as presented in Figure 2, with the hope of providing an overview of the challenges that railways face. The first two chapter focus on infrastructure management and its interaction with train operations. In the third chapter, we look at the downstream sector. Indeed the reform of railways rest upon the assumption that introducing competition has a rejuvenating effect. Such an effect is not straightforward. Competition in railways can both be in the market or for the market, the latter being the norm for public services and represents 70% of national passenger services in Europe[6]. Well, as pointed out by modern economic theories of procurement, the introduction of tendering will raise both *ex ante* concerns as summarised in Laffont and Tirole [1993] as well as *ex post* challenges, described for instance in Bajari and Tadelis [2001], which are not necessarily consistent with an efficient outcome.

On the many stages of production, we find that restructuring railways in Europe led to the creation of misalignments. They might be naturally corrected in

[6] Fourth report on monitoring development in the rail market, SWD(2014) 186 final, Brussels. Later referred to as RMMS 2014

time. But, as it has been the case in the electricity sector, those misalignments may also call for regulation given the complex nature of the transactions, or that the markets are not sufficiently contestable.

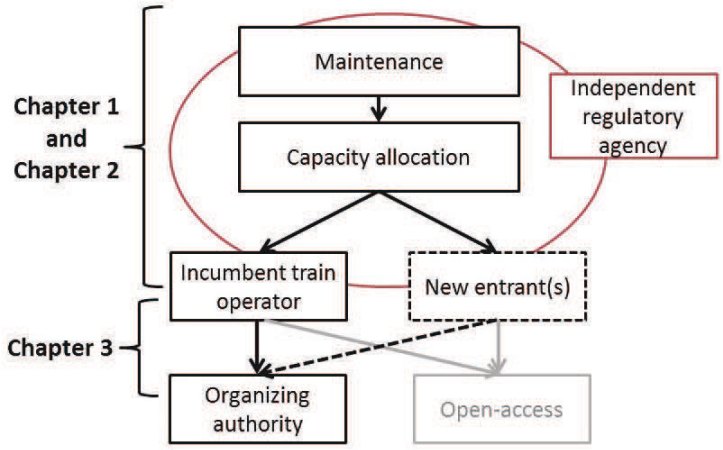

Figure 2: Outline of the dissertation

The rest of the dissertation is organised as follows. The first chapter of this dissertation studies more in depth the reform in Europe, and builds on the heterogeneity of policy choices made by European countries to assert the motivations and implications of such choices. In particular, we observe various structure where chosen and that they correspond to alternate means of coordination. In the second chapter we address more precisely this issue, that is how coordination between maintenance and commercial use may be achieved between the upstream and downstream firms in a unbundled sector. Those results help us identify conditions for downstream competition to be viable. The third chapter focuses on one of the downstream markets to see how competition can be beneficial. It studies the current regulation of public procurement for regional passenger services and its limits due to the limited expertise of a public.

Table Outline

In Tables 1 and 2, we summarize the research questions, the data and the methods used as well as the main results from each chapter of this dissertation.

Table 1: **Research Questions, Methodology and Main Results: Summary of Chapter 1 and 2**

Research Questions	**Methods and Data**	**Main Results**
• **Chapter 1**: Restructuring infrastructure management in Europe: The plurality of reforms • **Research Questions**: How to explain the plurality of structures and what are the expected effects over infrastructure management efficiency ?	• Econometrics: OLS with RE (Panel) and Instrumental variables. • Data from RMMS and Eurostat • Dataset on the cost of infrastructure management for 25 countries between 2009 and 2012.	• The bias created by endogeneity tends to weaken the arguments in favour of full unbundling in railways. • Having created a regulator more ahead of the reform decreases the costs of the infrastructure manager.
• **Chapter 2**: Vertical Separation in Rail Transport: How Prices Influence Coordination • **Research Questions**: What are the challenges of coordinating supply and demand of capacity on the network ? Can coordination be achieved using prices ?	• Study of the French capacity allocation process • Game theory : Normal form game	• Uncertainty on final demand motivated the design of a flexible allocation process for capacity, which *de facto* limits commitment • Uncoordinated market outcomes generate a loss of performance for railways. • Through price regulation the infrastructure manager can be steered toward a coordinated outcome. • The success of such a regulation depends on the ability of the downstream market to bear mark-ups.

Table 2: **Research Questions, Methodology and Main Results: Summary of Chapter 3**

Research Questions	Methods and Data	Main Results
• **Chapter 3**: The Role of Expertise on efficiency in Public Procurement: The Case of Regional Railway Transport in France • **Research Questions**: Do we observe heterogeneity in efficiency across public buyers and can it be linked to expertise-related moral hazard issues ?	• Econometrics: Stochastic Frontier Analysis. • Dataset for regional railway transport in France between 2009 and 2012 for the 20 regions.	• Significant efficiency differences between the regional local operators. • Efficiency decreases as the share of non-verifiable cost increases and over the duration of the contract. • Those adverse effects can be linked to the absence of competitive tendering but also a lack of monitoring expertise of the public buyer and • To reduce the information asymmetries, the public buyers could centralise information and apply yardstick competition

Chapter 1

Restructuring infrastructure management in Europe: The plurality of reforms*

1.1 Introduction

In the last two decades, major structural reforms have been implemented in most network industries. Considerable attention has been devoted by economists to the analysis of gas, telecom or electricity industries but less attention has been paid to railways so far. Yet, driven by European institutions, the railway transport sector in Europe has also gone through major institutional and organizational reforms during the last twenty years.

The objectives of the reforms were clearly identified by the European Commission[7] and are summed up by the following paragraph : "*The railway sector is in decline and its market share is falling. Rail is felt not to respond to market changes or customers' needs. However, rail has characteristics which could make it an increasingly attractive form of transport in Europe. Many*

* The author would like to thank Miguel Amaral, Anissa Boulemia, Eshien Chong, Mathias Laffont and John Moore for their comments.

[7] Commission White Paper of 30 July 1996: "A strategy for revitalising the Community's railways"

possibilities already exist for improving and developing services, and new areas of opportunity may open up. To meet these challenges, the Community needs a new kind of railway."

The two means for such a change are identified as the following :

- "*Introducing market forces into rail: Strengthening the market will give management and workers incentives to reduce costs, improve service quality and develop new products and markets.*
- *It required the separation of infrastructure management and transport operations into distinct business units, with separate management and balance sheets;*"

A third and implicit step in this process is the introduction of regulation, both through European legislation and the introduction of a regulator. Directive 91/440/EC was the first milestone to this process and introduced a degree of vertical separation in the sector. It has required an accounting separation between the management of essentials facilities (i.e. the management of the railway network) and the operation of rail services, which were deemed potentially competitive given the smaller barriers to entry. Since then the railway sector which consisted of vertically integrated monopolies has progressively unbundled. In this framework, vertical separation of infrastructure management is considered as a requirement and not as an end. And to our knowledge, no strong empirical evidences suggest that the overall impact of vertical separation both on efficiency or on consumer surplus is positive (or negative) in the sector.

As Directive 91/440/EC allows for different degrees of vertical separation, a consequence is that different governance modes coexist today in Europe depending on the willingness of each Member State to effectively reform its railways. This heterogeneity, both in drive and actions undertaken, raises the question of the comparative merits of the different vertical structures characterising rail transport in Europe, which as we will argue correspond to different

degrees of vertical separation. Economic literature in its various strands gives a better insight into the problems encountered by the choice of vertical structure. But whatever the structure chosen, a key point we focus on is that this reform led to the creation of a new player which is the infrastructure manager who has a central role in the performance of the railway sector as the sole supplier of rail capacity to the downstream firms, that is it provide an essential input to train operators. Due to the fact that infrastructure managers are quite new in the rail landscape, few studies have focused on drivers of their performance.

Further institutional changes have been made with this reform. In order to regulate infrastructure management, vertical separation was followed by the introduction of new institution in this market: independent regulatory agencies. Once again, the date of creation of the regulator varies from one country to another. There is a gap of more than twenty years between the creation of the first and last independent regulator. The third step of the process - competition - is being gradually introduced on the downstream market. European law has opened the market to competition for international passenger services and freight transport services but in most countries, it represents only a marginal share of the traffic. As of this day, there is still no mandatory competition for national passenger services, although some countries have anticipated this reform. Therefore the development of competition across Europe is at various stages depending on the country.

In this chapter, we analyse the drivers of the performance of infrastructure managers across Europe. A special emphasis is put on the organizational and institutional choices made by member states, not only the choices made for the vertical structure, but also those made in terms of regulation and competition. The second section is a literature review on the implications of vertical separation and presents how the choices made by Member States varies from one country to another for the three steps of the process we mentioned. Based

on those findings, we build predictions on how those characteristics affect the performance of infrastructure managers across Europe. In the third section, we put to the test our predictions using a dataset consisting of 25 European countries between the year 2009 and 2012 and assess the impact of the various choices on the maintenance and renewal costs of the infrastructure manager. In our empirical strategy, we focus our efforts on taking into account the endogeneity issue. Indeed, given that the overall objective of the railway reforms was to increase the performance of the railway sector the results might otherwise suffer from a bias due to reverse causality. To address the issue of endogeneity, we estimate two stage least squares regression using instruments derived from electricity markets.

1.2 Governance in the European rail sector

1.2.1 Vertical separation

Firstly, and before presenting the various vertical structures in Europe, we review in this section the theoretical implications of vertical separation. A large body of the literature in Economics has analysed the pros and cons of vertical separation in network industries, especially from an industrial organization's perspective (e.g. Vickers [1995], Sappington [2006]). Sources of relative performances of vertical separation and vertical integration are numerous and difficult to isolate but can be classified into three main dimensions: competition effect, production cost synergies and transaction costs.

A main driver for vertical separation relies on the need to improve competition in the downstream market (Sappington [2006]) and unbundling is often considered as a requirement for fair competition. Indeed, a vertically integrated company has the incentive to exploit its position to protect its competitive advantage and deter new entrants. It could translate, in particular, into lower

infrastructure service quality and/or higher infrastructure access charges for (potential) competitors. A vertically integrated monopoly may, for example, be able to impose excessive delays to the access of inputs to new entrants on the downstream market. This classic drawback associated with vertically integrated settings is well summarized by Reiffen and Ward [2002]. The authors recall that well-established economic principles indicate that a regulated monopolist with an affiliate in an unregulated business may have an incentive to deny the affiliates competitors access to an 'essential' input, or more generally, degrade the quality of service of the input supplied to the competitors. The particular situation of non-price discrimination (sabotage) by a vertically integrated monopolist has been analysed by a number of papers, since the seminal study by Economides [1998]. Mandy [2000], Beard et al. [2001] or Mandy and Sappington [2007], among others, provide detailed analysis of the potential and the impact of sabotage by a vertically integrated supplier. In the same vein, Sappington and Weisman [2005] analyse the incentives to develop "self-sabotage" whereby a vertically integrated monopolist intentionally raises the upstream costs and/or reduces quality, including for its downstream subsidiary. Empirical evidences of sabotage can be found, for example, in Reiffen et al. [2000] or Reiffen and Ward [2002] within the cellular phone market in the US.

The negative effects on downstream competition have to be balanced against efficiency consideration. Indeed, a driving force of vertical integration relies on the technological interdependencies between upstream and downstream markets. Behind academic papers examining that question is the assumption that a vertically integrated structure may entail significant economies of scope due, for example, to the existence of common fixed costs. In the rail transport sector, a first set of studies examines the cost synergies between infrastructure management and trains operations (see, for example, Ivaldi and McCullough [2001], Ivaldi and McCullough [2008], Growitsch and Wetzel [2009] or Mizutani and Uranishi [2013]). Empirical results highlight that vertical disclosure

might be associated with higher costs due to the existence of scope economies between rail infrastructure management and train service operations.

Yet some effects might improve the performance in a vertical separated setting. As pointed out by Nash [1997], both firms will become more specialized in their respective fields, which could result in better incentives toward performance. Besides, large integrated firms may experience significant problems in implementing internal incentive schemes to reduce production costs (due to the fact that, in such settings, aligning the incentives of the upstream and downstream firms could be subject of great difficulties).

Despite those considerations on potential gains in terms of performance, the economic literature often offers the following trade-off: an integrated mode should be preferred when the potentials for scope economies outweigh the economic losses in terms of competition. This central trade-off has been formally addressed by Crew et al. [2005] in the general case of network industries.

A third approach to study the phenomenon of vertical structure can be found in transaction cost economics. Following the work by Williamson [1975, 1985] we expect that the vertical structure should depend on the characteristics of a transaction. Amongst the characteristics to take into account are the specificity of assets used, that is to what extent the assets are made to support a particular transaction, and the complexity of transactions. Both those characteristics may generate transaction costs in case of vertical separation. The heuristic model of Riordan and Williamson [1985] highlights the trade-off between production and transaction costs. A separated structure entails smaller production costs, due for instance to specialisation, but generates transaction costs which are increasing as specificity of the assets grows or the transaction becomes more complex. Thus the merits of a market decline as those characteristics become more predominant.

In the case of railways, the framework developed by transaction cost economics has been used to study the relation between infrastructure management and train operations. Indeed the physical specificity of the asset used by an infrastructure manager seems straightforward, as it cannot be redeployed. For instance, Preston [2002] suggest that the separation between infrastructure and operation would lead to the transfer of site-specific asset and is a motivation towards a unified governance structure. Yet this specificity is a long term issue but not the only one that should drive the choice for governance structure.

In the short run, Pittman [2005] also reaches the conclusion that the specificity of railways compared to other network industries makes vertical separation harder to implement because of concerns of the day-to-day business of operating the rail sector. This is consistent with the findings of Merkert and Nash [2013] who stress that day-to-day operation along with timetabling is seen as a complex and intense part of the interface between infrastructure management and operating trains after having conducted a survey in three European countries. Previous studies (Mizutani and Uranishi [2013] or Van de Velde et al. [2012]) have also underlined the role of complexity of the transaction by using the density of traffic on a network as a proxy for the complexity of vertical coordination. They find that the merits of vertical integration are increasing with traffic density. Similarly, Finger [2014] argues that Germany preferred a holding structure to deal with the complexity of the network which shows some alignment between the vertical governance structure and the characteristics of the transaction.

In the second chapter of this dissertation, the focus is set on rail capacity rather than the infrastructure. This highlights that the train paths have a strong temporal specificity. Given the need to coordinate maintenance and commercial traffic, any changes made in the schedule may entail a loss either for the infrastructure manager or the train operators. These losses may stem from the uncertainty and should affect the production cost (maintenance and

renewal costs) of the infrastructure manager.

1.2.2 Plurality of vertical Governance structure in Europe

In the case of Europe, a particularity is that the debate cannot be restricted to vertical separation or vertical integration as a binary choice, but has to take into account the various forms of vertical structure in Europe. This plurality creates a continuum of structures between those two extremes. All countries except for Ireland now have some form of vertical separation set in place between the infrastructure manager and train operators. Secondly because various forms of vertical structure have been put in place by the countries in Europe. The European commission has identified 6[8] structures that could apply to infrastructure management in Europe which are the following:

- (1) Fully legally, organizationally and institutionally independent infrastructure manager undertaking allocation
- (2) Integrated infrastructure manager with guarantees of independence in relation to the railway undertaking
- (3) Integrated infrastructure manager working alongside an independent body in charge of capacity allocation
- (4) Independent infrastructure manager allocating capacity having delegated certain infrastructure management functions
- (5) Legally (but not institutionally) independent infrastructure manager undertaking capacity allocation owned by a holding company which also owns one of the operations
- (6) Infrastructure manager in charge of allocating capacity and railway

[8] Source: Commission staff working document accompanying the Fourth report on monitoring development in the rail market, Brussels, SWD(2014) 186.

undertaking still integrated.

What stems from the various structures across Europe? The categories chosen by the European Commission highlight that rather than two actors, the railway market can actually be seen as a three stage production process: maintenance works, capacity allocation and train operations; the first two stages being commonly referred to as the upstream infrastructure management, and the latter, train operations, being the downstream competitive market which is composed of an incumbent operator and, depending on the country and the sub-market, new entrants. One of the challenges in railways is that both maintenance works and train operations need access to the tracks, in other words that they consume capacity. One of the objectives of the process of allocating capacity is therefore to coordinate maintenance and commercial use of the network[9]. A key difference between the structures is that their ability and means to achieve coordination should vary. Based on those coordination mechanisms, it is possible to regroup the various structures in four categories as pointed out in figure 1.1.

We consider four noticeable vertical structures in Europe which are: full unbundling, separation of capacity allocation, a holding structure and full bundling We consider the following categories:full unbundling (1); separation of capacity allocation (3) and (4); a holding structure (2) and (5) and full bundling (6). as in Finger [2014]. Figure 1.1 depicts how interactions take place between the three actors and if coordination takes place in the market, in a holding structure or in a firm. Referring to Coase [1937] and transaction cost economics (Williamson [1975, 1985]), the difference is that if interactions take place within a firm, there can be a conscious coordination through authority, whereas in the case of a market relationship, coordination is achieved using prices. In between is the coordination within a holding, with interactions between two entities that are legally but not institutionally independent.

[9] The process is described in more details in the second chapter of this dissertation

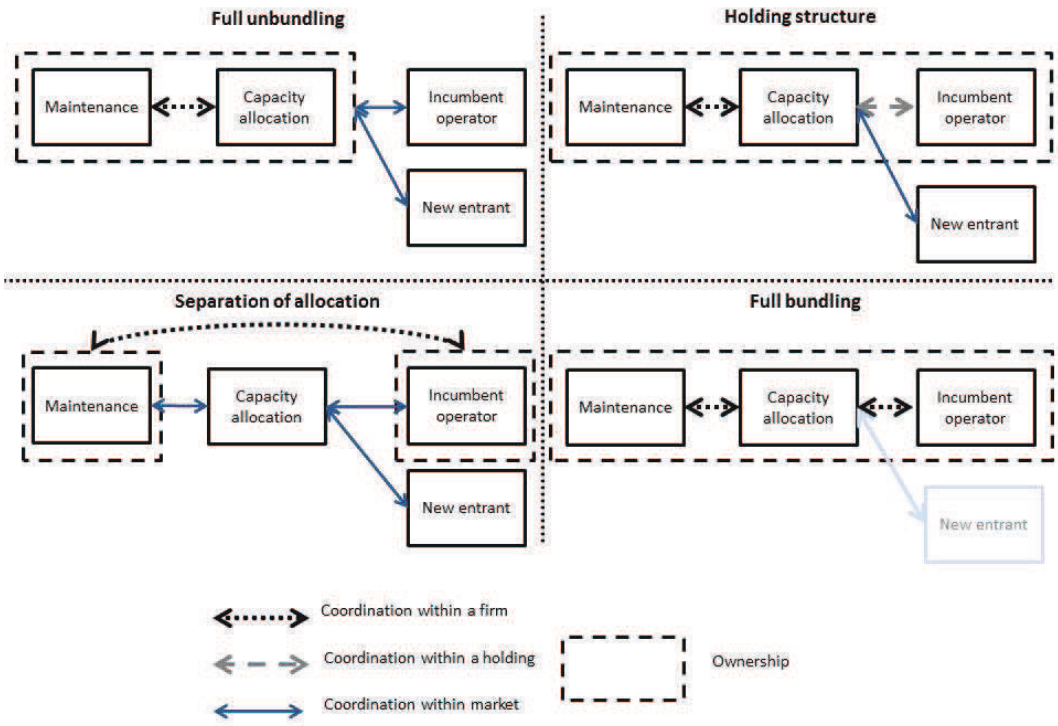

Figure 1.1: The four vertical governance structures in Europe

More than the four categories, we suggest a ranking for the structure, based on how coordination can be achieved. Setting full bundling and full unbundling at each extremity of the spectrum seems straightforward given the literature that has analysed the railway sector, but this is not the case for the holding structure and the separation of allocation structure. In between, the criteria used was if coordination had to be achieved using a market mechanism, or if it is done within a firm. We consider that the holding structure has more means to achieve coordination than when capacity allocation has been separated from the rest. Indeed it can more easily have access to alternate coordination mechanisms to match maintenance and commercial use, because *in fine* any arbitrage can be done within the firm, at the level of the holding structure. This is in line with Growitsch and Wetzel [2009] which considered in a study that the holding structure is an integrated one. The ranking allows a more subtle measure of the effects of vertical separation, besides the effects of having a fully unbundled structure. The structure of each country is presented in Table 1.2. Note that the data used is from 2012 to match our sample. Since then France switched to a holding structure since 2015. The motivations behind this choice confirms the criterion for this ranking, since one of those motivations was the difficulties experienced to achieve coordination in this governance structure between *SNCF Infra* in charge of maintenance works, *RFF* (Réseau Ferré de France) in charge of allocating capacity and the *DCF* (Direction des Circulations Ferroviaires) who would run day-to-day operations. This is in line with our interpretation that coordination is made easier in the holding structure than in the case of separation of capacity allocation.

In terms of impact on the maintenance and renewal costs of infrastructure management, we can expect that the costs of the infrastructure manager will vary depending on the loss of synergies and the increase in coordination costs in the case of vertical separation, which might be compensated by a specialization effect and better incentives toward performance. The expected effect

Table 1.1: Classification of functional forms depending on the availability of coordination mechanisms

Full unbundling	Separation of capacity allocation	Holding structure	Full bundling
Sweden (1988)	France (1997)	Germany (2000)	Ireland
Great Britain (1994)	Hungary (2006)	Italy (2001)	
Finland (1995)	Lithuania (2006)	Poland (2001)	
Norway (1997)	Slovenia (2007)	Belgium (2005)	
Portugal (1997)	Estonia (2009)	Austria(2005)	
Denmark (1998)		Latvia(2007)	
Netherlands (2002)		Switzerland (2009)	
Slovakia (2002)			
Bulgaria (2002)			
Czech Rep. (2003)			
Spain (2005)			
Greece (2008)			

Table 1.2: Vertical governance structures across Europe (Year of the reform in parenthesis)

are summed up in Table 1.3 and lead to an ambiguity on the overall effect.

	Full unbundling Separation	Sep. allocation ; Holding $\longleftrightarrow$	Full bundling Integration
Synergies	$\nearrow$ costs		$\searrow$ costs
Coordination costs	$\nearrow$ costs		$\searrow$ costs
Specialization effect	$\searrow$ costs		$\nearrow$ costs
Setting clear incentives	$\searrow$ costs		$\nearrow$ costs

Table 1.3: Expected effects of the vertical structure on costs

Besides an ambiguous overall effect, the existence of opposite effects for each structure might also have created a bias when the structure was chosen. We conjecture that the necessity of preserving synergies should have been all the more pressing to the policy makers that the firm was perceived as efficient and introducing a big change would have had a negative impact. On the contrary, having a specialized firm, or more particularly the possibility to set clear incentives are arguments that policy makers would be receptive to if they face a firm too big to regulate. This is particularly the case if the objective of the public firm had been previously affected by the objectives of the politicians that controlled them. For instance Boycko et al. [1996] note that public firms are prone to excess employment. In a similar fashion, Gagnepain and Ivaldi

[2007] find that the objective function of the regulator in the case of public urban transport might be to maximize the costs when regulation is captured by a third party, in this case the union of the transport operator. Friebel et al. [2010] note that productive efficiency gains in railways are correlated with staff cuts. So having a fully unbundled infrastructure manager can be a radical way for the state to take back the control of the railway sector. In other words this solution is preferred when the state monopoly if perceived as inefficient.

Therefore, a key issue which we have to address in our empirical strategy is the one of endogeneity. Given the known properties of the vertical structures, estimation results might suffer from a bias created by reverse causality. In particular if we conjecture that when a firm was perceived as less efficient, a more radical change was introduced, that is full separation, then the results of a regression not taking into account the endogeneity issue would affect negatively the results of vertical separation.

As for the overall effect on the costs of maintenance and renewal of infrastructure management, if we rely on the classic trade-off between the loss of synergies versus efficiency gains in the downstream market due to the introduction of competition, we should only capture the loss of synergies because we only look at the costs of the upstream firm. Therefore, despite having raised the ambiguous effects of the choice of vertical structures we make the following prediction:

Prediction 1:
A more unbundled governance structure will increase the costs of maintenance and renewal of infrastructure management, all other things being equal.

1.2.3 Regulation in railways

As noted previously, the railway reform was also accompanied by the creation of an independent regulatory agency (IRA). The creation of an IRA is in line with reforms that have taken place in other network industries. One of its key mission is to enable the emergence of fair competition by guaranteeing non discrimination against new entrants. The role of the agency is all the more relevant since the incumbent has been newly split into an infrastructure manager and a (usually) dominant operator on the downstream markets. Therefore they oversee market design for rail capacity. We conjecture that this duty does not have a clear impact on the performance of infrastructure manager. On the one hand, transparent market design should improve the performance of the overall sector. On the other hand, the regulator will pay close attention to any opportunistic behaviour of the infrastructure manager, rather than train operators which are not directly under its oversight. This might create some rigidities which will have a negative impact on upstream efficiency. This effect should be greater when the vertical structure is more integrated. Indeed if the benefits of vertical integration are the possibility to use alternate coordination mechanism than the market, a regulator should coerce the incumbent firm to rely on the market mechanism in order to avoid discrimination.

A second objective of the regulator can be to monitor efficiency and give incentive to the natural monopoly that is infrastructure management. Extensive literature has focused on the role regulation has to play in enforcing efficiency for a monopoly and overcoming asymmetric information as explained in Laffont and Tirole [1993]. Yet in the case of railways, the independent regulatory body were not given *per se* this prerogative which falls to the Member States. Regulatory bodies do have a role to play in enforcing accounting separation in the sector. Therefore they can reduce potential cost padding Rogerson [1992] in the case of a multi-product monopoly and have an effect on the cost of the service for the new entrants. The more integrated the structure and the more

this prerogative becomes significant.

Finally its independence is a means to limit the opportunistic behaviours from political power and the temptation to hold prices down below full economic costs (Levy and Spiller [1994]). In the previous section, we stressed that one of the difficulties to define the pricing scheme was meeting the budget constraint. And indeed it is rarely the case that access charges should cover the infrastructure cost. In the directive 2012/34/EU, it is explicit that the common rule should be that access charges are set equal to the marginal cost, although as an exception the regulated firm may levy a mark-up to recover its full cost of managing the infrastructure. Otherwise financial equilibrium is set through subsidies. This financing constraint means that all Member States (and Switerzeland) have a performance contract or regulation contract that binds them to the infrastructure manager. The contract will usually specify a certain amount of subsidies, and set in response an expected level of quality for the infrastructure, as well as the perimeter of the network which is expected to be maintained. In such context, one the duty of the independent regulator may be to monitor the terms of the contract, but not to define them.

Therefore the model of the IRA does not necessarily fit with railways in Europe. As it has been highlighted by Benedetto et al. [2015], the primary concern of regulators is really to guarantee the absence of discrimination. But in regards to performance and efficiency there is a dual regulation between the government and the IRA. In the sample they study (15 countries all around the world), Cherchye et al. [2015] find that only in the UK was the rail regulator awarded the duty to enforce efficiency. In most European countries it is done so through a performance contract signed between the central state and the infrastructure manager, although those contracts do not necessarily set challenging targets for the infrastructure manager. Nevertheless ,this is why Crozet et al. [2012] define the government as the most important stakeholder in railways. Indeed it has the capacity to define the consistency of the network

and the resources of the infrastructure managers in a market that relies on public funding. This has led to different types of regulation in the beginning such as described by Crozet et al. [2012]. The creation of independent regulatory bodies was then made mandatory in European law by the directive 2001/12/EC, that is after the obligation to unbundle the network. The independence of regulatory bodies was then gradually enforced by the European Commission. Table 1.4 displays when an independent regulator was created in the European countries in our sample.

Sweden 1988	Estonia 1993	Great Britain 1993	Czech Rep. 1994	Denmark 1996	Germany 1998	Austria 1999
Switzerland 2000	Bulgaria 2001	Lithuania 2001	Finland 2003	Norway 2003	Poland 2003	Belgium 2004
Netherlands 2005	Slovakia 2005	Ireland 2006	Hungary 2007	Slovenia 2007	Portugal 2008	France 2009
Greece 2011	Latvia 2011	Spain 2013	Italy 2013			

Table 1.4: Year of creation of the independent regulatory body

To explain this heterogeneity in the date of creation, we can consider the three reasons given by Gilardi [2005] to explain the diffusion of IRAs :

- Pressure due to political uncertainty[10] in order to avoid inconsistencies in regulation,
- Diffusion in neighbouring countries, due for instance to learning effects, cooperation or imitation according to the author,
- Europeanization, that is as a response to pressures from international institutions, such as the European Union.

As pointed out in table 1.4, we can identify three categories of countries given that an independent regulatory body was made mandatory by the European law following the 2001 directive. A first group of countries are the one that had

[10] As pointed out by Gilardi [2005], political pressure is specific to each country and is significantly associated to the establishment of IRA's. This approach is defined as *bottom-up* by the author.

anticipated the restructuring of railways such as Sweden, Great Britain and Germany and therefore had anticipated the creation of the regulatory body as well. A group of countries complied with the directive soon after it came into force. Finally countries such as France, Spain and Italy did not get the token of regulatory independence until long after the directive was enacted. This is consistent with the work of Thatcher [2002] who points out that France and Italy were late adopters of the IRA model, all sectors considered.

Did the age of the independent regulator have an impact on the costs of the infrastructure manager? We look if the early creation of a regulator decreased the costs of the infrastructure manager. We conjecture that the effect of the regulatory body should be linked to his experience. As stated by Stern [1997]: "*Effective regulation is bound up intimately with the reputation of the regulatory agency and this reputation takes time to build up. Regulators need time to build-up expertise before they are given full decision making responsibility*". This is in line with the framework proposed by Glachant et al. [2013] suggesting that the ability of a regulator to monitor costs is increasing in time. Indeed time and iterations are needed to overcome the asymmetries of information. Age is also commonly used in the measure of regulatory governance. For instance, Martin and Jayakar [2013] note in their survey that 60% of the studies use a minimum age of the regulator as an components when creating an index of regulatory governance. Therefore, to control for the effect of the regulator on the costs of infrastructure management, we suggest to use both the age of the regulator and it age relative to the one of the reform. In the first case, we conjecture that there is a direct link between the age of the regulator ant its expertise. As this conjecture is debatable, we introduce the second variable. In the reforms that took place in electricity, Newbery [2002] stresses the positive influence of having a well defined regulatory framework when the unbundling process takes place. In that sense, having an independent regulator before or at least soon after the unbundling process should translate the importance given by the policy maker to regulation and be a proxy fore *de facto* regulation;

as opposed to a regulator that was not endowed with sufficient resources, or regulatory power.

Prediction 2:
A more experienced regulator will decrease the costs of maintenance and renewal of infrastructure management all other things being equal.

1.2.4 Market Opening

The third of the three step, market opening, is also at various stages across European countries. The heterogeneity comes from national passenger services, which is at very various stages as pointed out in figure 1.2[11], while the transport for freight and international passenger services has been opened to competition.

What are the implication for infrastructure management? Growitsch and Wetzel [2009], when looking at the existence of economies of scope in twenty-seven European countries, find that the integrated companies that benefit the most from economies of scope are those where the downstream market is the most opened to competition. This would mean either that the pressure for the infrastructure manager of facing customers could increase its efficiency despite its monopoly on rail track access or that the efficiency gains due to the introduction of competition downstream spreads throughout the structure and increases the overall efficiency in the case of a more integrated structure. Other benefits could be that the infrastructure manager does not have to face a monopsony and having a firm that can ration the downstream supply. This is particularly relevant in the case of a separated structure as pointed out in Cremer et al. [2006].

[11] Source: European Commission in : '*Impact assessment for the Fourth Railway package'*, 30.1.2013 SWD(2013) 10 final

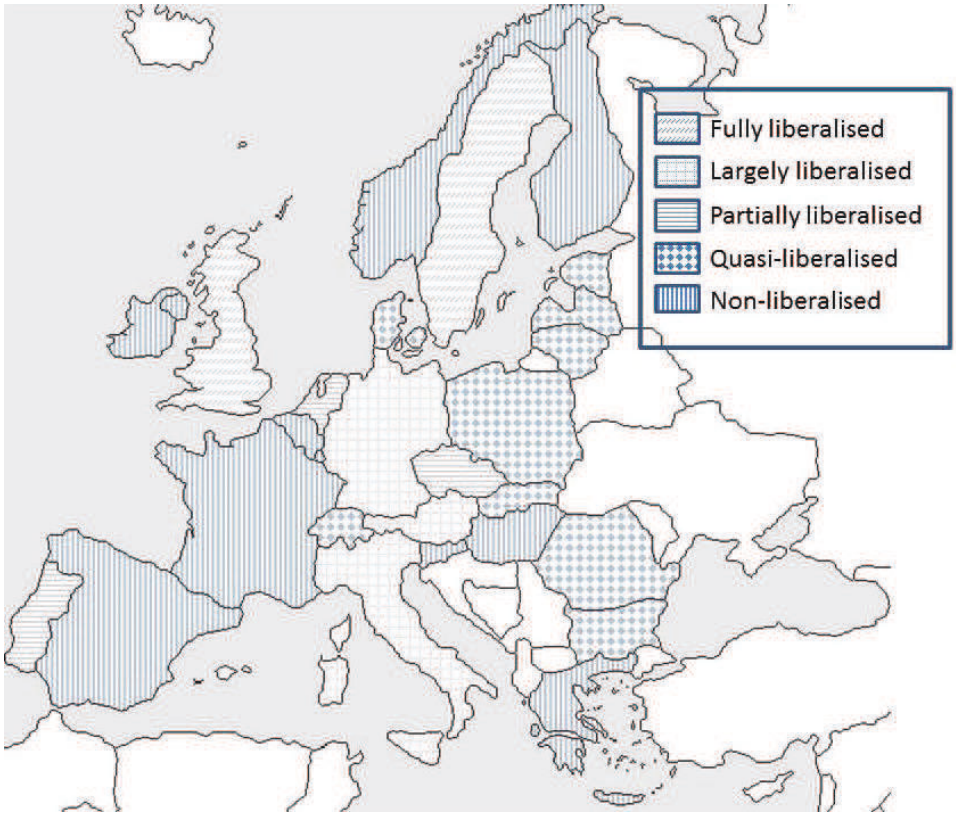

Figure 1.2: Market Opening for National Passenger Services in Europe

On the other hand, as the number of trading partners increases, so should the complexity of coordination as described in the Van de Velde et al. [2012] study. This is consistent with the findings of Bitzan [2003] saying that introducing multi-firm competition on a network increases resource costs. Moreover he infrastructure manager is subject to strong requirements in terms of transparency and non discriminatory access according the European law, enforced by an independent regulator. The guarantees for non discriminatory access on the network could have a cost for the infrastructure manager, this cost being both a condition for and a consequence of entry on the downstream market. For instance, Pittman [2005] argues that one of the pitfall of vertical separation is the difficulties to convey short term and long term incentives using access charges. Given that our study is focused on the costs of the network, we conjecture that

Prediction 3:

Increased liberalisation in the downstream market will increase the costs of maintenance and renewal of infrastructure management all other things being equal.

1.3 Empirical analysis

1.3.1 Description of the data

In order to test our predictions, we constructed a database covering four years between 2009 and 2012 and 25 countries with a total of 81 observations and is an unbalanced panel as the data on costs was not available every years for all countries.

The data that we used for costs comes from the questionnaires sent by the European Commission to Member States in order to monitor the railway sector[12]. This institutional database is, to our knowledge, the only public database that offers information on costs in railways and disentangles the cost structure of the infrastructure manager by making the distinction between renewal expenditures and other investments. Indeed, as in most network industries the cost perimeter for any applied work can be troublesome. Operational expenditures (OPEX) cannot be used solely, first because the scope of activities varies from one country to another. For instance some infrastructure managers will be operating train stations, others will only be managing tracks. Secondly using operational expenditure leads to neglect renewal expenditures which is categorize as a capital expenditure (CAPEX). Yet using total expenditures (TOTEX) leads to incorporating also new investments which would reflect poorly on the countries which are currently developing new infrastructures.

[12] And is referred to as RMMS

Due to those considerations, we therefore build our explained cost variable by summing maintenance and renewal (M&R) expenditures of the network which are respectively a sub part of the OPEX and the CAPEX. Note that summing-up these two expenditures allows us to circumvent the issue of accounting differences between countries where some construction work can be categorized alternatively as renewal or maintenance expenditures. It leaves outside our analysis' scope the expenditure for operating the network, such as traffic management, which is not available in public datasets. A caveat attached to the use of (M&R) expenditures is the existence of cycles to renew the network which is not necessarily linear in time. This is why for instance Wheat and Smith [2008] adjust the renewal cost of Great Britain to its steady state as to avoid a bias when measuring efficiency. But such an adjustment rests upon the hypothesis that all other countries are in their steady-state too. Ultimately, in the absence of detailed data on the age of networks, this bias cannot be avoided. It might be mitigated by the use of panel data where various renewal cycles may be captured across Europe.

Table 1.5: Descriptive statistics

Variables	Description	Mean	Std. dev.	Min	Max
M&R cost	Maintenance and renewal costs (M €)	906	1170	18	4810
Train.km	Train km per route km	19	14	6	68
Route	Length of the network	8122	8857	1196	41876
Nb tracks	Average number of tracks per route km	1.63	0.35	1.06	2.40
Electrified	Proportion of electrified tracks on the network	0.48	0.25	0.03	1

For the output and control variables the data used was published by Eurostat and by the UIC[13]. The regression we use is similar to the one used by Wheat and Smith [2008]. Indeed, although previous studies have estimated the cost structure of the rail sector, there hasn't been studies focusing on estimating

[13] International Union of Railways

a cost function of the infrastructure manager. As the results of Wheat and Smith [2008] were used by the British regulatory body to set efficiency targets, the choice of variable has been subject to a great deal of discussion.

To control for heterogeneity in prices, the costs were adjusted for purchasing power parity according to the data published by the OECD. But given that the break down of costs was not available, we cannot estimate the input prices which prevent us from estimating a cost function. Yet, we want to make the case that estimating a cost function would not have been appropriate in our case. First because in a cost function we consider that the input markets are perfectly competitive and the firm has no influence on prices (Coelli et al. [2005]). This hypothesis does not hold in our case on two accounts: given its monopsony nature the firm should have an influence on input prices, also all infrastructure manager are public firms and may be subject to the constraints of public purchasing, keeping them away from competitive prices. Secondly a cost function is theoretically not compatible with the use of capital or operational expenditures. Indeed the variable of cost has to be constructed as to have the price for each input included in the total. Thirdly an implicit assumption is that the firm can change the proportion of inputs, which for political reasons might not be the case in public firms (see Boycko et al. [1996] for instance). Similarly, Friebel et al. [2010] highlight the fact that railways may suffer from over-employment and raises the question to what extent infrastructure manager are able to optimize their inputs due for instance to political pressure.

The baseline model is a log linearised Cobb-Douglas function[14] with the following variables:

$$\ln Cost_{it} = \beta_0 + \beta_1 \ln route_{it} + \beta_2 \ln train.km_{it} + \beta_3 \ln nbtracks_{it} + \beta_4 \ln electrified_{it} + \epsilon_{it} \quad (1.1)$$

[14] We tested the introduction *Translog* function, which is more flexible but the results were not conclusive. We attribute it to the data restriction we have.

Table 1.5 presents the variables[15] used to model the rail network as well as the descriptive statistics associated.

In order to select the output variable, we consider that the production of an infrastructure manager consists both in the number of train kilometres run on the network by train operators, but also maintaining the network to a certain consistency level. The maintenance and renewal cost will therefore depend on the length of the network and its characteristics such as the the number of tracks per route kilometre and the electrification. Note that although one could expect that a greater proportion of electrified tracks increase the maintenance cost, previous studies (such as Wheat and Smith [2008]) find a negative coefficient for this variable. This may be due to the fact that the electrification variable captures also the age of the network, an electrified network being younger, and a more recent network is cheaper to maintain. It may also capture the intensity of use of the network.

We complete our dataset with variables on organization and regulation in order to test our predictions:

Prediction 1:

- *Full separation* is a dummy variable equal to one if the vertical structure set in place is full separation between the infrastructure manager and the train operator(s).
- *Vertical structure* is an ordered variable which ranges from 0 in case of full separation to 3 in case of full integration. In-between, in case of separation with delegation the variables takes the value 1 and 2 in case

[15] A difference with the model developed by Wheat and Smith [2008] is that we take the average number of tracks instead of the proportion of single tracks. We believe this captures better the complexity of some network such as the Netherlands or Switzerland whose average number of tracks is above two.

of a holding structure.

Prediction 2:

- *Age regulation* captures the age of the independent regulator in each countries.
- *Difference* captures the lag between the date of unbundling of railways and the introduction of an independent regulator. Note that if the regulator was introduced before separation occurred, the variable will be negative

Prediction 3:

- *Competition* is an ordered variable which takes into account the degree of market opening for national passenger service based on a working document published by the European Commission[16]. It ranges from 1 in case the market is non liberalized to 4 where it has been fully liberalized.

Table 1.6 presents the descriptive statistics for the previous variables.

Table 1.6: Descriptive statistics of governance variables

Variables	Mean	Std. dev.	Min	Max
Full separation	0.48	0.50	0	1
Vertical Structure	0.88	0.96	0	3
Difference	1.03	6.03	-9	12
Age regulation	7.75	6.01	0	24
Competition	2	0.97	1	4

1.3.2 Addressing the issue of endogeneity

An issue that has been less highlighted in railways regarding the reforms is the one of endogeneity. Endogeneity may arise for several reasons among which is the case of omitted variables, measurement errors and reverse causality. In

[16] Page 16 of SWD(2013) 10 final: *Impact assessment*, Brussels, 30.1.2013.

the case we address, we consider the issue of reverse causality of importance. Reverse causality occurs when the causal effect does not only run from the predictor variable to the outcome variable, but may also go the other way around. The literature on the implementation of independent regulator gives on this regard good illustrations. As pointed out by Galperin et al. [2013], a regulatory agency which is introduced to increase the performance of the sector will be done so even more as the sector is performing badly.

In the case of the railway reform in Europe, the motivations of the European Commission for pushing forward with vertical separation are consistent with this issue. Indeed, vertical separation is seen as one of the elements to revitalize railways across Europe. Nash [1997] gives further arguments that can be used to stress the issue of endogeneity. With vertical separation, both firms will become more specialized in their respective fields, which could result in better incentives toward performance. Besides, large integrated firms may experience significant problems in implementing internal incentive schemes to reduce production costs. This is due to the fact that, in such settings, aligning the incentives of the upstream and downstream firms could be tough to implement. Therefore, the heterogeneity we observe between the European countries in the way the reform was carried out might have been driven by the performance of the sector at the time and by the determination of policy makers to introduce more radical changes in the governance structure.

From a technical point of view, endogeneity arises if the regressor is correlated with the error term. In order to tackle the issue of endogeneity and to obtain consistent estimator a solution is to use an instrumental variables (IV) regression. The instruments must capture the information in the regressors that are uncorrelated with the error term in a first stage regression, and allows to eliminate this bias in a second stage regression, thus providing consistent

coefficients[17].

There are two conditions for an instrument to be valid: it must be relevant and exogenous. For an instrument to be relevant we need it to be correlated with the regressors and in order to be exogenous, it need to be not correlated with the error term. As instrumental variables, we use the choices that were made in the electricity market. One of the criticisms made towards the reforms of network industries was that the same model of vertical separation was imposed without taking necessarily into account the specificities of each industry (see Beard et al. [2015] for instance). It has been argued that the wave of reforms across network industries might have been to a certain extent ideological. Therefore we conjecture that the choices made in the electricity market can explain in part those made in the railway sector by capturing the dynamic of reforms in a specific country. With a positive correlation between the choices made in terms of regulation and organization in the electricity and railway markets, the instruments satisfy *a priori* the relevance condition. Yet, there should not be any correlation between the error term of the regressions ran and the choices made in the electricity sector, therefore the instrument variables are *a priori* exogenous. Given that in our framework endogeneity stems from efficiency, it means that there is no correlation between efficiency in railways and the structure of the electricity market.

More precisely, we take three instruments from the electricity market. The first one, *Index Electricity* is an index built by the OECD to measure entry regulation: the higher the index, the less liberalised the market is. Details on the construction can be found in Koske et al. [2015]. This variable is used as an instrument for the choice of the vertical structure in railways. We use an indicator on the condition of market entry because the vertical structure in the

[17] A brief aside: given that our primary concern is to assess the impact of organisational variables on efficiency, a possible methodology would have been to estimate a stochastic frontier. Yet there are some difficulties associated with the use of a stochastic frontier because endogenous variables can be correlated with statistical noise, or with technical inefficiency, or both as pointed out in Amsler et al. [2016].

electricity market do not offer enough heterogeneity to differentiate between European countries, while our variable on entry should capture the will of countries to unbundle the electricity sector. Given that the variable is growing with entry constraints, we expect for the sign of our first stage regression to be negative with the dummy variable capturing full unbundling and positive with our ordered variable.

The next instrument variable is the age of the regulator in the electricity market. It is used as an instrument for the mirroring variables in the railway market. Once again, we anticipate a positive correlation and capture the countries that have adopted more rapidly an independent regulator. Note that in order to check the robustness of our results, we add further instruments later on. Lastly, we use an index of the overall regulation of network industries (developed by the OECD as well) as an instrument for the degree of liberalisation in railways as we did not find suitable instruments from the electricity market.

1.3.3 Results

In the following subsection, we report the regressions of cost functions and of our variable of interest on organization and regulation. The results are presented the following way. First we present the results from estimating a random effect regression with the organizational and regulation variable, testing alternatively the *full separation* (model 1) dummy and the *vertical structure* ordered variable (model 2), as well as the inclusion of our variable *difference* (model a) or *age regulation* (model b). We then display the results from the IV regressions respectively on the vertical structure, the age of the unbundling and the age of the regulator as well as the degree of market opening in national passenger services. All regressions include heteroskedasticity robust standard errors and year dummies which are not reported on the tables.

In order to check the robustness of our estimates coefficient, we include the Kleibergen-Paap rk LM statistic (and its associated P-Value) and the Kleibergen-Paap rk Wald F statistic to test respectively for under identification and weak identification. Firstly, testing for under identification lets us check that the instruments are indeed relevant, correlated with the endogenous regressors. Secondly, instruments are considered weak when they explain to little variation of the endogenous variable. In this case, the normal distribution provides a poor approximation. We also provide the results of an endogeneity test. The null hypothesis is that regressors suspected to be endogenous are in fact exogenous, meaning that we should prefer the estimates from the one-stage regression. Finally we provide some evidence that the instruments satisfy the exogeneity condition by running the test of overidentifying restrictions. To do so, we add an extra instrument to our regression and report the Hansen J-statistic where the null hypothesis is that the all the instruments are exogenous. More precisely, if the P-value associated is below 10%, at least one of the instruments (or both) are endogenous. The estimates using two instruments are presented in the appendix.

The random effects model

The first regression presented does not include instruments and therefore does not take into account endogeneity. The regression ran is a random effects model which is preferred to a fixed effect model, mostly because the variables we want to test do not necessarily vary over time It is the case for instance of the vertical structure, also it is confirmed by the Hausman test results in table 1.11 of the appendix, suggesting a random effect model is suitable.. The results of the estimation are presented in Table 7.

The coefficient of the variables in the cost function have the expected signs as well as the expected magnitude. Indeed since the model estimated is a

Table 1.7: Results of the random effects regression

M&R Costs	Model 1a	Model 1b	Model 2a	Model 2b
ln train.km	0.956***	0.927***	0.926***	0.911***
	(0.192)	(0.179)	(0.186)	(0.181)
ln route	0.977***	1.037***	0.959***	1.030***
	(0.097)	(0.090)	(0.092)	(0.089)
ln nb tracks	0.354	0.311	0.425	0.322
	(0.417)	(0.431)	(0.379)	(0.401)
ln electrified	-0.344***	-0.305***	-0.329***	-0.292***
	(0.099)	(0.088)	(0.102)	(0.095)
Difference	0.027*		0.030*	
	(0.015)		(0.016)	
Age regulation		-0.025**		-0.024**
		(0.012)		(0.011)
Competition	0.147*	0.180*	0.154*	0.180**
	(0.087)	(0.092)	(0.083)	(0.091)
Full separation	-0.177	-0.044		
	(0.179)	(0.195)		
Vertical Structure			0.129*	0.047
			(0.076)	(0.079)
Constant	-5.624***	-5.947***	-5.625***	-5.906***
	(1.029)	(1.063)	(0.946)	(0.990)
Observations	81	81	81	81
R^2	0.901	0.898	0.906	0.899

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15

log-linearised Cobb-Douglas function, the coefficient can be interpreted as the elasticity. Therefore it is intuitive that the coefficient of the variable *Route* be positive and roughly around one. Similarly, the coefficient of *train-km* is positive and below one, as could be expected because of the existence of fixed costs in railways. The coefficient for the number of tracks per route-km is positive. The coefficient for electrification is negative, which is a counter intuitive, but common result when estimating a cost function in railways.

Regarding the coefficient of the organizational and regulatory variables, we provide here a brief explanation of the results, although the coefficients are to be interpreted in light of the IV regression which we will present later on, as the causal relationship may not be straightforward as suggested earlier. We

find that the dummy variable for full separation is not significant while the coefficient for the ordered variable is positive and significant at a 10% level in one of the model suggesting a more integrated vertical structure does not entail lesser costs in terms of maintenance. Therefore we would have to reject our first prediction. The coefficient for the variable difference is positive, which suggest that the greater the lag between the reform and introducing a regulator, the higher the cost for maintenance in accordance with our second prediction, while an older regulator reduces the cost for maintenance. As we conjectured, we find that the degree of downstream competition increases the cost of maintenance and renewal.

Results on the vertical governance

Results from the 2SLS regression of the vertical structure on the costs of maintenance are shown in Table 1.8. We find that vertical separation has a negative impact on maintenance and renewal costs, while the more integrated the firm is, the greater those costs. Our instrument, measuring the barrier to entry in electricity appears to be suitable. Firstly, the test results conclude that the instrument is identified and not weakly. Also in the first stage of the regression, the instrument is of the expected sign : the greater the barrier to entry in electricity, the more integrated the structure in railways. Lastly, when introducing a second instrument, an index of in telecoms built by the OECD (see Koske et al. [2015]) in order to have another network industry to compare to, we find that the two instruments can be considered as exogenous (see Table 1.12 in appendix).

Interestingly, the endogeneity test confirms that the vertical structure is endogenous. Taking into account endogeneity amplifies the (positive) effect of a more integrated structure on the costs of maintenance and renewal, and enables us to find a negative effect on costs for a vertically separated structure. Our first prediction, based on the trade-off made explicit in the economics

Table 1.8: Results of 2SLS regressions on the vertical structure

	Full separation 2SLS		Vertical structure 2SLS	
	Full separation	M&R costs	Vertical structure	M&R costs
Full separation		-0.542***		
		(0.209)		
Vertical structure				0.259**
				(0.106)
ln train.km	-0.008	0.816***	0.359*	0.727***
	(0.103)	(0.118)	(0.213)	(0.116)
ln route	-0.109+	0.967***	0.200*	0.974***
	(0.067)	(0.063)	(0.111)	(0.058)
ln nb tracks	-0.528**	0.502+	-0.021	0.794***
	(0.232)	(0.315)	(0.469)	(0.295)
ln electrified	0.239***	-0.114	-0.535***	-0.104
	(0.067)	(0.082)	(0.178)	(0.095)
Index electricity	-0.378***		0.789***	
	(0.064)		(0.140)	
Constant	2.005***	-4.453***	-2.534**	-4.882***
	(0.716)	(0.757)	(1.124)	(0.638)
Observations	81	81	81	81
R-squared	0.848		0.865	
KP LM-stat		13.396		11.350
KP P-Val		0.0003		0.0008
KP Wald F-stat		34.847		31.570
Endog. χ^2		0.0534		0.0744

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15

literature regarding vertical separation, is not corroborated. The interpretation we suggest is that full vertical separation was considered a more radical, as presented for instance in Staropoli and Yvrande-Billon [2009], change and chosen in the countries where railways had an efficiency problem which creates a bias when measuring the effect of vertical separation in previous studies on the topic (see for instance Van de Velde et al. [2012]).

An explanation for this counter intuitive result may lie in the fact that, once the incumbent has been unbundled and transparency set in place, the opportunities for synergies may not be the same even in a more integrated form such as the holding structure. For instance, to guarantee third party access,

the regulator or the Member State may hold the infrastructure manager to greater standard in terms of safety, age of the network and availability of capacities which come at a higher cost in terms of infrastructure management. Any arbitrage that could be made to decrease the cost of track maintenance at the price of decreasing service quality (lowering the maximum speed, closing lines that are not used) becomes less likely to be made by the infrastructure manager.

Results on the timing of the reform

The results from the 2SLS regression on the timing of the reforms and the introduction of a regulator are presented in Table 1.9.

Regarding the introduction of an independent regulator, we find that age of the regulator in the energy sector does not fare as a valid instrument for the age of the regulator in the railway sector. But it is a suitable instrument when we take into account the difference between the reform and the introduction of a regulator. The measured impact for this variable is greater than the one we measured in the random effects model, and find that the variable is endogenous. The relevance of the variable lies in the fact that the introduction of a regulator was made mandatory after it was the case for the vertical structure. Countries that anticipated the creation of a regulator may therefore be those that are more keen on granting the regulator with powers to supervise the infrastructure manager. In that sense, it may not be *per se* the age of the regulator that is relevant but the timing of its introduction which proxies the regulatory power which it was awarded, as well as the lag in terms of asymmetric information it has to overcome. The fact that endogeneity increases the measured effect of this variable suggests that in the countries where efficiency was perceived as low, a greater interest was set on having a well defined regulatory framework.

Table 1.9: Results of 2SLS regressions on the timing of the reform

	Age regulator 2SLS		Difference 2SLS	
	regulation	M&R cost	difference	M&R cost
age regulation		0.218		
		(0.236)		
difference				0.084**
				(0.033)
ln train.km	1.181	0.559+	-2.266**	1.006***
	(1.115)	(0.348)	(0.934)	(0.162)
ln route	1.009	0.756**	0.948	0.897***
	(0.813)	(0.341)	(0.802)	(0.068)
ln nb tracks	4.365	-0.028	-4.435*	1.295***
	(3.061)	(0.930)	(2.424)	(0.294)
ln electrified	-0.262	-0.196	3.119***	-0.514***
	(0.752)	(0.168)	(0.728)	(0.130)
age IRA electricity	0.133		0.349***	
	(0.134)		(0.116)	
Constant	-9.683	-3.333	0.829	-5.517***
	(7.385)	(2.635)	(7.337)	(0.703)
Observations	81	81	81	81
R-squared	-0.29		0.881	
KP LM-stat		0.958		8.246
KP P-Val		0.3278		0.0041
KP Wald F-stat		0.988		9.046
Endog. χ^2		0.0136		0.0494

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15

Given that our F-stat is less than 10, this could indicate that our instrument is weak (see Stock and Watson [2003]). In order to correct for it, we introduce an extra instrument presented in Table 1.13 in the appendix. Adding an extra instrument also allows us to run the test of overidentifying restrictions. Given that the P-value is 0.78 (as reported in Table 1.13), the null hypothesis is not rejected and the instruments can be considered as exogenous.

Note that we did not find valid instruments in the electricity sector for the degree of competition downstream that conclude that there is a significant impact on costs. We present in table 1.14 in the appendix the results from a 2SLS regression using as instruments the index build by the OECD that summarises the overall regulatory provision in seven network industries. The

results conclude that the competition variable does not exhibit any endogeneity which would suggest that we should consider the results from the random effects model.

Predictions	**Random effects model**	**2SLS model**
Prediction 1: A more unbundled governance structure will increase the costs of maintenance and renewal of infrastructure management all other things being equal.	Not significant	Yes
Prediction 2: The more experienced the regulator is, the more efficient infrastructure management is.	Yes	Yes, but only for the variable *difference*
Prediction 3: Increased liberalisation in the downstream market will increase the costs of maintenance and renewal of infrastructure management all other things being equal.	Yes	Not significant

Table 1.10: Summary of predictions and results

1.4 Conclusion

Our objective was to present an overview of the railway infrastructure management reforms across Europe. We observe a great diversity in all three steps of the reform, that is the vertical structure, introduction of regulation and of competition. Using this heterogeneity we can compare the merits of the various choices. But the plurality of forms and means by which countries have carried out this shift from a sole incumbent to (at least a minimum degree of) vertical separation raises the motivations behind these choices. Despite the fact that in most countries the reform has been impulsed by the European Union, it still raises an endogeneity issue, stemming from an uncertain causal relationship.

An important caveat attached to our results is that we estimate the impact of institutional and organisational choices on the cost of maintenance and renewal of the network. Therefore it does not allow us to perform a full analysis of the performance of railways across Europe and we cannot conclude in favour

of one organizational structure rather than another. But our results should be interpreted in light of the difference in estimates depending on whether the issue of endogeneity is addressed or not.

Regarding the vertical structure, we derive a ranking of vertical structures based on the study of coordination mechanisms and do not find that the more integrated the firm, the lower the costs of maintaining the network. On the contrary, we find that those costs decrease with vertical separation. This would mean that the effect of specialisation if often overlooked compared to one of potential synergies. Once again, we do not measure the impact on train operators, but the difference in our estimates when addressing endogeneity issues or not suggests that the negative effects of vertical separation in railways are regularly overestimated in economic studies.

On the creation of an independent regulator, we find that, once endogeneity is taken into account, it is not the age of the regulator that is significant but rather the difference between the date of the unbundling and the creation of the said regulator. This variable captures the importance of a well defined regulatory framework. It may also capture the "determination" of a country to implement an independent regulator and therefore capture to some extend the its endowment in regulatory powers. In a sector such as railways made up of publicly owned firms and often dependent on public subsidies, some countries first entrusted regulation to the central administration as described by Crozet et al. [2012]. In this case, regulation may suffer from conflicting interests which ultimately deter the results of vertical separation. This leads to straightforward implications in terms of policy. First it stresses the importance of independence in regulation. Second it suggest that the timing of the reform is important as well, and that an independent regulatory agency should be introduced at the same time as the unbundling of the network is made, if not sooner.

The third step we described was the introduction of competition in the downstream markets. On this subject the results we find are not entirely satisfactory as we fail to find a significant effect when we correct for endogeneity. Moreover, the introduction of competition is an area for future research on many accounts. First, the effects of the vertical structure on barriers to entry have not been formally measured, despite the suspicions that a more integrated structure might deter new entrants. Secondly, when competition is introduced in all downstream markets, it will become possible to assess the effectiveness of the strategy to revitalize rail across Europe.

1.5 Appendix

Table 1.11: Hausman test

H0: difference in coefficients not systematic				
	Model 1a	Model 1b	Model 2a	model 2b
χ^2	6.28	26.99	7.06	6.30
Prob $> \chi^2$	0.51	0.0007	0.42	0.61

Table 1.12: Results of 2SLS regressions on vertical structure with an additional instrument

	Full separation 2SLS		Vertical structure 2SLS	
	Full separation	M&R costs	Vertical structure	M&R costs
Full separation		-0.363*		
		(0.191)		
Vertical structure				0.225**
				(0.103)
ln train.km	-0.053	0.822***	0.413**	0.741***
	(0.088)	(0.109)	(0.197)	(0.110)
ln route	-0.203***	0.979***	0.313***	0.978***
	(0.047)	(0.053)	(0.093)	(0.055)
ln nb tracks	-0.531**	0.571*	-0.018	0.783***
	(0.250)	(0.316)	(0.491)	(0.286)
ln electrified	0.365***	-0.159**	-0.688***	-0.123
	(0.054)	(0.071)	(0.171)	(0.089)
Index electricity	-0.244***		0.628***	
	(0.069)		(0.149)	
Index telecom	-0.389***		0.469***	
	(0.102)		(0.163)	
Constant	3.458***	-4.743***	-4.287***	-4.941***
	(0.513)	(0.635)	(0.941)	(0.598)
Observations	81	81	81	81
R-squared	0.871		0.872	
KP LM-stat		21.701		17.337
KP P-Val		0.000		0.0002
KP Wald F-stat		26.403		15.835
Endog. χ^2		0.1731		0.1205
Hansen P-val		0.1591		0.4038

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15

Table 1.13: Results of 2SLS regressions on the timing of the reform with an extra instrument

	Age regulator 2SLS		Difference 2SLS	
	regulation	M&R cost	difference	M&R cost
age regulation		0.038		
		(0.091)		
difference				0.091***
				(0.025)
ln train.km	1.012	0.786***	-1.667*	1.021***
	(1.131)	(0.155)	(0.993)	(0.149)
ln route	0.923	0.961***	1.255+	0.887***
	(0.839)	(0.116)	(0.758)	(0.055)
ln nb tracks	4.132	0.583	-3.609+	1.348***
	(3.134)	(0.425)	(2.384)	(0.318)
ln electrified	-0.312	-0.240***	3.298***	-0.538***
	(0.762)	(0.066)	(0.701)	(0.122)
age IRA electricity	0.084		0.523***	
	(0.143)		(0.118)	
Index electricity	-0.654		2.325**	
	(1.250)		(1.099)	
Constant	-6.386	-4.982***	-10.888	-5.534***
	(10.219)	(0.940)	(10.013)	(0.752)
Observations	81	81	81	81
R-squared	0.834		0.812	
KP LM-stat		1.017		14.863
KP P-Val		0.6015		9.812
KP Wald F-stat		0.558		10.065
Endog. χ^2		0.6732		0.0001
Hansen P-val		0.0004		0.7849

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15

Table 1.14: Results of 2SLS regressions on the degree of downstream competition

	Competition	M&R cost
Competition		-0.010
		(0.091)
ln train.km	-0.138	0.838***
	(0.172)	(0.114)
ln route	0.123	1.008***
	(0.099)	(0.053)
ln nb tracks	0.765*	0.719***
	(0.385)	(0.273)
ln electrified	0.068	-0.251***
	(0.131)	(0.053)
OECD overall	-1.073***	
	(0.228)	
Constant	3.207**	-5.358***
	(1.606)	(0.531)
Observations	81	81
R-squared	0.885	
KP LM-stat		10.446
KP P-Val		0.0012
KP Wald F-stat		22.138
Endog. χ^2		0.3040

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1, + p<0.15

Chapter 2

Vertical Separation in Rail Transport: How Prices Influence Coordination*

2.1 Introduction

In the last two decades, major structural reforms have been implemented in most network industries. Considerable attention have been devoted by economists to the analysis of gas, telecom or electricity industries and, surprisingly, relativity little notice had been paid to railways so far. Yet, driven by the European institutions, the railway transport sector in Europe has also gone through both institutional and organizational reforms during the last twenty years. A main objective of those reforms is to break up the national monopolies in order to open up rail market services to competition. Directive 91/440/EC was the first milestone to this process by introducing a degree of vertical separation in the sector. It required an accounting separation between

* This chapter is based on a joint work with Miguel Amaral. The authors would like to thank Ricard Gil and Yannick Perez for their valuable comments as well as the participants of the seminar organized by the Laboratoire Ville Mobilité et Transport (LVMT) and the participants in our session at the EARIE (2015) conference, at the Congress of the Association Française de Science Economique (AFSE 2015), at the Third Florence Conference on the Regulation of Infrastructures and at ESNIE 2014.
A shorten version of the article was also published in Network Industries Quarterly (vol 16 - No 2 (2014)).

the management of the essentials facilities (the management of the railway network) and the operation of rail services[18].

The upstream entity, the infrastructure manager, was considered a natural monopoly and put under the supervision of a regulator. The downstream market, which consists of train operators was deemed potentially competitive. Since then the railway sector which consisted of vertically integrated monopolies has progressively opened to competition. Directive 91/440/EC allows for different degrees of vertical separation and, as a consequence, different governance modes coexist today in Europe.

Interestingly, four main modes of organization can be found in Europe:

- *Full unbundling*: full separation between the infrastructure manager (IM) and railway undertakings (RU);
- *Separation of allocation / Unbundling with delegation*: separation between the IM and RU, where the IM delegates infrastructure maintenance and operational management to a RU;
- *Holding structure*: separation where a holding company owns the IM and the RU;
- *Full bundling*: a unique firm operates infrastructure management and rail services.

This heterogeneity raises the question of the comparative merits of the different vertical separation degrees characterising rail transport in Europe. Yet, to

[18] Directive 91/440/EEC of 29 July 1991 states that "[w]*hereas the future development and efficient operation of the railway system may be made easier if a distinction is made between the provision of transport services and the operation of infrastructure; whereas given this situation, it is necessary for these two activities to be separately managed and have separate accounts; 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 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.*

our knowledge, no strong theoretical and empirical evidences suggest that the overall impact of vertical separation on consumer surplus is positive (or negative) in the sector. As a consequence, there is no clear answer to the optimal structure for the rail transport sector in Europe. A key objective of the paper is to shed light on this debate by focusing on a understudied issue, namely the coordination problems between the upstream (infrastructure access management) and the downstream (rail services activities) markets.

The issue of coordination stems from the developments in transaction costs economics (Williamson [1985]). Indeed it can be argued that coordination costs are a core determinant of the vertical structure in network industries and poor efficiency may arise from a misalignment of the governance structure. An important characteristic of the reforms across Europe is that it has created a shift in the means through which coordination may be achieved. Once the infrastructure management and the railway services have been separated, one must rely on prices instead of authority to achieve coordination as pointed out by Coase [1937]. A limit to a separated structure is therefore, as shown in Hart and Tirole [1990], that it may not be able to share the profits efficiently in a context where all the attributes of the good cannot be contracted upon.

There are evidence from the literature that such issues arise in railways. Using the transaction costs theory applied to railways, as suggested by Pittman [2005], Mizutani and Uranishi [2013] test the relevance of vertical separation depending on the network's usage density, used a a proxy for asset specificity. They argue, in particular, that the governance costs should increase rapidly with an increase in network usage as its operation becomes more and more complex. Hence, their proposition is that governance costs should increase if the industry is to be unbundled. Their empirical results indicate that vertical separation is associated to an overall decrease in costs. Nonetheless, when the usage of the network is very important, separation does becomes more costly. Those results corroborate the findings by EVES-Rail (2012).

According to Merkert and Nash [2013], who identify the various interfaces between infrastructure management and train operators, one the most complex part of the transaction, along with day to day operation, is the timetabling part, that is when capacity (also called a *train path*) is allocated by the infrastructure manager to railway undertakings. As we will describe in Section 2, there is the need to coordinate the capacity which is used for commercial use and the one that is used for track work. This issue is consistent with empirical work, as the more dense the use of the network is, the more difficult coordination becomes.

To our opinion, those two results highlight that coordination costs are significant in the railway industry. To tackle this issue, we focus on the French rail transport framework to identify the characteristics of coordination in railways. The transaction for rail capacity raises problems that have to do with limited commitment and non binding communication. Using the literature on those topics we see that a key challenge is to align the interests of both parties. We develop a model to examine the nature and the impact of coordination problems between infrastructure management and railway undertakings. In particular we identify the conditions under which the incentives of both parties diverge. In that sense our results indicate that vertical separation may lead to inefficient outcomes; unless the regulator is able to implement a credible regulation of mark-ups.

The paper is organised as follows. Section 2 provides an overview of the French capacity allocation process. In Section 3 we review the issues identified in the literature which affect coordination process in railways. Section 4 presents our model and section 5 offers concluding remarks.

2.2 Coordination costs in the railway industry : the French case

To illustrate possible coordination issues, we look into the capacity allocation process in France and in particular focus on the structure that was in place between 1997 and 2015. The governance structure chosen in France was what we have described earlier as unbundle with delegation. Therefore we can identify the four players in the sector which are the delegated infrastructure manager, the infrastructure manager, the railway undertaking and the final consumer as described in Figure 2.1. This will highlight the complexity for the railway sector to have market clearing regarding capacity, that is for the capacity made available by the infrastructure manager to match the one needed by railway operators. And therefore why coordination issues arise.

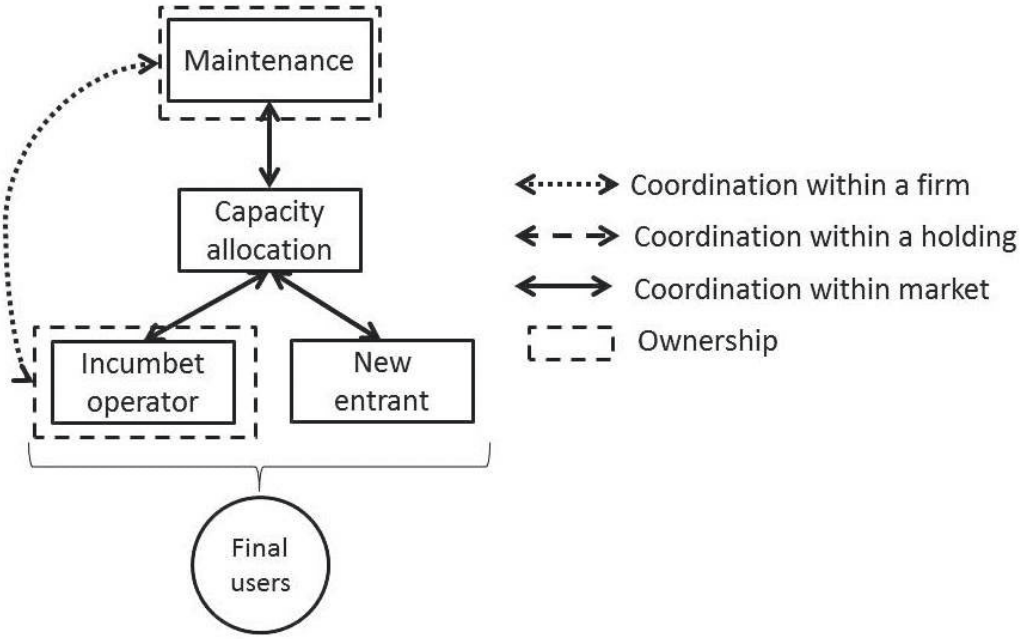

Figure 2.1: Organisation of the French railway sector between 1997 and 2015

The French allocation process can be divided in three stages which are the structuring stage, construction and adaptation.This process is depicted in Fig-

ure 2.2. Before we go over each stage into detail to identify the cause of coordination issues, it is worth noting that it is a long process which can take up to four years.

The first stage consists in **structuring the timetable**. In the French case, this takes the form of a consultation phase beforehand starting 4 years before the circulation date. The objective is for the infrastructure manager to understand the long term needs of the operators. During this stage, there is no binding agreements between the infrastructure managers or the operators. Yet there might be forms of commitment made by both sides, especially on the infrastructure side. Our understanding is that it is during this stage that the IM has to define its general maintenance policy. The infrastructure manager has to contract with the delegated infrastructure manager over the capacity needed for maintenance which will then be unavailable for commercial use.

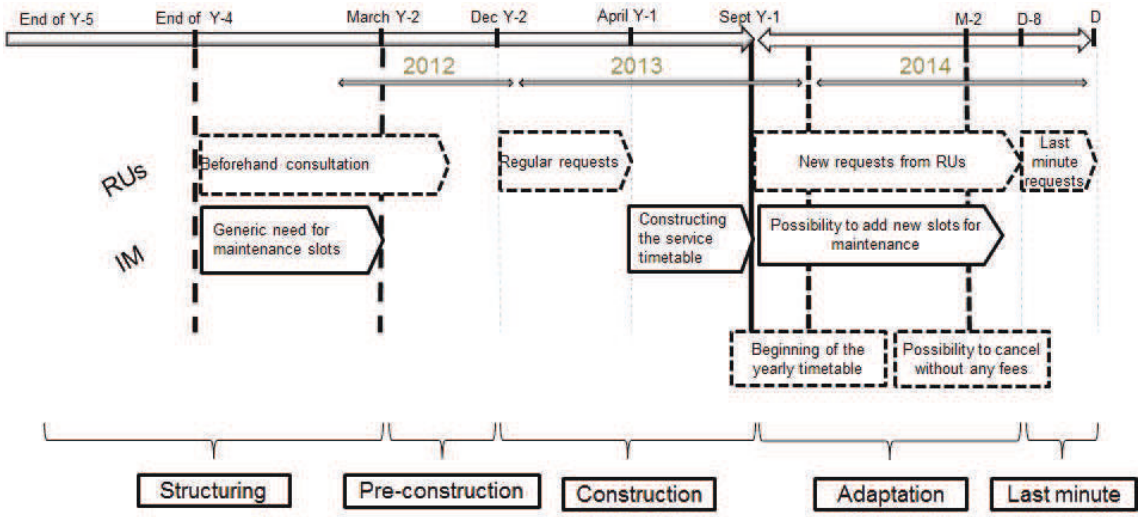

Figure 2.2: The French capacity allocation process

There are several dimension over which both the infrastructure manager and the delegated one have to agree. First off they have to decide what the overall level of maintenance will be. The more trains run on a portion of track, the

more maintenance will be needed. Therefore if the infrastructure manager anticipates a high demand he may ask for the maintenance to be done at night, which is more costly but leaves more capacity available during the day. Another consequence is that he can not mass the maintenance in time. A known way to limit the cost of maintenance is blocking a portion of tracks for a longer time but only having to deploy the maintenance team once[19]. Both infrastructure managers also have to decide where and when some specific asset will be deployed.

This is why we assume that for a given size of the network and when capacity is becoming scarce enough to require a more complex maintenance plan, the more trains run on the network and the more expensive it becomes to maintain the tracks in good condition. For instance in the CATRIN report the authors derive an elasticity of the marginal cost to traffic, and this elasticity increases with traffic. The infrastructure manager has to commit on the price of its maintenance when contracting with the delegated infrastructure manager, but the only information it has are the non binding signals sent by the railway operators. Regarding the operators the 4 year time-span may correspond to its decision to invest in new rolling stocks or other productive inputs such as specialized labor, even though at this point train operators are not committed yet to their customers.

The second stage which is **constructing the service timetable** after the formal requests are made. The length of this second stage has been bounded by directive 2001/14/EC: there should be a working timetable once a year and the deadline for capacity request should be at most 12 months before the beginning of the new timetable[20]. This means that some requests for capacity can still be made up to two years before the train effectively runs. Especially for the freight transport services this time span means that there is environ-

[19] Rapport de la Cour des Comptes
[20] See annex III of the directive 2001/14/EC on the schedule for the allocation process

mental uncertainty as it is very difficult for a railway operator to forecast its demand for transport services in advance. At this point there are many incentives for the undertakings to ask for more than what they need, especially in the freight market for several reasons.

First of all it might be because they are anticipating a contract that has not been signed yet. Secondly they might also be anticipating some negative responses or hazard during the service. This hazard might be due to technical conditions, or may be entailed by the infrastructure manager changing its maintenance slots. Thirdly it might be a strategic behaviour to overbook capacity from the other operators. For instance in France the former incumbent for freight railway services was condemned for such practices in 2012 by the French competition Authority[21]. Table 2.1[22] shows the ratio between capacity used over the capacity that had been booked for Fret SNCF. It is important to note that the figures themselves were not considered as evidence of anti-competitive overbooking (capacity hoarding) acknowledging the fact that operators need spare capacity to face all kinds of hazards. In this case the capacity overbooked represented around 20% of the capacity allocated.

Year	2006	2007	2008	2009	2010	2011
Capacity used	78,27%	78,19%	75,80%	82,06%	77,38%	74,14%

Table 2.1: Ratio of capacity used over capacity demanded for Fret SNCF

We refer to the third and last stage as **adaptation**. During this stage there is trade-off to be made between allocation certainty and flexibility. On the one hand the directive states that the infrastructure manager should be able to "*levy an appropriate charge for capacity that is allocated but not used*[23]". On the other hand the process for allocating capacities should "*have regard to the business requirements of both applicants* (*i.e.* the railway undertakings)

[21] See Decision by the French Competition Authority 12-D-25 of December, 18^{th} 2012 relating to practices used in the freight railway transport sector

[22] Source: French Competition Authority, Decision 12-D-25, paragraph 131

[23] See article 36 of directive 2012/34/EU.

and the infrastructure manager[24] *"*. It does seem that flexibility is amongst the business requirements of a railway operator, especially in the freight railway sector as it is facing competition from other modes of transport. The French infrastructure manager had made this arbitrage by deciding that reservation fees will not be reimbursed if the capacity is not given back 2 months prior to the date the train is scheduled to run [25].

More recently, the rules have evolved toward more flexibility during the adaptation stage. The infrastructure manager has introduced a penalty that increases depending on the notice given when capacity is cancelled. The numbers show, concerning freight transport, that on average over the 6 first months of 2015, the infrastructure manager cancelled unilaterally 4 % of the train paths that had been awarded after the construction stage in September Y-1, while freight operators cancelled on average 14% of the train paths[26]. Regarding passenger trains, there is less variance and the figures are respectively 2% and 1%[27]. Those figures highlight the discretionary power that both parties have to go back on the capacity they asked for, or granted to the train operators, in order to optimize their production.

Coordinating the need for capacity and its demand is both a long and uncertain process. It is worth noting that the trade-off between lowering the overall cost for maintaining the network and making capacity available will become more relevant as the use of the network is increasing. This is in line with the result of Mizutani and Uranishi (2012). Issues may also arise *ex post* that is on the date the train is scheduled to run with real-time coordination between operators and the organism in charge of traffic management. In the absence of a performance scheme where both network users and its manager

[24] See recital 52 of directive 2012/34/EU.

[25] This corresponds to the maximum penalty that can be issued according the French law

[26] Those figures are not constructed the same way as those given earlier for the incumbent, SNCF Fret, and should not be compared.

[27] Source: *Retour d'expérience Incitation Réciproque* V1.16 September, 30^{th}, 2015

internalize the negative effects of disrupting the timetable, these costs might grow. Yet this matter relays more to market opening rather than vertical separation itself. Therefore our read on the situation is that with vertical separation, the coordination costs will arise because of the need for flexibility along the allocation process while commitments in specific inputs are made.

2.3 Coordinating for capacity in railways

From the contractual theories of vertical integration, we can derive that key driver of switching from an integrated vertical structure to independent units is that coordination has to be achieved through prices. As pointed out by Coase [1937], prices might replace authority as a mean to coordinate, or as in the model of vertical integration developed by Hart and Tirole [1990] a separated structure cannot set profit sharing mechanisms between the upstream and downstream entities. Indeed, vertical separation may create a breeding ground for hold-ups and therefore limited investments of parties that anticipate opportunistic behaviours (Klein et al. [1978] and Williamson [1985]).

In the case of railways in the short or medium term, the problem faced by both participants can be related to the news-vendor problem that was first described by Edgeworth [1888] and latter formulated by Arrow et al. [1951]. Where you have to produce a time sensitive product, demand forecast will be an issue. If you produce too much capacity it will go to waste, on the opposite, if you produce too little demand will not be fulfilled without the opportunity to adjust. This capacity issue will be empathized along a supply chain with an extra stage of coordination between the upstream and downstream entities as described in Hart and Tirole [1990] and Rey and Tirole [2007] in the case where firms cannot commit on quantities.

Legros and Newman [2009] gives us an insight as to when coordination will

be an issue. It is the case when both parties along the vertical chain have divergent objectives in terms of quantities. As pointed out earlier in the allocation process in France, it seems to be the case as railway undertakings will prefer that too much capacity be produced so that they can adapt, while the infrastructure manager can reduce its unit cost in the short run by producing less. The trade-off is therefore similar to the one made explicit by Legros and Newman [2009], between minimizing the coordination costs while producing at a higher price when the firm is vertically integrated.

Along with the trade-off on quantities, the nature of communication is also of relevance. In the case we have described earlier, we have seen that commitment on both sides is very limited despite frequent interactions between upstream and downstream firms. Because firms cannot commit, the importance attached to the message sent when the time tabling is being built, is necessarily smaller and damages the quality of information' transmission. This limited commitment can be be related to cheap talk as described by Crawford and Sobel [1982] in their seminal article. Cheap talk studies precisely the case where agents can only give information on their future action using a non binding signal. It does avoid some coordination failures. But one of the key insight is that there can only be a credible signal if the interests of both parties are aligned enough as pointed out by Farrell and Rabin [1996]. Therefore in a non dynamic framework, this highlights the difficulties of coordination in railways due to divergent objectives in terms of quantities.

A strand of literature has also studied the incentives for information sharing in a supply chain following the first article by Novshek and Sonnenschein [1982] and Gal-Or [1991]. As presented by Li [2002], there are limited incentives for the downstream firm to share its information for two reasons. First off there is a direct effect to sharing the information with the upstream firm which might react in a strategic manner, for instance by rationing its supply of train paths in order to mitigate the risk of producing too much. There also is

an information leakage effect. By observing the reaction of the upstream firm, the other downstream firms can deduce the information their competitors have shared. For instance, if a train operator observes that the infrastructure manager is making a lot of capacity to go to a particular destination, it can deduce the demand of its competitors, and that they believe that this destination will be lucrative.

The incentives to withhold information are going to increase when the governance structure does not guarantee the independence of the infrastructure manager. Indeed when the separation between infrastructure and operation is not strong enough, the infrastructure manager can feed directly information to the incumbent train operator. It is consistent with the findings of Decision by the French Competition Authority relating to practices used in the freight railway transport sector[28]. This decision stressed that the leakage of information enabled the incumbent operator to understand the strategy of new entrants. Such practices necessarily trigger less cooperation from new entrants. In this regard, it is worth noting that the forth railway package shifts the timetabling process as an item of the "minimum access" package[29]. Hence it may be subject to a more stringent regulation in terms of transparency.

A key issue behind transparency lies in the risk of foreclosure. As pointed out in the framework set by Rey and Tirole [2007], *"vertical foreclosure can be motivated by the desire to restore a market power that is eroded by a commitment problem".* In other words, setting vertical restraints can be a means for the infrastructure manager to overcome the coordination issue in a manner similar to the benefits of vertical integration (Hart and Tirole [1990]). Transparency will mitigate the risk of foreclosure, and improve market outcome. Normann et al. [2015] provide experimental evidence that both open communication and direct communication between upstream and downstream firms

[28] Decision 12-D-25 of December, 18^{th} 2012

[29] The minimal services that the infrastructure manager is obligated to supply to an operator requesting access to the network.

can solve the commitment issue for the supplier.

Communication in railways has been implemented *via* the time-tabling procedure. This procedure is essential to achieve central planning and optimize the use of the infrastructure as highlighted by Quinet [2003]. But it raises concerns due to limited commitment and information asymmetry. It is not certain if downstream firms have the incentive to reveal the private information they have on the capacity they need. In such a case, information sharing might face some limits, especially in a context of constrained capacities. The question has therefore been discussed if auctions could be introduced in order to reveal preferences of users and award the capacity to the operator which values the most its use. Auctions could limit the cases of capacity hoarding from the dominant train operator by making this strategy more costly. But it should be noted that there are no clear result on the feasibility of introducing auctions. If Caillaud [2003] describes a framework for introducing auctions in railways at the train path level; Nash [1999] on the other hand stresses the difficulties of introducing auctions in railways. Beyond the ease of implementing auctions in railways, such a solution would mostly solve very short run issues, arising from conflicts between train operators, but do not necessarily achieve coordination in the medium run.

It is worth noting that similar concerns have been identified in other network industries, such as electricity or gas market as pointed out in Marty [2012]. The strategies of vertically integrated operators have been criticized on the accounts that they artificially created congestion, or under invested in the development of capacity to deter potential competitors in the downstream sector. Such strategies give credence to the possibility of opportunistic behaviours coming from an integrated infrastructure manager in the case of railways.

In the following section, we suggest a different solution which can be imple-

mented by the regulator in charge of reviewing the access charges on the network. We take into account the need to align interest between the upstream firms and the downstream firms in order to enhance coordination in railways.

2.4 The model

2.4.1 Description of the model

We consider the two players in the railway industry, that is the infrastructure manager (IM) on the upstream market in charge of producing and allocating the railway paths and the railway undertakings (RUs). The production cost of the IM depends on the maintenance policy. If more capacity has to be made available then this extra capacity will be costly. For two level of quantities produced q_H and q_L with $q_H > q_L$ we denote k the average cost the IM has to commit to in order to increase its production. (The cost of production will increase by $k(q_H - q_L)$). In the downstream market, railway undertakings[30] are competing to sell railway transport, where each unit of railway transport requires one path. The average marginal cost an operator has to commit to in order to increase its production from q_L to q_H is denoted c.

We focus on the downfalls of separation arising from the lack of coordination between both players creating a possible tension between the quantities served by the IM and the effective use of capacity by the RUs. These tensions rise due to uncertainty in the final demand and the opportunistic behaviour arising from flexibility on both sides. To take into account the uncertainty of the demand for transport services, we assume there are two states of nature denoted L and H where H corresponds to the state of nature with a positive shock in demand compared to L which is the standard demand for rail trans-

[30] Formally the infrastructure manager will be interacting with one representative railway undertaking.

port services. The prior distribution for each state is common knowledge and is defined by $Pr(H) = \pi$ and $Pr(L) = (1 - \pi)$.

For a given state of nature the optimization program will lead the RUs to an equilibrium price for rail transport services and a quantity. We denote Mu_H (respectively Mu_L) the mark-up the downstream firms are able to levy above the cost of production when demand will be high. q_H and q_L are the quantities served associated to those prices[31]. We make the assumption that $Mu_H > Mu_L$ and $q_H > q_L$. When the demand is high, the quantities served will increase as well as the prices.

The network manager can either choose to produce q_H or q_L. We assume that the access charges pricing scheme allows to recover marginal costs of production, and that the IM may levy a mark-up if it had anticipated a high demand, and the demand is indeed high. This is a consequence of the right to price above marginal cost in order to recover its full cost as stated in directive 2001/14/EC[32]. In our model the IM is allowed to levy a mark-up if the market can bear it, that is if the demand is high but also if the IM had anticipated a high demand. We denote the mark-up of the IM Mu^{IM} and assume it is fixed exogenously implying that a regulator has to give its assent to any mark-ups from the IM[33]. Such a pricing strategy is an adaptation of Ramsey-Boiteux pricing (Boiteux [1956]) where we consider two markets for capacity: one where demand is high and downstream users become more captive in which case the IM is allowed to price above marginal price, the other one with a low price elasticity on the downstream market and triggers marginal

[31] Note that using Cournot competition to model the interaction between the downstream firms, we would have that $Mu = \frac{(QP'(Q))}{n}$ where P(Q) is the inverse demand function and P'(Q) its derivative. Therefore the shift in mark-up will depend both on the shift in elasticity and the overall increase in demand.

[32] Article 8.1 of directive 2001/14/EC: In order to obtain full recovery of the costs incurred by the infrastructure manager a Member State may, if the market can bear this, levy mark-ups on the basis of efficient, transparent and non-discriminatory principles, while guaranteeing optimum competitiveness in particular of international rail freight.

[33] By assuming that Mu^{IM} has no effects on q_H we neglect part of the downfalls of double marginalization.

cost pricing.

To highlight the potential coordination issues, we model the outcome of the market using a normal form game. We assume that both agents have to commit to a quantity produced and to a maximum price. Therefore both the infrastructure manager and the railway undertaking have a set of action $A = L; H$. Once they observe demand, they have limited possibilities to adjust their offer. They can only sell less than what they had produced and may only lower their price. Given that we need the supply of capacity to equate its demand, we need the following rules on the coordination:

1. If a firm played H and the other firm plays L, then it has to sell less, that is selling q_L instead of q_H

2. If a firm played L and the state of nature is H, it has to sell at a lower price, that is levying Mu_L instead of Mu_H (or 0 instead of Mu_{IM} in case it is the IM)

The first rule signifies that if not enough train paths where produced, extra trains cannot run on the network, or conversely if the operator has not planned on supplying high quantities on the downstream market, train paths will go to waste. In other words, it is too late to produce more once you observe the production from the other player. The second coordination rule means that the firm have to commit to their pricing strategy in case demand is unexpectedly high. For the IM it implies that the regulator does not accept to renegotiate the access charges. On the downstream market, it suggests that the railway undertaking does not have the means to change its commercialisation strategy over the time span. This can be the case either because it has define in advance its pricing strategy, or because for instance it planned on offering a low cost service and has adjusted its rolling stock accordingly.

We enforce those two rules in order to have market clearing. The payoffs are represented in Table 2.2.

IM \ RU	H	L
H	IM: $\pi(q_H Mu^{IM}) - (1-\pi)(q_H - q_L)k$; RU: $\pi q_H(Mu_H - Mu^{IM})$ $+(1-\pi)[(q_L Mu_L - (q_H - q_L)c]$	IM: $\pi q_L Mu^{IM} - (q_H - q_L)k$; RU: $q_L[Mu_L - \pi Mu^{IM}]$
L	IM: 0 ; RU: $\pi q_L Mu_H + (1-\pi)q_L Mu_L - (q_H - q_L)c$	IM: 0 ; RU: $q_L Mu_L$

Table 2.2: Payoffs matrix

2.4.2 Outcome of the game

Since there is no administrative control with vertical separation we could have two uncoordinated equilibria. In both those outcomes, either the network manager or the operator are left with unsold goods. This leads to inefficiency in the railway sector. We focus our analysis on the conditions for those two inefficient outcomes to occur, and in particular how the mark-ups levied affect the coordination first for the infrastructure manager, then for the operators.

If the outcome of the game is HL, that is the IM anticipates a higher demand than the RU, then the cost of spoilage would be $(q_L - q_H)k$ which has to be balanced with the potential mark-up in case the demand is high. In order for HL not to be a Nash equilibrium, we need that L be the best reply to L. Thus the condition is that:

$$Mu^{IM} < \frac{1}{\pi}\frac{q_H - q_L}{q_L}k \tag{2.1}$$

On the other hand, if we want LH not to be a Nash equilibrium we need to make sure that the potential mark-up is greater than the potential costs. The condition is that :

$$Mu^{IM} > \frac{(1-\pi)}{\pi}\frac{q_H - q_L}{q_H}k \tag{2.2}$$

As we can see, this leads to an upper and lower bound[34] for the mark-up of the IM if we want to avoid uncoordinated equilibria. The upper bound represents the fact that if the IM is able to levy too-high a mark-up when the conditions are met, then there will be more incentive to serve higher quantities, despite the low odds of the demand actually being high. The right-hand side of equation (2.1) represents the expected loss in revenue when the IM systematically plays a high quantity. This loss should be greater than the reward it gets when demand is actually high.

The lower bound is the necessary counter part so that the IM assumes the risk of a low demand and leading to spare capacity. The right-hand side of equation (2.2) is the expected loss for the IM that does not systematically produce low quantities. This expected loss has to be lower than the mark-up. In equation (2.2) we identify the risk premium that has to be awarded to the IM.

Result 1 : Should $Mu^{IM} \in [\frac{(1-\pi)}{\pi}\frac{q_H-q_L}{q_H}k; \frac{1}{\pi}\frac{q_H-q_L}{q_L}k]$ then no uncoordinated outcome can be a Nash equilibrium.

Result 1 states that we can limit the behaviours of the IM leading to inefficient outcomes by bounding its mark-up. Given that the pricing scheme of the infrastructure manager can be reviewed by an independent regulator in the European Union, it should be possible to bound the mark-up. The lower bound gives us the amount of incentive the IM needs in order to avoid producing low quantities. If we rule out the possibility to have access charges above marginal prices when the market can bear it, then the network manager will never risk producing a high amount of capacity. The upper bound is needed so that the infrastructure manager does not always anticipate high quantities. If the allowed mark-up is too high, the regulated monopoly will be over-producing.

[34] Given that $q_h > q_L$ then we always have that $\frac{(1-\pi)}{\pi}\frac{q_H-q_L}{q_H}k < \frac{1}{\pi}\frac{q_H-q_L}{q_L}k$

In our setting, the mark-up compensates for the absence of outside option for spare capacity. If there was an outside option for at least a part of spare capacity, this would shift the interval downwards. Should competition in the downstream market bring such an outside option, then the IM would have an incentive to produce higher quantities, at the risk of being an overproducing monopoly. If we consider that the IM can only have price above its marginal cost in order to recover its full cost, as stated in directive 2001/14/EC, then an extra condition for this result to hold is that the difference between full cost and marginal cost be within the bounds of the interval we have defined.

With the full separation, the pricing scheme can be an important tool for the regulator to mitigate the risk of an uncoordinated outcome to occur as well as inducing the capacity made on the network. With a holding structure, those incentives will be dulled by the IM internalizing part of the downstream market's overall profit as described for instance in Cremer et al. [2006]. This may raise as well the issue of cross-subsidies. Then the IM is less legitimate to levy mark-ups if it is seen as a part of a squeeze strategy towards alternate train operators.

We now check the condition under which the RUs favor a coordinated equilibrium as well:

- H is the best reply to H if:

$$(Mu_H - Mu_L) > \frac{q_H - q_L}{q_L}[\frac{(1-\pi)]}{\pi}c + (Mu^{IM} - Mu_H)] \tag{2.3}$$

- L is the best reply to L if:

$$(Mu_H - Mu_L) < \frac{1}{\pi}\frac{q_H - q_L}{q_L}c \tag{2.4}$$

Similarly to the IM case, on the one hand (2.3) is the risk premium for the operator. The increase in revenue when demand is high has to be greater than

the possible increase in cost for unused capacity ($\frac{(1-\pi)]}{\pi}c$) and the sure increase in input prices ($Mu^{IM} - Mu_H$). On the other hand (2.4) is the condition needed so that the train operator is not always over-producing. Both conditions give us an upper and lower bound in order to avoid an uncoordinated outcome. Note that for those two conditions to be met at the same time we need that $Mu^{GI} - Mu_H < c$.

Result 2 : Should $(Mu_H - MuL) \in [\frac{q_H - q_L}{q_L}[\frac{(1-\pi)]}{\pi}c + (Mu^{IM} - Mu_H)]; \frac{1}{\pi}\frac{q_H - q_L}{q_L}c]$ then no uncoordinated outcome can be a Nash equilibrium.

Contrary to the IM's case, the increase mark-ups is an outcome of the market and can not be regulated when there is downstream competition. The upper bound of the interval directly refers to the shift in elasticity of the firm and its ability to capture it. As competition increases in the downstream market, the difference in mark-ups should tend to zero and the upper bound of our interval has less chance of being met. Conversely with little competition on the downstream market, the odds of having the downstream firms favoring the possible increase in its revenue to the sure increase in cost are more important even with a low probability of demand being high.

The lower bound of our interval is more problematic as it sheds to light a possible cause for an inefficient outcome. This threshold states that the increase in mark-up should cover the uncertain increase in costs due to over-producing plus the mark-up of the infrastructure manager. As competition is increasing, then the mark-ups a downstream firm are able to levy will not be high enough to compensate for this risk. Therefore the more competition, the more the downstream firms tend to play the low quantity outcome. Rewriting the threshold as :

$$Mu^{IM} < \frac{q_H Mu_H - q_L Mu_L}{q_H - q_L} - \frac{(1-\pi)}{\pi}c \tag{2.5}$$

We identify the condition in our model under which the downstream market can bear any deviation from marginal cost pricing. The mark-up of the in-

frastructure manager becomes harder to bear for the railway undertakings as their market power decreases and as competition is increasing it becomes more prejudicial to the quantities served on the market.

Result 3 : If

$$(1-\pi)[\frac{q_H - q_L}{q_H}k + c] \leq \frac{q_H Mu_H - q_L Mu_L}{q_H - q_L}\pi \tag{2.6}$$

then there exist a mark-up Mu^{IM} such that HH is a Nash equilibrium.

This result stems from the conditions for H to be the best reply for both the IM (2.2) and the RU (2.5). For both conditions to be met, we need that:

$$\frac{(1-\pi)}{\pi}\frac{q_H - q_L}{q_H}k \leq \frac{q_H Mu_H - q_L MuL}{q_H - q_L} - \frac{(1-\pi)}{\pi}c \tag{2.7}$$

which is equivalent to our third result. Therefore, if this condition is met, it is possible for a regulator to set a pricing scheme for the infrastructure manager such that this coordinated outcome of the game is a Nash equilibria. In order for the high quantity equilibrium to be sustainable, we need that the expected average increase in revenue in the downstream sector be greater than the expected unnecessary increase in costs, which consists in the cost of production for the downstream firm c and the minimum risk premium awarded to the infrastructure manager $\frac{q_H - q_L}{q_H}k$. This result states, as suggested earlier, that if the margins on the downstream market are limited due, for instance, to competition, then the high output equilibrium becomes harder to sustain.

It also stresses the importance of the need for flexibility in the downstream market. In our model it is the absence of outside option for the railway undertakings that leads to too little production. The unit cost c is the cost of committing to more input and is a sunk cost for a railway operator. The absence of outside option can be seen as a fairly strong assumption. It depends on the specificity of the assets used by the downstream firm.

Similarly, regarding the low quantity outcome, the conditions for it to be a Nash equilibrium are the following:

- Condition IM: $\pi Mu^{IM} \leq \frac{q_H - q_L}{q_L} k$ (2.1)
- Condition RU: $\pi(Mu_H - Mu_L) \leq \frac{q_H - q_L}{q_L} c$ (2.4)

They correspond to the upper bounds of the interval we have defined respectively in result 1 and result 2. Remember that the condition of the IM can always be met by using price regulation, while the bound for the RU might be endogenous to the market. In particular, we next investigate what happens when the difference in margin for the RU is sufficiently important as to give an incentive to always be other producing, that is if:

$$\pi(Mu_H - Mu_L) \geq \frac{q_H - q_L}{q_L} c \tag{2.8}$$

In this case, is it always possible for the regulator to avoid that the equilibrium be uncoordinated, that is to set the mark-up of the infrastructure manager such that the outcome will be the high quantities equilibrium?

(2.8) can be rewritten as:

$$\pi \frac{q_H Mu_H - q_L Mu_L}{q_H - q_L} \geq c + \pi Mu_H \tag{2.9}$$

This equation means that the RU will choose the uncoordinated outcome LH if its average increase in revenue when demand is high is greater than the cost of producing plus the unsold units because there is not enough train paths to run its trains even though demand is good. The necessary condition (2.9) as stated in result 3 for HH to be a Nash equilibrium is is satisfied as long as: equationMu$_H + c \geq \frac{1-\pi}{\pi} \frac{q_H - q_L}{q_H} k$The interpretation is that if the cost of spillage on the left hand side of the inequation is greater than the risk premium that has to be awarded to the infrastructure manager, then we can switch to the high output equilibrium.

Result 4: If $\pi(Mu_H - Mu_L) \geq \frac{q_H - q_L}{q_L} c$ and $Mu_H + c \geq \frac{1-\pi}{\pi} \frac{q_H - q_L}{q_H} k$ then there exist a mark-up Mu^{IM} such that HH is a Nash equilibrium. Furthermore, the only Nash equilibrium possible is the coordinated outcome with high quantities.

In other words it is possible to trigger the high output equilibrium if the two following conditions are met when the train operator switches to producing high quantities : the increase in revenue is greater than the cost increase, and at the same time the cost of spillage for the operator is greater than the risk premium for the infrastructure manager.

2.4.3 Discussion

In the model developed, we address the problem of coordination between an infrastructure manager and a train operator which arguably is one of the important drawback to vertical separation. We argue that a better market design and in particular a better design of prices might mitigate that risk making this vertical structure more sustainable. Yet the setting we use is not relevant to compare an integrated and a separated structure. Firstly there are several dimensions to be taken into account when analysing the vertical organization. In our model we elude the potential losses due to economies of scope with vertical separation. Secondly we implicitly assume a dichotomous repartition of coordination costs where only a separated form would face such issues and that in a integrated mode, coordination can be achieve. Nevertheless if the conditions of our fourth result are not met, even an integrated firm could decide to have uncoordinated quantities.

The second issue that is not addressed in the model is the one of optimal amount of capacity produced by the infrastructure manager. By making the

analogy between capacity production and quality, the problem we describe fits the one described by Spence [1975]. Either in the case of an integrated monopoly or a separated infrastructure manager, prices do not necessarily convey the preferences respectively of the final users or downstream firms in terms of capacity. That is to say that defining q_H and q_L would be a challenge.

Similarly, to develop further this model, we would need to take into account the asymmetric information between the upstream and downstream markets as well as within the downstream firms. The operators should have a better understanding of the demand and our model does not take into account the strategic behaviours this could entail. But the underlying reasoning we can put forward is that if the pricing scheme set out by the regulator aligns sufficiently the interests of the upstream and downstream firms, then the interactions during the allocation process could lead to meaningful communication.

We also do not capture the disincentive to share information with the infrastructure manager if it means sharing it with the competing operators. This effect could exist with train operators that have a perfect rationality. But it will especially be the case with more integrated structures that cannot enforce a *Chinese wall* in terms of information sharing between infrastructure management and train operations. Therefore if an integrated structure should facilitate information structure within the holding, it can have a negative effect on the information shared by new entrants.

Regarding the downstream market, the model does not take into account the effect of competition in the downstream market and on the profits made. As pointed out earlier, as competition increases, the mark-ups a firm can levy on a market should decrease and tend to zero. Nonetheless the mark-ups which are contingent on the state of demand might vary differently. If a train operator is always able to levy a mark-up when demand is high, even though a firm cannot make any profit when demand is low the results and the possibility to

sustain a high output equilibrium would still hold.

Another lever to increase the possibility of having a high output equilibrium is to decrease the cost of spoilage in case there is no coordination. Indeed in our model, coordination issues are linked to the absence of outside option once the firm has committed to a certain level of production. Yet this relies on the fact that a train operator can not redeploy its productive factors to another use. In a European rail market where interoperability remains limited such a hypothesis is realistic. Indeed there still are differences in terms of stock, technology, signalling systems and safety regulations[35]

We could argue that as the market becomes more competitive the outside option is increasing for the infrastructure manager who could assign the capacity to a different railway undertakings. Regarding the outside option of the railway operator it should be able to redeploy inside a same country. Increasing interoperability across networks could also limit the level of sunk costs an operator has to bear if productive factors can be used in an other country.

This argument is in line with the findings of Marty [2012] in the case of the electricity sector. According to the author, the risk of foreclosure, that is an integrated network operator adopting an opportunistic behaviour in our framework, is greater when the inter-connexions between networks are limited. The existence of an outside option means that an actor can leave the market which is one of the conditions of having a contestable market according to Baumol et al. [1982]. In our framework, interoperability also benefits the infrastructure manager, in case for instance of vertical separation, which cannot be held up by a dominant train operator.

[35] Source: EUR-Lex - l24015

2.5 Conclusion

In this paper, we focus on coordination problems arising between the upstream (infrastructure access management) and the downstream (rail services activities) markets, and in doing so contribute to the debates over the relevance of vertical separation/separation in network industries. Our paper highlights that a key step of the coordination between the infrastructure manager and the operators is the capacity allocation process. The railway sector is facing uncertainty on the final demand due to the overall length of this process. Therefore the attribution process was made very flexible, since both sides do not have to commit to one another but they do exchange information over their need. Yet the firms might have to commit themselves to the total input (maintenance operations) they will be using in the case of the upstream firm, and the total output in the case of the downstream firm. This could lead to uncoordinated market outcomes, thus a loss of performance for the railway sector.

This issue of limited commitment is central in the literature on vertical interactions, as it creates a risk that coordination becomes more costly and lead to vertical restraints from the upstream firm in order to restore its monopoly power. In order to avoid such an outcome, the objective should be to align the interest of both the upstream and downstream firms so that communication can become meaningful and lead to a better communication. In a regulated industry such as railways, the regulator can enforce pricing rules in order to enhance coordination. Thus the intuition given by Hogan [2002] that "*market are made and don't just happen*" does not apply only to the electricity sector but also railways.

In our model we analyse the conditions to avoid any uncoordinated outcomes, depending on the the mark-ups levied upstream and downstream. When the upstream mark-ups are under the supervision of a regulator, the infrastructure manager can be steered toward a coordinated outcome. In order to do so, the

regulator has to award a risk premium to the infrastructure manager who anticipates high quantities and depart from marginal pricing. Therefore with an effective price regulation the network manager will not choose outcomes that are inefficient for the railway sector.

Regarding the downstream market, the conditions for the railway undertakings to favour a coordinated equilibrium depend on the market structure and might be out of the control of a regulator. Especially as the market power of downstream firms decreases the lower threshold necessary to have a high output equilibrium is more problematic. There are two implications to the lower threshold we have identified. Firstly we obtain the condition for the downstream market to bear any deviation from marginal cost pricing for the access charges. It is directly related to the market power of downstream firms. Secondly an increase in the outside option of downstream firms for their spare inputs would increase their incentive to deliver high outputs, even with downstream competition. We believe this outside option could be a higher interoperability between European networks making it possible to reallocate productive inputs from one country to another.

In terms of public policy recommendations our results highlight the role of regulating access charges. Deviations from marginal cost pricing are needed to mitigate the coordination problems associated with vertical separation and price regulation should aim at aligning the interests of the infrastructure manager and train operators in terms of quantity. In which case the regulatory agency can have a positive impact on the outcome vertical separation.

Chapter 3

The Role of Expertise on efficiency in Public Procurement: The Case of Regional Railway Transport in France*

3.1 Introduction

When investigating the efficiency of public procurement, the economic literature has stressed the major role information plays. More specifically it identifies how the informational constraint impacts the outcome for the public buyer. Following the lead by Laffont and Tirole [1993], those considerations led to public recommendations on the contractual scheme that should be set in place. Those recommendations evolved from an incentive regulation, such as a fixed price contract, aimed at exerting a greater cost reducing effort from the operator (Cabral and Riordan [1989]), to offering a menu of contract for the operator to reveal its type in addition. In doing so, economists recognize the key role the public buyer has for an efficient public procurement, especially its

* This chapter is based on a joint work with Miguel Amaral. The authors would like to thank Matej Drev, Aude Le Lannier, Pierre Ravier as well as the participants at the 6th International Conference on "*Contracts, Procurement, and Public-Private Arrangements*", at the ISNIE Conference (2015) and at the ESNIE Days (2014) for their valuable comments .

ability to find the second best contract in a context of information asymmetry.

Yet, going back to the seminal paper by Jensen and Meckling [1976], the principal also has a monitoring function which is often overlooked at the expense of minimizing the residual loss. As pointed out by Saussier and Tirole [2015] in the case of public procurement, the public buyer often "lacks of data for monitoring changes in the public procurement system and analysing its performance, meaning that the *ex ante* and *ex post* control of contracts is consequently limited." This is in line with Brown and Potoski [2003] who suggest that a government needs both implementation capacities, in other words the ability to negotiate a contract but also the capacity to evaluate the performance of the operator. These two components being the expertise of the public buyer which, arguably, can be too limited (Saussier and Tirole [2015]).

The need for expertise can be explained by the strategic behaviour the operator might set in place regarding the information it passes on to the principal. For instance, Laffont and Tirole [1992] describe a framework were the operator can do cost padding and therefore increases its informational rent. Indeed Rogerson [1992] argues that, even in the case of a price cap contract, a firm with overhead costs will claim it has high costs even though its costs are low, spend the money and shift its assignment. In a dynamic framework, more strategic options are available to the operator. Fudenberg and Tirole [1995] present a strategy of income smoothing when the players anticipate that they will be evaluated mainly on their recent performance, considered as more representative. In the case of public procurement such strategies are expected before the contract is renewed (Iossa and Rey [2014] and Affuso and Newbery [2000]). Any good performance exhibited in front of the public buyer has to be weighed against the future assignment of more ambitious targets (Weitzman [1980]).

The changes that occurred more than a decade ago in the regional railway

sector in France make it an interesting case to study the outcome of public procurement. In a global movement towards more decentralized public policy decisions in France, the regions (*i.e* the largest administrative division in France) have become responsible for organizing public transport on their territory from 2002 on. Each of the 20 regions now award directly and for a limited duration a Public Service Obligation (PSO) contract to an "*operator*" to run its regional railway services. But since the incumbent still benefits from a legal monopoly when it comes to running national passenger services, the transport operator is the state owned monopoly, organized in regional directorates specific to each PSO contract. The regions do not either have the possibility to operate the service in house so procurement contracts are set out in a non-contestable market. This contrasts with the policy chosen for other local public services in France such as local urban transport where procurement contracts are put to competitive tendering. On the other hand, the setting is similar to electricity distribution in France, where the service is delegated by the municipalities to ERDF, a subsidiary of the historical operator. A description of the market can be found in Coelli et al. [2013].

Our objective is to analyse the effect of expertise on the efficiency of PSO contracts in the context of the French regional railway services. To this aim, we identify the potential strategic behaviours regarding the display of performance that can be put in place by the operator in the presence of asymmetric information. If the intensity of such strategic actions vary depending on which transport authority the operator is facing, then the expertise of the public buyer should have an impact on the outcome on public procurement. Based on the common structure of the contracts as well as the setting of the market, we identify theoretical predictions.

The theoretical propositions are corroborated relying on an original data set gathering information on regional contracts over the period 2009-2012 and the use of a stochastic cost frontier analysis. Our work is in way similar to those

of Gagnepain and Ivaldi [2002], Dalen and Gómez-Lobo [2003] or Gautier and Yvrande-Billon [2013] who study the role of regulatory schemes on the cost efficiency in the case of public urban transport. To the extent that in our case, the nature of the contract does not vary and we benefit from a very homogeneous set of operators, we can observe the actions of the operator without firm specific effects.

Our estimations' results exhibit the existence of differences in efficiency amongst the regional companies of the incumbent and also exhibit that the provisions in the PSO contracts do not produce the expected effect in terms of cost reduction incentives. Regional operators adapt their strategy both in time during the contract execution and also according to the authorities they are facing to maximize the rent they can extract out of the PSO contracts. In other words, the operators make the best of the design of the contracts. At last, our result highlight that the relative efficiency is decreasing during the duration of the contract. The heterogeneity we observe in terms of efficiency also corroborate the intuition that efficiency benchmarking is of relevance in the sector and as so, contributes to the debates both over the implementation of yardstick competition and on the relevance of the decentralization of regulatory decisions.

The paper is organized as follows. In section 1 we provide a description of the regional railway transport sector in France. Section 2 presents the testable propositions regarding the impact of the contract's design on the technical efficiency. In section 3 we develop our empirical strategy and present the original data set used for the estimations. In section 4 we present and discuss the empirical results before discussing the policy implications in terms of regulation and decentralized public procurement to finally conclude in section 5.

3.2 Regional railway transport in France

3.2.1 Organisation of the sector

The regions (*i.e* the largest administrative division in France[36]) have become responsible for organizing public transport on their territory since 2002 and the law called SRU[37]. The transport services encompass suburban and intercity trains, inside the region as well as with neighbouring regions in some cases. Each of the 20 regions awards directly and for a limited time a Public Service Obligation (PSO from now on) contract to an operator to run its railway services. The length of the contracts ranges between 6 to 10 years with an average of 8.15 years in the period we study. As the incumbent (SNCF) still benefits from a legal monopoly[38] when it comes to running national passenger services, the operator is the state owned monopoly, organized in regional directorate specific to each PSO contract. Given that the regions do not have either the possibility to put the contract to competitive tendering or to operate the service in house, the market is non-contestable[39] despite the fact that procurement contracts are set out.

The new organisation of the sector coincides with a higher provision of services for regional public transport since 2002. As reported in the Haenel [2008] report the number of train-kilometres has increased by 30% between 2002 and 2012 and the number of passenger-kilometres has increased by 51% as well. This is mainly due to the fact that the regions have both increased the supply

[36] The average area is 26,000 km^2.

[37] "*Loi relative à la solidarité et au renouvellement urbains*" - SRU - (Law n° 2000-1208 of December 13^{th}, 2000).

[38] "*Loi d'orientation des transports intérieurs*" - LOTI - (Law n° 1982-1153 of December, 30^{th} 1982) consecrates in the article 18 the monopoly of SNCF for passenger services.

[39] This setting is currently under discussion since the fourth railway package (a bundle of European legislation regulating the railway sector that should soon be adopted) could open the market to competition in a decade or so. The fourth railway package should be the final step to the process of competition reforms that the European Commission started in 1991. This would of course be a major reform for competition policy in a sector that has been developed by publicly owned monopolies.

of services and have put in place their own pricing strategy on the travel cards. The expansion of regional railway transport has had a cost for the transport organizing authorities since their contribution has increased by 96% during the same period which is a result of the cost increase of new services and often lower prices for the customers. As a consequence the regional railway transport has become an increasingly subsidized sector over the years: the commercial revenues cover on average 27% of the operating expenditures[40]. The overall budget for PSOs contracts added up to 2.8 billion Euros in 2012 and represented on average 18% of the budget of the regions[41], (which makes it on average their second largest budget item after education which represents on average 3 billion Euros per year.)

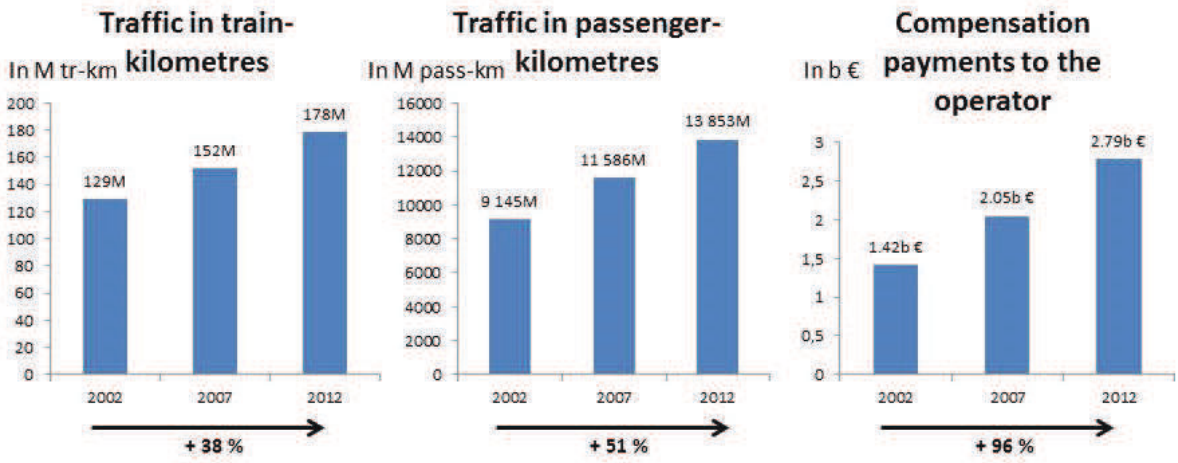

Figure 3.1: Evolution of the scope of PSO contracts in regional transport since the 2002 decentralisation

3.2.2 Scope of the contracts

The scope of activities that are attributed by the contracts to the operators is overall the same between the 20 regions. The operators of the PSO contracts are in charge of all the operations regarding the service, that is operating and maintaining the rolling stocks, supplying the energy (fuel and electricity) to

[40] Source: *Ville Rail et Transport* magazine.
[41] Website of the *Association des Régions de France*.

run the trains and managing the access to the infrastructure (train paths and stops at stations). Regarding the commercialization of the service, the operator is also in charge of selling the tickets, collecting the revenues from fares and enforcing ticket checks on board. The staff necessary to run the services is provided by the operator and are not employed by the regions. Regarding the rolling stocks, it is owned by the operator although the regions subsidize the whole cost of purchasing new rolling stocks. The rolling stocks are attached to one transport authority, but can occasionally be redeployed in another region when operating the services requires it. Therefore the rolling stock is not a highly specific physical asset whose development is borne by the operator.

In regards to the means to produce the service, due to the organisation of SNCF at that time, the operator delegates part of its tasks to other directorate of the monopoly. This represents an important part of the cost drivers. The management of the rail network is done by a separate entity since 1997[42] as is the management of train stations[43]. Also many operations are delegated to task-specific national directorate of the incumbent. For instance the staff for operating and maintaining the trains is usually employed at a corporate level. So regional operators have access to common inputs and at the same prices. An important consequence is that the regional operators are quite comparable in the inputs and the technology they use.

The scope of the contracts highlights that the regions chose to delegate a vast majority of operating decisions to the operator, even though their status of transport organising authority could have led them to assume more decisions. As a consequence the operator is not too constrained to organise and optimize its production and commercialize the services. The regions do make

[42] Between 1997 and 2015, the infrastructure manager was an independent firm called RFF. Since the beginning of 2015, the railway sector has vertically re-integrated in France following a legislative bill on the sector past in 2014. The infrastructure manager is now a subsidiary of the incumbent under the name SNCF Réseau, but has to be managed independently.

[43] The subsidiaryy of the monopoly in charge of maintaining train stations: Gares et Connexions was created in 2010.

use of their prerogative to set the price of travel cards for daily commuters[44]. For the rest, the regions delegate to the operator the aspects of communication to customers and marketing activities. The transport authorities do not use the possibility they have to order train paths, leaving it to the operator. This enables the state incumbent to design a regional service consistent with its national trains and, in particular, to create a national network of inter-connecting trains to service smaller urban areas. Therefore the transport authorities have not developed a strong operational expertise, considering it is a task for the operator. And this distribution of tasks leaves some latitude to the operator to optimize its production.

3.2.3 Design of the regulatory mechanism

Regarding the financial transfers between the two contractors the design is the same but some parameters of the contracts are adapted to each regions which creates heterogeneity as we will describe later on. In all contracts the regional operator receives each year a compensation from the region which is equal to the difference between operating costs[45] and revenues from fares. On average, the revenues from fares cover only 27% of the costs in our database. Thus, and as already mentioned, the financing of rail transport services depends heavily on the compensation payments made by the regional authorities.

The payments scheme was set the same way across the regions and, at first glance, it has the flavour of a fixed-price contract. The transfer payments (T) by the regions can indeed be described as follows:

$$T = C1 + C2 - R$$

[44] The basic level of fares is set at a national level and approved by the ministry of transport.

[45] Investments are treated on a separated account. The only capital expenditures are the rolling stocks when they have not been fully paid by the regional authorities.

- *C1* are the controllable costs subject to a cap set at the beginning of the regulation period with its indexation formula defined for the duration of the contract;
- *C2* are the non-controllable costs billed *ad valorem*;
- *R* are the commercial revenues which are a function of a yearly objective and the realized revenues.

The regime applied to costs is the same across contracts where two types of costs are explicitly defined: the non-controllable costs (**C2**) and the controllable costs(**C1**). The allocation of budget items between the two types of costs is the same across the contracts. The non-controllable costs are transferred each year *ad valorem* as pass through costs to the region by the operator because the operator can not control the evolution of these operating expenditures. They encompass all the non-controllable costs such as infrastructure charges, taxes and amortisation of capital. All those items can be easily verified by the transport authorities.

The controllable costs, on the other hand, are under a cap. They include the remaining of operating expenditure, in particular, the operation and maintenance of rolling stocks, the energy consumption, expenditures for operating the train stations and commercializing the service. They represent on average 76% of the operating expenditures. The evolution of the cap is set at the beginning of the contract from one year to another according to a composite index formula based on the evolution of cost drivers in the sector. At last, it is worth noting that a provision in the contracts sets the remuneration of the operator as a percentage of the controllable costs. This remuneration is intended to cover the industrial risk resulting from the cap on controllable costs. Notwithstanding this last provision, the contract set in place could be categorized as a fixed price contract. Yet, the fact that the operator is explicitly remunerated to bear the risk on production mitigates the actual risk shifting

and dulls the incentives for the operator.

A key feature of the controllable costs is that most items are derived from internal billing and, as stated by the *Cour des Comptes*[46], *caution [which] should be used when interpreting the accounts by branch, constructed for a large part on the base of billing of service providers internal to the SNCF and who are not certified by an external auditor.* Besides the issue of internal billing, the other elements, that can be pointed out, are that the reporting lacks data on the working units used, making it complicated to assess the cost structure of the operator and that the regional authorities can not monitor when those working units are really affected to a task specific to the contract. In a nutshell there is a strong level of asymmetric information on costs. Especially, the costs that are controllable by the operator and are subject to a price cap in the PSO contracts, cannot be observed by the public buyer.

The denomination in the contract for the two categories of costs is therefore hazardous. The first category of costs is called controllable because they can be controlled by the train operator, but the feature we will use is that they are non-verifiable by the transport authority. It is the opposite for the second category. Although they cannot be controlled by the train operator, they can be easily verified by the transport authority. Figure 3.2 sums up the financial characteristics of the contracts.

In this context, setting the initial cost base requires a strong audit capacity on behalf of the regions. In the recent years, the regions have often stressed issues related to asymmetric information and, as a reaction, audits have been ordered to better understand the billing of the operator[47] and to better assess its performance, in the absence of competitive pressure and benchmarking

[46] Court of Auditors which is a quasi-judicial body of the French government charged with conducting financial audits of most public institutions

[47] The *Lorraine* region ordered for instance an audit in June 2013.

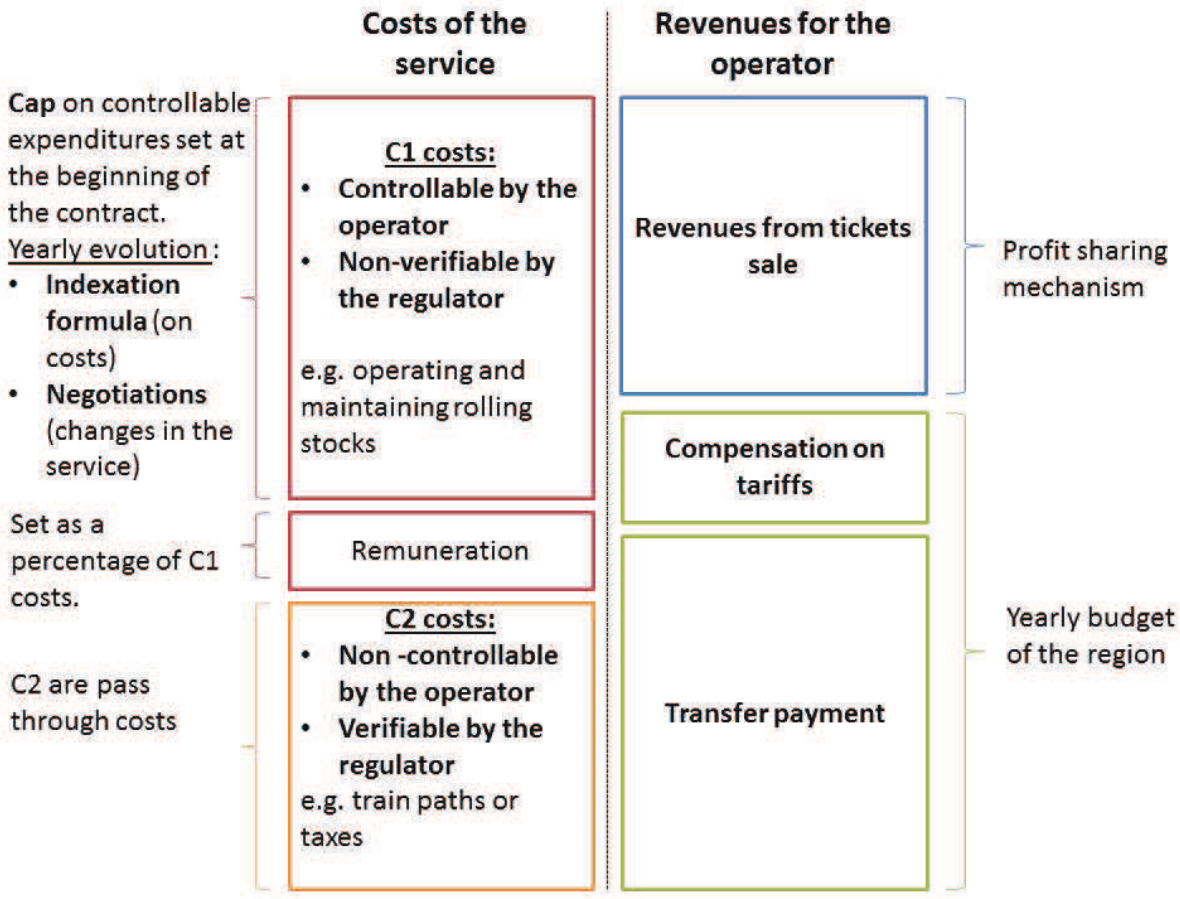

Figure 3.2: Financial design of the PSO contract

tools that could enlighten them on the subject. Regional transport authorities have a limited access to information on costs and it is uneasy for them to verify how far the costs correspond to any economic reality for the operator. Consequently, they might have a limited bargaining power to challenge the costs when they set the initial amount of expenditures included in the cap.

Furthermore, the development of new rail services does not take place in-between contracts but also after the contracts are signed. This means that the transport authorities and the operator have to meet again to negotiate on the terms of a new service according to proceedings defined in the contracts. Therefore negotiations on the cost of the service are not a one-shot deal but may arise frequently.

On the revenue side, the commercial revenues objectives are usually negotiated *ex ante* on a yearly basis and therefore not set in stone for the duration of the contracts. Second, in case the yearly objectives are not met, the difference is distributed between both contractors using a risk sharing mechanism specified in the contract. Depending on the percentage of deviation from the revenue target, the contracts specify which contractor has to bear the difference. The modalities of the risk sharing mechanism vary from one region to another in particular with respect to the threshold that leads to a renegotiation of the initial objectives. In the end, it would seem that there is a limited commitment on commercial revenues objectives over the whole regulation period and yearly deviations are subject to a risk sharing mechanism. As a consequence, our opinion is that operator bears only a small part of the commercial risk.

3.3 Theoretical framework and testable propositions

3.3.1 Expertise of the public buyer

In public procurement, many articles have insisted on the importance of the design of contracts on the efficiency of the operator in case of asymmetric information. In the public procurement literature it is often assumed, as in Laffont and Tirole [1986], that the cost of the operator depends on its innate efficiency θ and the effort of cost reduction e such that the cost of the operator is $C = \theta - e$. Fixed-price contracts are deemed to create more incentives to reduce costs (that is to increase the parameter e) because, on the contrary to a cost plus scheme, operator's profits depend on its ability to lower its level of costs during the duration of the contracts as described in Cabral and Riordan [1989]. On the other hand, those contracts are supposed to leave a rent to the operator whose innate efficiency θ is higher. A regulator might be reluctant to do so given the cost of public funds. From an empirical stand point, Gagnepain and Ivaldi [2002], Piacenza [2006] and Roy and Yvrande-Billon [2007] corroborate the predictions on the impact of high powered incentive schemes on technical efficiency in the urban public transport sector.

More generally, in order to optimize its agency relation with the operator, a principal has, according to Jensen and Meckling [1976], two tools at its disposal which it can juggle with, which are the contractual scheme, to align its interest with those of the firm, and a monitoring function. These two elements are considered substitutes and any control that is set in place by the principal has an agency cost. The economic literature has often overlooked the issue of monitoring at the benefit of minimizing a residual loss. The literature on public procurement has also stressed, as in Saussier and Tirole [2015], that the public buyer often "lacks of data for monitoring changes in the public procurement system and analysing its performance, meaning that the *ex ante* and

ex post control of contracts is consequently limited." This statement is in line with Brown and Potoski [2003] who suggest that a government needs both implementation capacities, that is to say the ability to negotiate a contract and also the capacity to evaluate the performance of the operator. These two abilities define the expertise of the public buyer which, arguably, could be limited (Saussier and Tirole [2015]).

3.3.2 Monitoring expertise

One of the main differences with the framework described in Laffont and Tirole [1986] is that we have to relax the assumption they make that accounting data can be used to regulate. Maggi and Rodriguez-Clare [1995] point out that there is a systematic distortion of information transmitted to the principal in an agency relationship, with the regulated firm spending considerable resources to convince the government that their costs are higher than they really are. The operator is able to use the accounts in a strategic way such as cost padding. And as pointed out by Rogerson [1992], even in the presence of a revenue cap mechanism, cost padding can be a beneficial strategy for the regulated firm when there are important overhead costs that have to be assigned between various activities even if it does not increase the profits of the regulated activity.

As explained by Bougheas and Worrall [2012] cost padding falls in between the two hypothesis made on costs in agency theory. Contrary to Baron and Myerson [1982] the regulator is able to observe part of the costs. But the costs cannot be perfectly observed as in Laffont and Tirole [1986]. When a firm has the possibility to do cost padding, a way to represent the cost of the operator is the following: $C = \theta - e_1 + e_2$ where the two actions that can be undertaken *ex post* by the manager are e_1, a cost reduction effort, and e_2, cost padding. Those two actions cannot be verified by the regulator. In the presence of

cost padding, it is more difficult to extract a rent as suggested by Laffont and Tirole [1992] and therefore can be used by the firm when the contract is signed.

Our interpretation of the contracts is that the heterogeneity in the ability to set the costs by the regions can be captured using the ratio of controllable costs over the total costs given that the operators have access to the same technology and inputs. This ratio ranges from 67% to 85% (with an average of 76%).

The informational gap should not just translate in a one shot deviation at the beginning of the contract but also during the contract when renegotiations take place, as the cap on controllable costs can be subject to renegotiations during the contract. As mentioned, those negotiations will occur, in particular, when there is a need to change the nature of the service, whether the situation is lasting (the transport authority may for example decide to open a new route) or temporarily (this is the case when a route has to be closed temporarily after heavy maintenance was decided on the network or when a train is cancelled due to operational difficulties). Some contractual clauses also allow renegotiating in the situation where the economic equilibrium of the contract changes[48]. In the end, we assume that the yearly evolution of the controllable costs does not follow strictly the indexation formula set at the beginning of the contract but is rather the result of a negotiation between the two parties.

Proposition 1:
The larger the share of costs, at the beginning of the contract, that are non verifiable by the transport authority, the less cost efficient the contract is for the transport authority, all other things being equal.

[48] Although the notion of economic equilibrium is not precisely specified in the contracts.

3.3.3 Implementing an incentive contract under limited expertise

The lack of auditing and monitoring abilities of behalf of the regional transport authority has an impact on the possible set of contracts it can use in order to align interest. In this regard, Glachant et al. [2013] define the notion of regulatory alignment. They insist on the fact that the choice of the regulatory scheme should be in line with the expertise of the regulator. In particular, a revenue cap is considered more complex than a cost plus contract to put in place when the expertise of the principal is too limited. The authority must have the capacity to implement the contract that is to negotiate the contract (Brown and Potoski [2003]). As stressed by Saussier and Tirole [2015], this is important "*in a context of ubiquitous asymmetries of information and in which contractual details are significant.*" In our setting we have identified adverse effects which may arise due to the remuneration percentage and affect the dynamics of performance during the contract.

Regarding the remuneration percentage there are two potential effects. It dilutes the risk for the operator, dulling the incentives to reduce cost and creates an incentive for cost padding. As shown by Cabral and Riordan [1989] the incentives for technical efficiency in a fixed-price contract stem from the fact that the operator becomes the residual claimant of a cost reduction. But in the case we study, a remuneration for the operator is set as percentage of the controllable costs. It ranges from 0.8% to 3.7% of the overall costs in the contract. A strategy for the operator may therefore be to increase its costs when the contract is negotiated, if banking on the cost plus dimension of the contract is deemed more profitable than the net benefits of reducing ones costs. The payment scheme would then lead to an Averch and Johnson [1962] effect on the controllable costs. The clause that sets the remuneration as a percentage of the controllable, yet non observable costs, is going act as a catalyst for any cost padding phenomenon.

Proposition 2:
The larger the percentage of remuneration on costs at the beginning of the contract, the less cost efficient the contract is for the transport authority, all other things being equal.

Regarding the dynamic of cost reduction incentives, since we are under a fixed price contract, changes in the cost that are billed to the regional authorities take place either because of the indexation formula or because of renegotiations taking place during the contract. Given the literature on the dynamics of cost reduction under a price cap, and if the operator does in fact produce an effort to reduce its costs, it should be at the beginning of the contract. This intuition is given by Joskow [2005]: "*A dollar of cost savings in year 1 is worth much more to the firm than a dollar of cost savings in year 5*". Firms have an incentive to make a greater effort of cost reduction at the beginning of the contract if they can benefit from a rent over a longer period of time, before the regions appropriate part of the rent they generated. In the case of urban public transport Gautier and Yvrande-Billon [2013] showed empirically that the profit flow is a source of incentive for reducing costs at the beginning of the contract. Should this be true in our case, this would mean less renegotiations at the start of the contract and therefore a contract that is more efficient. Because as pointed out it Guasch et al. [2008] the incentives are not dulled, and because renegotiations during the contract leads to higher transaction costs to adapt the contract under a fixed price mechanism as shown by Bajari and Tadelis [2001].

A second phenomenon when studying the dynamics of cost reduction in the literature on procurement is the effect the perspective of seeing the contract being renewed has. Players anticipate that they will be evaluated mainly on their recent performance due to the fact that it is considered more representative (see, for example, Fudenberg and Tirole [1995]). Fudenberg and Tirole

[1995] highlight in particular that there may be room for self-sabotage when the incentive to smooth the performance presented to shareholders is important. In the case of public procurement Iossa and Rey [2014] and Affuso and Newbery [2000] also expect that renewal decision are mainly based on recent performance. This creates an incentive for the firm to make a greater effort towards the end of the contract and not during the contract since past performance is of little relevance when negotiating the contract. Gautier and Yvrande-Billon [2013] find that an operator in place increases its technical efficiency at the end of the contract to maximize the odds of being re-conducted in the case of public urban transport in France.

In the case of regional railway transport in France, no such incentives exist since the probability of not being renewed is null and there is no possibility to operate the service in house for the transport authority. Therefore there is no incentive to increase the efficiency of the contract at the end of the contract. On the contrary, we might observe a decline in efficiency toward the end of the contract.

Given the negotiation process we have described earlier in the French regional railway transport, the operator can anticipate as well that the transport authority will use recent performance as a basis to negotiate the upcoming contract. Since the contracts are automatically renewed, the operator can adopt a strategy where it exhibits a higher degree of inefficiency at the end of the contracts. This strategy would lead potentially to a greater rent during the next contract, overcoming the loss incurred to adopt such a strategy (should the operator not be able to fully renegotiate its cost increase). In other words, the fact that the market is not contestable creates a potential ratchet effect on the performance of the operator as described by Weitzman [1980]. This ratchet effect is consistent with the findings of Laffont and Tirole [1988] stating that agents can only be given a low-powered incentive scheme in a repeated relationship without commitment. Those arguments lead us to our third the-

oretical proposition.

Proposition 3:
The efficiency of the contract is decreasing in time: the closer the end of the contract, the less cost efficient the contract is for the transport authority, all other things being equal.

3.4 Empirical model

3.4.1 Assessing the relative performance of the contracts

To test our propositions we need to have a measure of the efficiency of the contract and assess the effect of contractual variables over the efficiency. It stems from the previous section that if our propositions are verified and the contractual variables impact cost efficiency, then the expertise of the public buyer does to. Our empirical strategy is to rely on a measure of relative performance and more particularly on a stochastic frontier analysis (SFA), as this methodology fits our needs, this point being developed later on.

A prerequisite to assessing the relative performance is to set the objective function of the authority, as laid out in Cherchye et al. [2015]. In our view, a general objective for regional authorities should be to minimize the annual payment it transfers to the train operator for a given level of service (routes, frequencies) defined in the contract. A similar assumption is done for instance in Dalen and Gómez-Lobo [2003] or Wheat and Smith [2015] in the case of transport services procurement. It is to be noted that an underlying associated assumption of this objective is the exogeneity of the output. An output such as train-kilometres can however be considered exogenous given that the service is characterized by a public service obligation (in this sector, it has been defined as "*a requirement defined or determined by a competent authority in order to*

ensure public passenger transport services in the general interest" according to the European legislation[49]). The cost minimizing objective is rather straightforward given the hard budget constraint that public administrations usually face.

To test our propositions, we rely on a stochastic frontier analysis[50] (SFA). Since the seminal paper by Farell [1957], this method has been enriched on numerous occasions (Kumbhakar and Lovell [2001]). When compared to non-parametric methods, one of the key characteristics of SFA lies in the use of (at least) two error terms. The first one captures the statistical noise that may arise from data collection or production shocks. Contrary to DEA methods, not all noise is considered as (in)efficiency. The second error term is a univariate and is often used to measure technical efficiency as described in Aigner et al. [1977] and Meeusen and Van Den Broeck [1977]. Yet this univariate term can also be used reporting errors that can be assumed to be one sided. For instance Chaudhuri et al. [2015] use SFA to estimate unreported violent crimes in India. In the case studied, this term might not only capture the technical efficiency of the operator, but also strategic increase in costs during the contract, due for instance to cost padding.

This method has, in particular, been used to measure the technical efficiency in the case of railways. Most studies have focused on assessing the performance of the overall railway sector [51], and less on the performance of solely the downstream market as in Lévêque [2004] and Wheat and Smith [2015]. It is also to be noted that stochastic frontier analysis has also been frequently

[49] Article 2 of Regulation (EC) n° 1370/2007 of the European Parliament and of the Council of 23 October 2007 on public passenger transport services by rail and by road and repealing Council Regulations (EEC) n° 1191/69 and 1107/70.

[50] Non-parametric modelling such as *Data Envelopment Analysis* (DEA) is also commonly used in the literature to measure performance. A survey on the merits and disadvantage of the two methods can be found in Coelli et al. [2005] and Coelli et al. [2003]. In our case, the SFA methodology was more relevant to test our theoretical propositions as we will develop later on.

[51] the management of the infrastructure and the operation of transport services. A survey can be found in Mizutani et al. [2009].

used to measure the performance of the urban transport services (Piacenza [2006], Gautier and Yvrande-Billon [2013]).

An usual presentation of a stochastic cost frontier model is as follows:

$$C_{it} = f(Y_{it}; \beta).Ef_{it} \tag{3.1}$$

Where C_{it}, Y_{it} and β stand respectively for the cost level, the vector of outputs of firm i $(i = 1, 2, ..., I)$ at the period t and the vector of parameters to be estimated. The term Ef_{it} represents the efficiency of the operator i at date t, that is to say the ratio between the minimum level of cost that could be obtained for given outputs and inputs' prices.
As a consequence:

$$C_{it} \prec f(Y_{it}; \beta) \text{ and } Ef_{it} \prec 1.$$

Using SFA we can test the impact of contractual variable on the efficiency of the contract. Our objective is to test whether the efficiency differences can be attributed to the contractual scheme and not only to varying operating conditions. In other words, an objective of the paper is to analyse how far any departure from the objective of the transport authority (minimize the transfer payment given a certain level of output) can be explained by strategic behaviours of the operator, focusing on the ones that stem from the lack of monitoring ability of the regions and the design of the contract we have identified in the previous section.

3.4.2 Specification of the model

One of the peculiarities of parametric models rely in the need to specify a functional form which necessarily imposes constraints on estimation results. In this paper, we choose the log linear Cobb-Douglas cost functional form. A

TransLog cost function could be more appropriate as it is a more flexible form: the TransLog function imposes indeed less restrictions on the substitutability of inputs. One of the perks of the Translog functional form is that it gives information on the return to scale as discussed in Wheat and Smith [2015] for downstream railway services. Yet, such a functional form would require a greater number of degrees of freedom which we cannot afford given the limited size of our sample (see Urdanoz and Vibes [2013] for a discussion on this topic). The Cobb-Douglas function has also already been used in the railway sector such as Farsi et al. [2005] in Switzerland, Mizutani et al. [2009] in Japan and Lévêque [2004] in France.

Rewriting equation (1), the cost frontier to be estimated can be written as follows, in the case of a Cobb-Douglas functional form:

$$\ln C_{it} = \beta_0 + \beta_1 \ln Y_{it} + \sum_{n=1}^{N} \beta_n \ln Yc_{nit} + \sum_{n=1}^{M} \beta_n \ln X_{mit} + \epsilon_{it}, \tag{3.2}$$

C_{it} is the cost variable, Y_{it} the output, Yc_{it} a vector of N output characteristics, X_{it} a vector M of environmental variables and ϵ_{it} the error term. The introduction of a vector on output characteristics is suggested in Mizutani and Uranishi [2013] when assessing the relative performance of railway sectors in OECD countries. It is to be noted that the variables we consider as output characteristics may be treated as control variables in other empirical studies (see for example Smith and Wheat [2012]).

To test our theoretical predictions, we use an original dataset on the 20 French regions between 2009 and 2012 which is 7 years after they became transport organizing authorities (2002). It is worth noting that during the period, no institutional or organisational reform took place in the railway sector, making the comparison more comfortable. During the time period we study, each region was at least running its second PSO contract with the incumbent. The descriptive statistics of the variables we use can be found in table 1.

The cost variable. The variable we use is the net contribution of the regions, in other words the cost billed each year minus the commercial revenues of the service. We consider the net contribution of regions to be relevant for our analysis since it is the variable the regions base their decision on and it is consistent with PSO contracts where the objective of the regulator is to provide a certain quantity of service to users under a budget constraint. It is worth noting that our variable depends on the commercial revenues generated by the service but, as mentioned, the pricing strategy is outside of the control of regional operators. Therefore, variations of commercial revenues are mostly dependent on the ability of the regional operator to optimize and market the service.

In the case we study we only observe the costs that are billed to the transport authorities by the operator on a yearly basis. We can not ascertain the exact profits of the operator and therefore its cost efficiency or its evolution. What we capture in our univariate error term is part the cost efficiency and part the univariate strategic display of performance of the operator which we have described in our theoretical framework. In both cases, we argue that we capture the overall efficiency of the contract as seen by a cost minimizing public buyer and estimate a net cost function.

The output. The indicator we use to quantify the production of the regional companies of the incumbent is the vehicle-kilometres ($TRKM$). As pointed out in Figure 3.3, there is a large heterogeneity in the size of the contracts. This variable is communally used in empirical studies focusing on the public transport sector (in railways:Lévêque [2004], Farsi et al. [2005] and in urban transport: Berechman [1993], Kerstens [1996], Dalen and Gómez-Lobo [2002], Piacenza [2006], Roy and Yvrande-Billon [2007]). One of the advantages of this supply-oriented variable is that, for a large part, it can be considered as exogenous to the extent that the expected service level is defined *ex ante* in the contractual agreement.

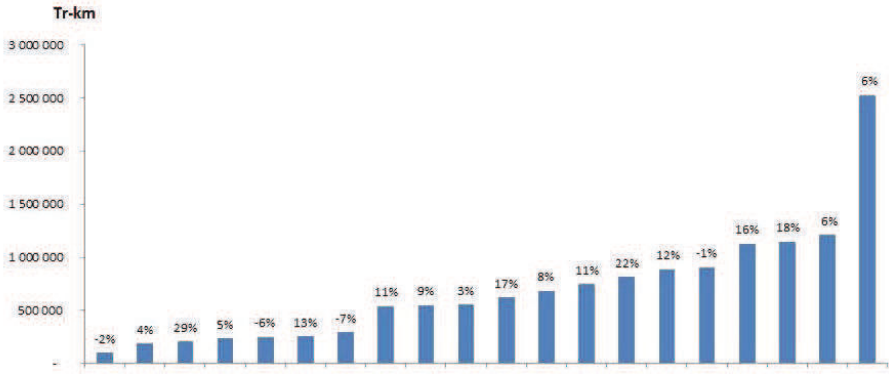

Figure 3.3: Train-kilometers in 2012 and the evolution since 2009 in percentage

Output characteristics. A drawback associated with supply-oriented indicators relies in the fact that it does not take into account the number of passengers using the service. Ignoring demand might lead to misleading conclusions where an operator is seen technically efficient although its trains are empty[52]. To tackle this issue and, as it has been done in recent studies on rail downstream services (Wheat and Smith [2015]), we introduced a demand-related output variable ($LOAD$) giving the average load per train. Since our dependent variable is the compensation paid by the region each year, net of the commercial revenues, this variable allows us to control for the demand and not only the supply of public services. We assume that the average load per train has a positive influence on the costs.

Inputs prices. Inputs prices on capital, labour and energy are not included in the cost function for the reason that corresponding information was not available. However, as we are analysing the regional directions of a unique firm, one can reasonably assume that they should have access to the same inputs at similar prices. Energy purchases are, for example, made at the corporate level, and there is also a common salary grid. Other activities, such as maintenance of rolling stocks and traction for the trains are also performed by an entity of

[52] On the contrary if the trains are empty, it would reduce the commercial revenue, thus making the operator look more inefficient, depending whether or not ticket prices are above the marginal cost of an extra passenger

the firm operating at a national level and then billed to the local divisions. The absence of price data should not impair the results. On the contrary, taking into account prices derived from internal billing could neutralize the effect of cost padding on efficiency. Indeed, it is assumed when estimating a cost function that prices are competitive (Coelli et al. [2005]) when they are taken into account. On the contrary we assume that they are distorted due to the monopoly position.

Control variables. To control for the heterogeneity of exogenous production constraints, several variables have been introduced. The objective of these variables is to proxy for the complexity of networks and the geographical distribution of the stations, which may alter the operating conditions of the operator and impacts on the operating expenditures.

The first variable *NODE* has been designed to proxy the presence of nodes in the area and the complexity of the network. To our knowledge we are the first to use such a variable. It was designed to capture the existence of major rail junctions in the regional networks. In order to construct this variable, we added the number of tracks entering the cities of more than 200 000 inhabitants, multiplied by their UIC coefficient. The UIC coefficient is an international classification for railway lines where the coefficient varies from 1 to 9 and is defined depending on the traffic on the line. The variable was designed such that the higher the value of the variable, the more complex the infrastructure is. We expect this variable to have a negative influence on cost efficiency.

The two other variables *STATIONperROUTE* and *STATIONperAREA* intend to capture how the network of stations is designed. *STATIONperROUTE* is the number of stations divided by the length of the rail network in each region. Our intuition is that a denser network should increase service complexity and decrease commercial speed thus having a negative impact on efficiency. *STA-*

TIONperAREA is the number of station divided by the surface area of the region. We consider this variable to be a proxy for the density of urban territories and, as for *STATIONperROUTE*, it is expected to have a negative effect on efficiency.

At last, we introduced a variable giving the average number of stops at a railway station (*ASTOP*) on a route in the region. As for the variable *STATIONperROUTE* and *STATIONperAREA*, we expect that an increase of the average number of stops has a negative impact on technical efficiency since it will be associated with more station charges and a higher energy consumption due to more frequent accelerations[53].

The descriptive statistics of the variables used for estimating the cost frontier can be found in Table 3.1.

Table 3.1: Descriptive statistics

Variables	Average	Median	Min	Max	Std. dev.
BUDTER: Compensation (M €) paid to the operator	131.2	121.5	42	417	78.7
TRKM: Train-km per year	8 534 445	7 817 962	2 803 419	28 300 000	5 128 461
LOAD: Average load per train-km	70.17	68.91	27.73	120.73	19.79
ASTOP: Average number of stops per route	8.10	8.89	6.81	11.57	1.34
NODE: Complexity of the network around the large cities	34.80	27.85	1	106	31.29
STATIONperAREA: Number of stations per surface area	0.59	0.42	0.30	1.95	0.44
STATIONperROUTE: Number of stations per length of the network	0.12	0.11	0.06	0.25	0.04

[53] Experts pointed out to us that the accelerations are a non-negligible source of energy consumptions. In some cases, it justifies a specific formation for the drivers to reduce the operating expenditures.

3.4.3 Specification of the error terms and propositions

Our net cost function can be rewritten as (**Model 1**) :

$$\ln C_{it} = \beta_0 + \beta_1 \ln TRKM_{it} + \beta_2 \ln LOAD_{it} + \beta_3 \ln ASTOP_{it} + \beta_4 \ln NODE_{it} + \beta_5 \ln STATIONperAREA_{it} + \beta_6 \ln STATIONperROUTE_{it} + \epsilon_{it} \quad (3.3)$$

where $\epsilon_{it} = V_{it} - U_{it}$. V_{it} are assumed to be *i.i.d.* $N(0, \sigma_V^2)$ random errors, independently distributed of U_{it} capturing the effects of measurement errors, statistical noise and random exogenous chocks. U_{it} capture univariate deviations to the frontier (inefficiency) and are assumed to be independently distributed as truncated normal distribution. This decomposition in two terms was defined by Aigner et al. [1977] and Meeusen and Van Den Broeck [1977]. To estimate the model 1, we use the specification defined by Battese and Coelli [1992] to take into account the panel structure of our data. One of the perks of SFA is that we can check if a variable has an impact on the efficiency of the operator and therefore test our propositions.

A first methodology would consist in estimating the parameters of the stochastic frontier and the efficiency scores of each firm (*i.e.* the distance to the frontier) as a first step. In a second step, these results would be regressed over the variables explaining the inefficiency of operators using, for instance, an OLS regression. This two stages methodology has been previously used in empirical studies in the public urban transport sector to estimate the inefficiency of operators (Jorgensen et al. [1997]). However, as pointed out by Dalen and Gómez-Lobo [2003] for example, this methodology exhibits an incoherence as the efficiency score are assumed to be independently and identically distributed in the first stage of the regression whereas, in the second stage, those score are assumed to be dependant to firm specific variables, hence they cannot be *i.i.d.*

For that reason, we rely on the model developed by Battese and Coelli [1995] where both the parameters of the cost frontier and the impact of firm specific

variables over the efficiency score are estimated simultaneously. In this model, the efficiency term U_{it} has a truncated normal distribution $N(M_{it}, \sigma_U^2)$ such that $M_{it} = Z_{it}\delta$. Z_{it} is a vector of variables that may have an impact on the efficiency of firms and δ the vector of parameters to be estimated. V_{it} are assumed to have an $i.i.d$ normal distribution $N(0, \sigma_v^2)$ and distributed independently to the technical efficiency terms U_{it}. The vector of parameters δ and β are estimated using a maximum likelihood method as well as the associated parameters $\sigma^2 = \sigma_V^2 + \sigma_U^2$ et $\gamma = \sigma_U^2/\sigma^2$. The parameter γ which is included between 0 and 1 gives the influence of technical efficiency on the overall variance of ϵ_i. The closer γ is to one, the bigger the effect of U_{it} is.

In order to test our theoretical predictions, we introduce four so-called *Z-variables*, namely *RATIO, REMU, LEFT* and *TWO* and measure the impact they have on the distribution of the efficiency error term U_{it}.

RATIO represents the ratio of controllable costs over the overall costs defined at the beginning of the contracts. In other words, this variable represents the share of expenditure subject to a fix flat rate over the overall expenditure that is billed by the operator on a yearly basis. As it has been pointed out earlier, we use this variable as a proxy for the monitoring expertise of the transport authority. According to our **first proposition**, we expect that the higher the variable, the lower the relative efficiency[54].

REMU is obtained by multiplying the variable *RATIO* by the percentage of remuneration on controllable costs defined *ex ante* in the contract. This variable therefore gives the percentage of remuneration awarded to the operator over the overall costs. Consistently with our **second prediction**, we expect a high value of this variable to have a negative effect on efficiency, should it indeed dull the incentive of a fixed price contract.

[54] This implies a positive coefficient as it shifts the distribution of the efficiency term to the right-hand side.

LEFT represents the remaining years in the contract over the total length of the contract. According to our **third proposition**, efficiency is expected to decrease during the contract so the sign of the coefficient should be negative. Despite the first contracts having been signed in 2002 in all the regions[55], the contracts are at different point in time over the period we study (2009 - 2012) as pointed out in Figure 3.4.

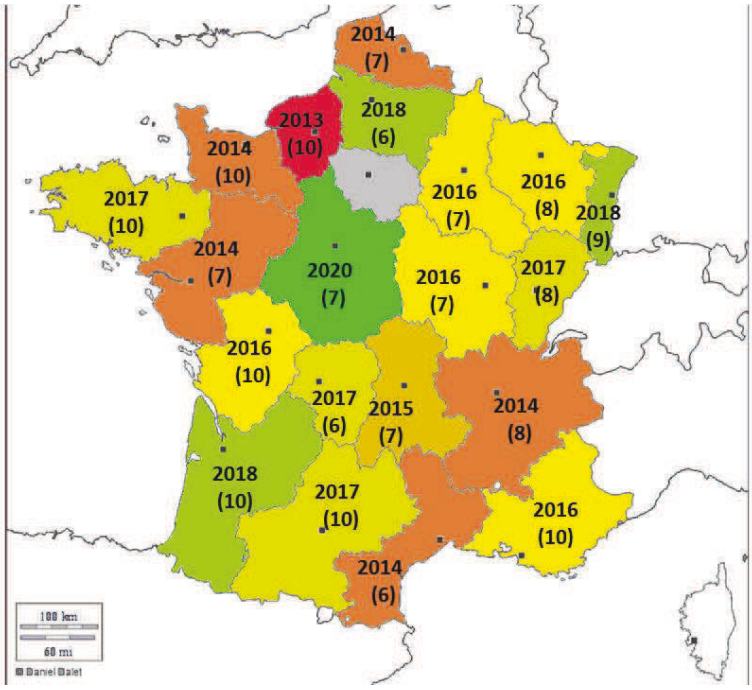

Figure 3.4: Contracts termination in our database (year 2012)

TWO is a dummy variable which is equal to 1 if we are in the two remaining years of the contract and zero otherwise. We introduced this variable to capture the existence of a ratchet effect at the end of the contract consistent with our **third proposition**. Should the operator deteriorate its performance at the end of the contract, before it enters the negotiation for a new contract, the coefficient for this variable will be positive.

[55] Six regions took part of an experimentation process which led to a first set of contracts being signed in 1996 and 1997.

The corresponding descriptive statistics can be found in Table 3.2.

Table 3.2: Descriptive statistics of variables on the contracts

Variables	Average	Median	Min	Max	Std. dev.
RATIO: Percentage of controllable costs	76.27%	76.00%	67.10%	85.16%	5.29%
REMU: Percentage of remuneration on the overall costs	1.73%	1.70%	0.65%	2.81%	0.39%
LEFT: years left on the contract divided by the total duration	49.85%	50%	0%	100%	21.97%
TWO: Dummy variable = 1 if two years or less are remaining	0.25	0	0	1	0.44

As stated earlier, we rely on the methodology developed by Battese and Coelli [1995]: V_{it} is a stochastic term with and *i.i.d.* distribution $N(0, \sigma_V^2)$,independent from U_{it}. U_{it} is a random variable associated to the efficiency and we assumed its distribution to follow a truncated normal form $N(\delta Z, \sigma_V^2)$ such that:

$$U_{it} = \delta_1 Z_{1it} + \delta_2 Z_{2it} + W_{it}$$

where W_{it} is a random variable with a truncated normal distribution of zero expectation and variance σ_U^2. Thus, rewriting equation (3):

$$\ln C_{it} = \beta_0 + \beta_1 \ln TRKM_{it} + \beta_2 \ln LOAD_{it} + \beta_3 \ln ASTOP_{it} + \beta_4 \ln NODE_{it} + \beta_5 \ln STATIONperAREA_{it} + \beta_6 \ln STATIONperROUTE_{it} + V_{it} - U_{it} \quad (3.4)$$

To test our prediction, we estimate three models with different specifications of the efficiency term. More precisely we test two by two our contractual variables because respectively *RATIO* and *REMU* on the one hand, and *LEFT* and *TWO* on the other hand are correlated[56].

Model 2:

$$U_{it} = \delta_1 RATIO_{it} + \delta_2 LEFT_{it} + W_{it} \quad (3.5)$$

[56] A fourth combination was possible but the results are not significant so we do not report them.

Model 3:

$$U_{it} = \delta_1 REMU_{it} + \delta_2 LEFT_{it} + W_{it} \tag{3.6}$$

Model 4:

$$U_{it} = \delta_1 REMU_{it} + \delta_2 TWO_{it} + W_{it} \tag{3.7}$$

3.4.4 Empirical results

We estimate the cost function shown in equation (4) with the specifications (5), (6) and (7) using the maximum likelihood. An usual specification test has been realised to check the robustness of the estimations. This test compares the constrained OLS model where $\gamma = \sigma_u^2 = 0$ with the models we have estimated. The test statistics is $LR = -2[\ln L_0 - \ln L_1]$ where L_0 is the log-likelihood of the constrained model and L_1 of the unconstrained model. As indicated in Table 3.3, we can reject H_0 and, therefore, conclude on the presence of inefficiency in the models we have estimated.

Table 3.3: LR-Test results

$H_0 : \gamma = \sigma_u^2 = 0$

	Log-likelihood	LR-Stat	$\Pr(> \chi^2)$
Model 1	71.83	102.24	$1.31 * 10^{-13}$***
Model 2	71.83	78.10	$2.48 * 10^{-3}$**
Model 3	71.83	76.09	0.017*
Model 4	71.83	77.81	$3.29 * 10^{-3}$**

Significance: $*p \prec 0.10$; $**p \prec 0.05$; $***p \prec 0.01$

Our first empirical result indicates that there is significant efficiency variations from one region to another (see figure 3.5). When normalizing the best score such that, each year, the most efficient region has a score equal to 1, we estimate an additional cost of 394 M € per year for the transport authorities. This amount represents 15% of the annual contributions paid by the regions over

the time period we consider (2009-2012)[57]. This first empirical result helps illustrate the magnitude of efficiency deviations in our database. What our estimates reveal is that, despite the fact that regional transport services are operated by divisions of a unique firm, it appears that their relative performances differ significantly. To our opinion, this result may contribute to the idea that regional authorities could benefit from the development or regulatory tools based on the measurement of relative efficiency.

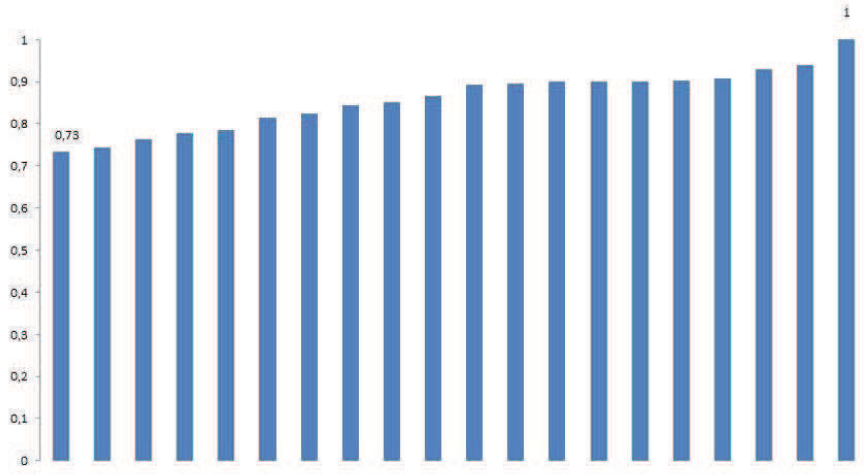

Figure 3.5: Average efficiency scores of model 1(2009-2012)

The results of the estimations can be found in Table 3.4. Model 1 is the cost frontier as defined by Aigner et al. [1977] where we do not consider a vector Z_{it} of variables that may have an impact on cost efficiency. It is worth noting that, although significant, the coefficient of $TRKM$ is lower than 1. This might indicate positive return to scale. It might also be due to the fact that the variable $LOAD$ captures also the production of the operator and that the two coefficients have to be interpreted together. The coefficient of the variable $LOAD$ is also positive. All things equal, train with more passengers on board generate more costs.

[57] The compensation payments for regional transport for the 20 regions all together were on average of 2623 M € per year.

As expected, the coefficient associated with the variable *ASTOP* (number

Table 3.4: Results of the regressions

Variables	Model 1	Model 2	Model 3	Model 4
Constant	-10.75***	-10.28 ***	-11.73***	-10.40***
	(0.83)	(0.48)	(0.52)	(0.43)
log(TRKM)	0.66***	0.66***	0.75***	0.66***
	(0.06)	(0.04)	(0.05)	(0.04)
log(LOAD)	0.37***	0.31***	0.36***	0.36***
	(0.08)	(0.06)	(0.06)	(0.05)
log(ASTOP)	0.91***	0.85***	0.70***	0.80***
	(0.17)	(0.11)	(0.12)	(0.07)
log(NODE)	0.05*	0.05***	0.02	0.04***
	(0.02)	(0.01)	(0.01)	(0.01)
log(STATIONperROUTE)	-0.61***	-0.54***	-0.58***	-0.64***
	(0.12)	(0.08)	(0.08)	(0.07)
log(STATIONperAREA)	0.21***	0.22***	0.16***	0.27***
	(0.06)	(0.03)	(0.03)	(0.03)
time	-0.13***	-	-	-
	(0.04)			
RATIO	-	0.36***	-	-
		(0.03)		
REMU	-	-	11.55***	14.09***
			(1.39)	(3.26)
LEFT	-	-0.22**	-	-0.26*
		(0.08)	(0.01)	(0.12)
TWO	-	-	0.06*	-
			(0.03)	
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.01*	0.01***	0.01***	0.02***
	(0.04)	(0.01)	(0.01)	(0.01)
$\gamma = \sigma_u^2/\sigma^2$	0.84 ***	1 ***	0.97***	1***
	(0.05)	(0.01)	(0.26)	(0.01)
Log likelihood	102.24	78.10	76.09	77.81
Mean efficiency	0.83	0.84	0.81	0.84

Significance : $*p \prec 0.10$; $**p \prec 0.05$; $***p \prec 0.01$
Standard deviation in parenthesis

of stops per route) is positive and significant. This result can be partly attributed to the fact that the higher the average number of stops, the higher the station charges (included into the operating costs). Also, with an increase in the average number of stops, train have to accelerate more frequently, thus increasing operating costs. In this regard, we would expect the variable *STATIONperROUTE* (number of stations divided by the length of the network) to

have a positive impact on operating costs. Yet, the coefficient is significant but negative. A possible interpretation is that a finer-meshed network facilitates rolling stocks management, thus decreasing operating costs. At last, consistently with our expectations, we obtain a positive and significant relationship between operating costs and the variables *NODE* and *STATIONperAREA*, corroborating the fact that complex transport networks are associated with higher operating costs.

On the economic drivers of inefficiency, our results indicate that both the coefficients of variables *RATIO* and *REMU* are positive and statistically significant (see models 2, 3 and 4) thus corroborating our first two propositions. In other words, our results indicate that the higher the share of controllable costs over total costs and the higher the percentage of remuneration on the total costs given to operators, the bigger the deviations from the best practice.

Besides, it appears from our estimations that the coefficient of the variable *LEFT* is negative and statistically significant, consistently with our third proposition: the closer the end of the contract, the lower the technical efficiency.

In the same vein, the coefficient associated with the variable *TWO* is positive and significant, suggesting that efficiency is on average lower during the last two years of the contract. This finding support our prediction according to which a ratchet effect exists at the end of the contracts, where the operator might decrease its performance when the negotiations over the new contract have begun.

3.5 Discussion and conclusion

3.5.1 On decentralized regulation of public contracts

Our results on the expertise of the public buyer can be also interpreted in the light of the regulatory structure that was chosen. One of the benefits of decentralization can be to increase the accountability of the public buyer as pointed out by Seabright [1996]. In the case we study, this has led to an increase in the provision of the public service by 51% between 2002 and 2012. Yet as we had pointed out earlier, the higher provision came at a price for the transport authority since the budget allocated to rail regional transport almost doubled over the period.

One interpretation is that the creation of 20 public buyers might not have been associated with the transfer of the necessary expertise and resources to negotiate the PSO contracts and in particular to oversee the financial evolution of PSO contracts. Our results also highlight that the performance of the contracts is asymmetric amongst the regions, due to the various levels of expertise. This heterogeneity might be amplified by what Laffont and Pouyet [2004] define as a competition effect between regulators when they regulate the same firm. It is the case if the action taken by the operator are substitutes across regions and fits our case given the blurred accounting we have described.

To avoid such drawbacks and increase the efficiency of public procurement, a possible regulatory tool is for the transport authorities to centralize the information. On this matter, Auriol [2000] stresses that "*a clever way to reduce information asymmetries consist in using the existing correlation between firms producing the same type of product or services. In which case we assume they are facing the same costs or at least comparable costs. Even if the regulator does not know the true value of these costs, it can use this common structure*

to overcome its informational disadvantage and be a step closer to social optimum. This is what we call yardstick competition[58]*."*

In the recent years, the regions have multiplied initiatives to make the performance of the operator more comparable. Those initiatives are emerging from the *Associations des Régions de France* which is used in the transport sector as a platform to share best practices. The output of the collaboration has been a common base of indicators which led to the publication of rankings, with an important focus on service quality. This platform also led to the production of reports on the best practices to monitor and enforce the PSO contracts. Those are positive steps in the transition from passive monitoring of the contract to a more active monitoring of the PSO operator. Yet, no formal regulation mechanism based on the performance of the operator has emerged from this collaboration as of this day.

Besides, using cost minimization as a tool to examine PSO contracts is consistent with the mechanism set out by Shleifer [1985] in its seminal article on yardstick competition. Shleifer [1985] defined a regulatory mechanism where the optimal payments made to a firm depend on cost comparison with other comparable firms under the jurisdiction of the regulator. From a theoretical point of view, the effort made by firms to reduce their costs should be greater when facing yardstick competition than with a more traditional regulation.

Despite the fact that regulation based on relative performance has been less developed compared to other sectors[59] such as the electricity (Jamasb and Pollitt [2000]) or to a lesser extent in the telecommunication sectors (Sappington

[58] *"Une facon astucieuse de réduire les asymétries informationnelles consiste à exploiter la corrélation qui existe entre des entreprises produisant le même type de bien ou de service. On suppose dans ce cas qu'elles font face au même coût ou du moins des coûts comparables. Meme si le régulateur ignore la valeur de ces coûts il lui est possible exploiter leur structure commune pour surmonter son désavantage informationnel et se rapprocher de l'optimum social. C'est ce qu'on appelle la concurrence par comparaison."*

[59] See Cherchye et al. [2015] for a survey on the use of relative performance in regulatory mechanisms.

[2002]) there are documented successful applications in public transport. For instance, as described by Mizutani et al. [2009], it has been implemented in the railway sector in Japan where there is no local competition but 14 local vertical monopolies. The regulator in charge of reviewing the fares has set up several performance measures, one of them being targeted at assessing operating costs. According to the relative performance of firms, the regulator then decides to validate or not the fares level. According to Mizutani et al. [2009] this mechanism has had very positive results since yardstick competition led to a 12.4% decrease of variable costs.

A second application to a comparable sector is described by Dalen and Gómez-Lobo [2003] in Norway for the regulation of bus services. Similarly to the French case, the responsibility for local transport had been decentralized. As pointed out by the authors, the counties adopted a cost model, applying it to all companies within a county in order to determine the annual transfers. The threat of tendering network services alone allowed some counties to put in place a subsidy cap. And this threat seems to have been enough since during the 10 years period of their study yardstick type contract helped reduce cost inefficiency faster where it was implemented and only 1.7% of the total production was subject to competitive tendering in the end.

As pointed out by Lévêque [2004] who has applied to the regional rail transport sector the theoretical and empirical tools developed for yardstick regulation, we believe also that the comparison of the regional entities can be used to assess the performance of each PSO contract. One of the reason lies in the fact that the heterogeneity of firms is not a constraint because regional contracts are awarded to the divisions of a unique company: the incumbent SNCF. Even if that company gives managerial freedom to the local operators, the productive structure should be the same, and therefore there would be homogeneity amongst decision making units which is a prior condition to the application of yardstick competition according to Sobel [1999]. Still, a challenge would be to

develop a collusion free mechanism to regulate a single firm.

3.5.2 Conclusion

In this paper we estimated a cost frontier to measure the technical efficiency of the 20 regional operators in regional railway transport. Our results indicate that there are significant efficiency differences between the regional local operators as the estimated efficiency scores ranges from 0.73 to 1, despite the contracts being operated by a single firm. Overall, according to our results, the cost could be reduced by 15% if all the other 19 regional operators adopted the best practice of the most efficient one. Part of the heterogeneity in costs we observe can be attributed to the network characteristics we identified, such as the spatial distribution of rail stations and the complexity of the rail networks. Our results also highlight the relationship between contractual scheme, regulatory commitment and cost efficiency.

In particular, the contractual design does not produce the expected incentives as the share of cost under a revenue cap is found to have a negative impact on efficiency. We relate this result to the lack of monitoring tools at the disposable of transport authorities and to the design of the contract itself which also remunerates the operator as a percentage of the controllable costs under a cap. Furthermore, and consistently with our predictions, we found that efficiency decreases over the contractual life. According to our predictions, this dynamic effect may result from a ratchet effect towards the end of the contract since the market is non-contestable.

A first policy recommendation could be derived from our results. In our view, an increase of the expertise capacity of regional authorities would indeed help to reduce the contractual drawbacks identified in our research. For that purpose, the transport authorities could benefit from the informational

externalities generated by the efficiency benchmarking of their regional operator. To some extent, our paper illustrates that such a regulatory tool could be technically implementable in the sector. This recommendation is in line with the position paper of the Independent Regulators' Group - Rail on competitive tendering for public service contracts[60] suggesting that yardstick competition "*can be seen as a means to introduce some "virtual" competition into industries where market competition is either not viable or not desired.*"

[60] IRG-Rail (15) 3 - Position paper on current proposals on competitive tendering for public service contracts, 14-15 April 2015.

Conclusion

This dissertation started by exposing the underlying logic of the reforms to revitalise railways in Europe: first separating vertically the management of infrastructure from train operations, a necessary condition to later introduce rejuvenating competition in the downstream market. The theoretical foundations of this approach are derived from the theory of contestable market by Baumol et al. [1982] which has had an influence in all network industries in Europe. But this strategy still raises many questions: on the relevance of vertical separation, on the design of markets and on the introduction of competition in the downstream market, and therefore on regulation. Using the fact that the dust has not settled yet and the reforms are still under way in Europe, this dissertation analysed some of the misalignments generated by the changes in organisation to address those questions. This conclusion first sums up the main findings in the dissertation and their policy implications, before bringing up some of the limits of our work and areas for future research.

Summary of main findings and policy implications

On the plurality of reforms

The first observation we made is that the process of unbundling railways took various forms in Europe which can be classified in three main forms. The first wave of countries chose to establish a full unbundling, that is having an infrastructure manager that was legally, organisationally and institutionally separated from the incumbent train operators. Concerns over such a radical change drove the next countries to set in place other vertical governance structures. The holding structure which consisted in having an infrastructure manager owned by a holding company, also owning the incumbent train operator and a model called separation of capacity allocation, that is having an independent infrastructure that was stripped to its core function: allocating capacity, without being in charge of maintenance.

The plurality of vertical structures can be seen as a means to overcome some of the associated drawbacks of vertical separation such as the loss of synergies but also the ones that are given by the Theory of the Firm, limited coordination mechanisms due to the impossibility to enforce *ex post* commitment. On this regard, the second chapter illustrates the challenges raised by vertical separation in order to achieve coordination. Indeed, in order to match maintenance and commercial constraints, the capacity allocation process in railways was made flexible to suit the needs of the firms. But we show that this flexibility can entail wasted capacities in the absence of commitment, and increase the overall cost of maintenance or reduce the output. If the holding structure, or the separation of capacity allocation might mitigate such a risk and can achieve coordination through mechanisms other than prices, the corollary of making this hypothesis would then be the risk of foreclosure for a new entrant. All in all, the arguments made are similar to the pros and cones of choosing

vertical separation versus vertical integration: there is a central trade-off between enabling competition versus losing economies of scope.

In the first chapter, we want to test if this trade-off also applies to the intermediate governance structure existing in Europe. If it is the case, then one would expect that the cost for maintaining the network in an intermediate form to be less costly than a fully unbundled structure. Yet one must keep in mind that a perceived advantages of choosing a fully unbundled structure could lie in its *radicalness*, creating a shock to rejuvenate the sector dominated by a state monopoly and dismantling a firm too big to regulate. Due to reverse causality, it might create a bias when we estimate the impact of the vertical structure on costs. In order to take into account endogeneity, we estimate a two-stage least square regression with instruments derived from the organisation of electricity sector. Our results are twofold. First we find that full unbundling tends to reduce the cost of maintenance contrary to expectation. Also we find that endogeneity does create a bias that reduces the positive impact of full unbundling when not taken into account.

Supplementing market mechanisms

Having concluded that full unbundling does not entail an increase in the cost of maintaining the network, the next condition is that there are mechanisms that allow coordination and avoid some market failures. In the introduction it was stressed, as stated by Hogan [2002] that "*markets are made and don't just happen*" in the case of network industry, which leads to the need of a well established regulatory framework before changes are made (Newbery [2002]). In order to test the impact of regulation in the case of railways, we tested in the first chapter two variables : the age of the regulator as well as the difference between the age of the unbundling and the age of the regulator. We find

that, once we correct for endogeneity, the variable having a significant impact on costs is the second one. Having anticipated the creation of a regulator before proceeding with the unbundling decreases the costs of maintenance on the network.

This results highlights the central role we believe regulation has in the success of an unbundled railway market and which goes beyond enforcing cost efficiency. Indeed, an independent regulator should aim at aligning interests of the upstream and downstream firms when they become to divergent. This goal can be achieved through different means, for instance in its role of appeal body, the regulator can contribute to reduce opportunistic behaviours and enforce more complex contracts, when a non specialized jurisdiction can not. The review of the pricing scheme is another medium, and the regulator should make sure that prices send the appropriate signals to the downstream market.

An illustration is given in chapter 2, on how regulation can supplement market mechanisms in the case of railways. The intuition behind this chapter is that switching coordination mechanisms from hierarchies to prices - a consequence of vertical separation according to Coase [1937] - can entail difficulties in matching the consumption of capacity for maintenance and for commercial purposes. This issue is quite similar to the *news-vendor problem* of managing a time sensitive product. In the case of railways, we find that a provision in the pricing scheme framework set by European law, namely the possibility to apply a mark-up to access charges when the market can bear it, can mitigate the risk of having a mismatch between supply and demand for capacity. More precisely we find that, when the upstream mark-ups are under the supervision of a regulator, the infrastructure manager can be steered toward a coordinated outcome.

In terms of policy recommendation, this result stresses the role the regulator has to play in reviewing access charges. This review has to be comprehensive

and aim at aligning the interests of the infrastructure manager with those of the downstream firms. An analogy can be made with the problematic of the *battle of the sexes*. As a minimum, the role of the regulator can be to achieve a correlated equilibrium - as introduced by Aumann et al. [1974] - in the industry and avoid having wasted capacities. Furthermore, the regulator can modify the pay-offs of the infrastructure manager and influence the outcome of the *game* in order to ensure coordination.

Gradual introduction of competition

In the second chapter we also develop the insight that regulation has to take into account the degree of competition. The ability of the downstream firms to bear deviations of access charges from marginal cost ultimately depends on their ability to make profits on the downstream market. It should therefore vary with the intensity of competition. In particular, as competition increases, there is a risk that the market will be less able to bear the mark-ups and it could lead to a decrease in outputs. This result depends on how competition is introduced. For instance in our framework, the drawbacks of vertical separation come from the absence of outside options for train operators which is the case if it is impossible to redeploy rolling stocks from one market to another once the investment has been made. This drawback could therefore be limited if the interoperability was increased across European countries and a train operator could easily widen its span, thus becoming independent from the decisions of one infrastructure manager or from a shock in demand.

This brings us to an important caveat attached with the introduction of competition. Despite the separation of infrastructure, technical barriers to entry remain, making the downstream market not perfectly contestable. Those barriers lead to complications to introduce competition on the market - when

competing market players offer the same products- as described in chapter 2 but also competition for the market - when an operator is selected after a bidding process to serve demand - as it is the case in the third chapter of this dissertation on regional railway transport in France.

In France, regional transport is subject to public service contracts but despite a tender procedure, the contract is awarded to the incumbent which still benefits from a legal monopoly. In the absence of competitive pressure, the efficiency of the procurement, process relies heavily on the expertise of the Regions, acting as transport organising authorities (the public buyers). Building up on the typology of capacities for a public buyer identified by Brown and Potoski [2003] we find, in the case of regional transport in France, two items to be particularly important: the *ex ante* design of the contract and monitoring the execution contract. Taking advantage on the common structure of contracts, we identify the parameters that should vary with the expertise specific to each region. We then deduce from the use of a stochastic frontier if those parameters have an impact on the cost efficiency of the contract. In particular we find that the expenditures may be subject to cost padding, cancelling the incentive properties of a revenue cap. Our second result is that efficiency decreases over the duration of the contract, stemming from renegotiations that occur when there is the need to change the scope of the transport services.

While acknowledging the peculiarity of the French case - a tendering procedure but no competition - public buyers would meet similar difficulties if competition does not turn out to be viable when introduced. In this case, transport authorities would (continue to) face severe information asymmetries, which in the case of regional transport in France, translates by an increase of 15% of the cost of the service according to our estimation. To circumvent this issue, our work has the following implication in terms of policy: transport authorities could benefit from informational externalities generated by the use of benchmarking tools. In other words, introduce yardstick competition.

Yardstick competition was defined by Shleifer [1985] where the optimal payments made to a firm depend on a comparison of costs with other comparable firms. In this matter, the independent regulator could be in charge of centralising information and bring its expertise to the regions who need assistance in assessing efficiency. Or, as it has been done for instance in Japan (see for instance Mizutani et al. [2009]) set an explicit regulatory mechanisms to create virtual competition where non exist. It is all the more relevant that the fourth railway package contemplates the perspective of having only one operator expressing its interest even in the case of an open tendering procedure[61] . A transport authority would then have to engage in direct negotiation with the operator, with all the caveats attached.

Limits and areas for future research

One of the main limits of this dissertation lies in the dataset we constructed for our empirical analysis and the limited size of the samples we use in the first and third chapter. When it comes to infrastructure managers, data is *de facto* limited by the fact that there is only one firm by country and leads to a small sample of countries. Similarly for regional transport in chapter 3, the absence of allotment in PSO contracts in France limits the number of observations to the number of comparable regions, that is 20 observations[62] per year. And it could be reduced to 11 due to the administrative changes that occurred in France. In order to circumvent this issue, we use a panel, in both case for four years between 2009 and 2012. Once again, we acknowledge that it is not a long time period. We found that data availability is very limited before 2009. In the case of infrastructure management, it stems partly from the fact separation is effective in all European countries since this date. Another reason probably

[61] See the text provisionally agreed at the trilogue on 19 April 2016 (8061/16ADD1REV2)

[62] Out of the 22 regions in France, regional transport services in Corsica and Ile-De-France are too different to be compared

lies in the fact that state monopolies are not necessarily keen on transparency and policy makers are gradually understanding the virtue of information. For both chapters, the data we use is based on monitoring reports and the first edition was in 2009. In the first chapter, the data comes from a report from the Commission to the Council and the European Parliament. In the second case, the initiative was taken directly by the regions with the purpose of increasing comparability. The increasing availability of data should benefit future empirical work on rail roads.

A second point worth mentioning in this conclusion is that this dissertation is focused on conditions and implications of introducing competition, but little is said about the actual emergence of competition itself. For instance we do not ascertain empirically the effect of the vertical structures on entry. Incidentally, the optimal way to open competition downstream is still uncertain and a further assessment would have to be made to identify the pros and cons of competition on the market and competition for the market. For example, the opening to competition of the market for international passenger services has translated so far into a very limited number of new entrants. In France, the only operator not owned by the incumbent offers two new routes to Italy, but chose not to compete with any existing services and could raise doubts on how competitive the market will become. Furthermore the implications in terms of coordination between the upstream and downstream firms would not necessarily be the same depending on the fact that competition in the market or for the market is dominant.

We do not analyse either if competition has been beneficial in the countries where it has been introduced. As it is, we cannot assess the overall success of the strategy in Europe to revitalise railways. With the introduction of competition in all European countries for national passenger services in the coming year, and the increasing data availability we have described, such an assessment will become possible. Especially, the heterogeneity we have described in

the first chapter should provide interesting counter-factuals on the effects of vertical governance structures on the emergence of competition.

In a nutshell, the scope of our results is to say that with the right regulatory tools, vertical separation should work. But, given that is does not necessarily imply that downstream markets will become all of a sudden contestable, regulation of downstream services would be, at least temporarily, justified. On both those matters, we believe that the introduction of an independent regulatory agency act as a supplement to dawning market mechanisms.

References

Affuso, L. and Newbery, D. M. G. (2000). Investment, Reprocurement and Franchise Contract Length in the British Railway Industry. CEPR Discussion Papers 2619.

Aigner, D., Lovell, K., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1):21–37.

Amsler, C., Prokhorov, A., and Schmidt, P. (2016). Endogeneity in stochastic frontier models. *Journal of Econometrics*, 190(2):280–288.

Arrow, K. J., Harris, T., and Marschak, J. (1951). Optimal inventory policy. *Econometrica: Journal of the Econometric Society*, 19(3):250–272.

Aumann, R. J. et al. (1974). Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics*, 1(1):67–96.

Auriol, E. (2000). Concurrence par comparaison: Un point de vue normatif. *Revue Economique*, 51(3):621–634.

Averch, H. and Johnson, L. L. (1962). Behaviour of the firm under regulatory constraint. *American Economic Review*, 52(5):1052–1069.

Bailey, E. E. and Baumol, W. J. (1983). Deregulation and the theory of contestable markets. *Yale Journal on Regulation*, 1:111.

Bajari, P. and Tadelis, S. (2001). Incentives versus transaction costs: A theory of procurement contracts. *RAND Journal of Economics*, 32(3):387–407.

Baron, D. P. and Myerson, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica*, 50(4):911–930.

Battese, G. and Coelli, T. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, 20(1):325–332.

Battese, G. E. and Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: with application to paddy farmers in india. In *International Applications of Productivity and Efficiency Analysis*, pages 149–165. Springer.

Baumol, W. J., Panzar, J. C., Willig, R. D., Bailey, E. E., Fischer, D., and Fischer, D. (1982). Contestable markets and the theory of industry structure.

Beard, T. R., Kaserman, D. L., and Mayo, J. W. (2001). Regulation, vertical integration and sabotage. *Journal of Industrial Economics*, 49(3):319–33.

Beard, T. R., Macher, J. T., and Vickers, C. (2015). This time is different (?): Telecommunications unbundling and lessons for railroad regulation. *Review of Industrial Organization*, 49(2):1–22.

Benedetto, V., Smith, A. S., and Nash, C. A. (2015). Rail regulation in europe: an industry-based survey. *Working paper*.

Berechman, J. (1993). *Public Transit Economics and Deregulation Policy*. North-Holland, Amsterdam.

Bitzan, J. D. (2003). Railroad costs and competition: The implications of introducing competition to railroad networks. *Journal of Transport Economics and Policy*, 37(2):201–225.

Boiteux, M. (1956). Sur la gestion des monopoles publics astreints á l'equilibre budg{etaire. *Econometrica*, 24(1):22–40.

Bougheas, S. and Worrall, T. (2012). Cost padding in regulated monopolies. *International Journal of Industrial Organization*, 30(4):331–341.

Boycko, M., Shleifer, A., and Vishny, R. W. (1996). A theory of privatization. *Economic Journal*, 106(435):309–319.

Brown, T. L. and Potoski, M. (2003). Contract–management capacity in municipal and county governments. *Public Administration Review*, 63(2):153–164.

Cabral, L. M. and Riordan, M. H. (1989). Incentives for cost reduction under price cap regulation. *Journal of Regulatory Economics*, 1(2):93–102.

Caillaud, B. (2003). Allocation des sillons ferroviaires: la possibilité dâenchères. *Économie publique/Public economics*, (10).

Cambini, C. and Jiang, Y. (2009). Broadband investment and regulation: A literature review. *Telecommunications Policy*, 33(10):559–574.

Chaudhuri, K., Chowdhury, P., and Kumbhakar, S. C. (2015). Crime in india: specification and estimation of violent crime index. *Journal of Productivity Analysis*, 43(1):13–28.

Cherchye, L., De Rock, B., Estache, A., and Verschelde, M. (2015). Efficiency measures in regulated industries: History, outstanding challenges and emerging solutions. *Working Paper*.

Coase, R. H. (1937). The nature of the firm. *Economica*, 4(16):386–405.

Coelli, T., Estache, A., Perelman, S., and Trujillo, L. (2003). *A Primer on Efficiency Measurement for Utilities and Transport Regulators*. The World Bank Institute, Washington.

Coelli, T., Rao, P., O'Donnell, C., and Battese, G. (2005). *An introduction to efficiency and productivity analysis*. Springer Science and Business Media.

Coelli, T. J., Gautier, A., Perelman, S., and Saplacan-Pop, R. (2013). Estimating the cost of improving quality in electricity distribution: A parametric distance function approach. *Energy Policy*, 53:287–297.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431–1451.

Cremer, H., Crémer, J., and De Donder, P. (2006). Legal vs ownership unbundling in network industries. *CEPR Discussion Paper*, (5767).

Crozet, Y., Nash, C., and Preston, J. (2012). Beyond the quiet life of a natural monopoly: Regulatory challenges ahead for europe's rail sector. *CERRE, Brussels*.

Dal Bó, E. (2006). Regulatory capture: a review. *Oxford Review of Economic Policy*, 22(2):203–225.

Dalen, D. M. and Gómez-Lobo, A. (2002). Regulatory contracts and cost efficiency in the norwegian bus industry: Do high-powered contracts really work? *Discussion paper no. 6/2002 Norwegian School of Management*.

Dalen, D. M. and Gómez-Lobo, A. (2003). Yardsticks on the road: Regulatory contracts and cost efficiency in the Norwegian bus industry. *Transportation*, 30(4):371–386.

Dupuit, J. (1853). Du monopole des chemins de fer. *Journal des Economistes*.

Economides, N. (1998). The incentive for non-price discrimination by an input monopolist. *International Journal of Industrial Organization*, 16(3):271–284.

Edgeworth, F. Y. (1888). The mathematical theory of banking. *Journal of the Royal Statistical Society*, 51(1):113–127.

Farell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, 120(3):253–281.

Farrell, J. and Rabin, M. (1996). Cheap talk. *The Journal of Economic Perspectives*, 10(3):103–118.

Farsi, M., Filippini, M., and Greene, W. (2005). Efficiency measurement in network industries: application to the swiss railway companies. *Journal of Regulatory Economics*, 28(1):69–90.

Finger, M. (2014). Governance of competition and performance in european railways: an analysis of five cases. *Utilities Policy*, 31:278–288.

Friebel, G., Ivaldi, M., and Vibes, C. (2010). Railway (de) regulation: a european efficiency comparison. *Economica*, 77(305):77–91.

Fudenberg, D. and Tirole, J. (1995). A theory of income and dividend smoothing based on incumbency rents. *Journal of Political economy*, 103(1):75–93.

Gagnepain, P. and Ivaldi, M. (2002). Incentive regulatory policies: The case of public transit systems in france. *RAND Journal of Economics*, 33(4):605–629.

Gagnepain, P. and Ivaldi, M. (2007). Contract choice, incentives, and political capture in public transit. Technical report, mimeo.

Gal-Or, E. (1991). Vertical restraints with incomplete information. *The Journal of Industrial Economics*, 39(5):503–516.

Galperin, H., Viecens, F., Mariscal, J., Trillas, F., and Montoya, M. A. (2013). Independent regulators: Theory, evidence and reform proposals. *info*, 15(3):39–53.

Gautier, A. and Yvrande-Billon, A. (2013). Contract renewal as an incentive device. An application to the french urban public transport sector. *Review of Economics and Institutions*, 4(Winter):Art 2.

Gibbons, R. (2005). Four formal(izable) theories of the firm? *Journal of Economic Behavior and Organization*, 58(2):200–245.

Gilardi, F. (2005). The institutional foundations of regulatory capitalism: the diffusion of independent regulatory agencies in western europe. *The Annals of the American Academy of Political and Social Science*, 598(1):84–101.

Giocoli, N. (2014). *Predatory Pricing in Antitrust Law and Economics: A Historical Perspective*. Routledge.

Glachant, J.-M., Khalfallah, H., Perez, Y., Rious, V., and Saguan, M. (2013). Implementing incentive regulation through an alignment with resource bounded regulators. *Competition and Regulation in Network Industries*, 14(3):265–291.

Grossman, S. J. and Hart, O. (1986). The cost and benefits of ownership: A theory of vertical and lateral intégration. *Journal of Political Economy*, 94(4):691–719.

Growitsch, C. and Wetzel, H. (2009). Testing for economies of scope in european railways: An efficiency analysis. *Journal of Transport Economics and Policy*, 43(1):1–24.

Guasch, J.-L., Laffont, J.-J., and Straub, S. (2008). Renegotiation of concession contracts in Latin America: Evidence form the water and transport sector. *International Journal of Industrial Organization*, 26(2):421–442.

Hadley, A. T. (1897). *Railroad transportation: Its history and its laws*. GP Putnam's sons.

Haenel, H. (2008). Ecrire l'acte II de la revolution ferroviaire regionale. *Rapport au Premier ministre*.

Hart, O. and Moore, J. (1990). Property rights and the nature of the firm. *The Journal of Political Economy*, 98(6):119–1158.

Hart, O. and Tirole, J. (1990). Vertical integration and market foreclosure. *Brookings papers on economic activity. Microeconomics*, 1990:205–286.

Hogan, W. W. (2002). Electricity market restructuring: reforms of reforms. *Journal of Regulatory Economics*, 21(1):103–132.

Holmstrom, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization*, 7:24–52.

Holmstrom, B. and Tirole, J. (1991). Transfer pricing and organizational form. *Journal of Law, Economics, and Organization*, 7(2):201–28.

Iossa, E. and Rey, P. (2014). Building reputation for contract renewal: Implications for performance dynamics and contract duration. *Journal of the European Economic Association*, 12(3):549–574.

Ivaldi, M. and McCullough, G. J. (2001). Density and integration effects on class I U.S. freight railroads. *Journal of Regulatory Economics*, 19(2):161–182.

Ivaldi, M. and McCullough, G. J. (2008). Subadditivity tests for network separation with an application to U.S. railroads. *Review of Network Economics*, 7(1):1–13.

Jamasb, T. and Pollitt, M. (2000). Benchmarking and regulation: international electricity experience. *Utilities Policy*, 9(3):107–130.

Jensen, M. and Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4):305–360.

Jorgensen, F., Pedersen, P. A., and Volden, R. (1997). Estimating the inefficiency in the norwegian bus industry from stochastic cost frontier models. *Transportation*, 24(4):421–433.

Joskow, P. L. (2005). Incentive regulation in theory and practice: electricity distribution and transmission networks. *AEI-Brookings Joint Center Working Paper*, (05-18).

Joskow, P. L. and Schmalensee, R. (1983). *Markets for power: an analysis of electric utility deregulation*. Massachusetts Inst. of Technology Press, Cambridge, MA.

Kerstens, K. (1996). Technical efficiency measurement and explanation of french urban transit companies. *Transportation Research A*, 30(6):431–452.

Klein, B., Crawford, R. G., and Alchian, A. A. (1978). Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics*, 21(2):297–326.

Klein, B. and Murphy, K. (1988). Vertical restraints as contract enforcement mechanisms. *Journal of Law and Economics*, 31(2):265–297.

Koske, I., Wanner, I., Bitetti, R., and Barbiero, O. (2015). The 2013 update of the oecd's database on product market regulation. Technical report.

Kumbhakar, S. and Lovell, K. (2001). *Stochastic Frontier Analysis*. Cambridge University Press.

Laffont, J.-J. and Pouyet, J. (2004). The subsidiarity bias in regulation. *Journal of Public Economics*, 88(1-2):255–283.

Laffont, J.-J. and Tirole, J. (1986). Using cost observation to regulate firms. *Journal of Political Economy*, 94(3):614–641.

Laffont, J.-J. and Tirole, J. (1988). The dynamics of incentives contracts. *Econometrica*, 56(5):1153–1175.

Laffont, J.-J. and Tirole, J. (1992). Cost padding, auditing and collusion. *Annales d'Economie et de Statistique*, pages 205–226.

Laffont, J.-J. and Tirole, J. (1993). *A Theory of Incentives in Procurement and Regulation*. MIT Press, Cambridge.

Leaver, C. (2009). Bureaucratic minimal squawk behaviour: theory and evidence from us regulatory policy. *American Economic Review*, 99(3):572–607.

Legros, P. and Newman, A. F. (2009). A price theory of vertical and lateral integration. CEPR Discussion Papers DP7211.

Leibenstein, H. (1966). Allocative efficiency vs." x-efficiency". *The American Economic Review*, 56(3):392–415.

Lévêque, J. (2004). An application proposal of yardstick competition for the regional markets of the french railway system. *AET European Transport Conference*.

Levy, B. and Spiller, P. T. (1994). The institutional foundations of regulatory commitment: A comparative analysis of telecommunications regulation. *Journal of Law, Economics, and Organization*, 10(2):201–246.

Li, L. (2002). Information sharing in a supply chain with horizontal competition. *Management Science*, 48(9):1196–1212.

Maggi, G. and Rodriguez-Clare, A. (1995). Costly distortion of information in agency problems. *The RAND Journal of Economics*, 26(4):675–689.

Mandy, D. and Sappington, D. (2007). Incentives for sabotage in vertically related industries. *Journal of Regulatory Economics*, 31(3):235–260.

Mandy, D. M. (2000). Killing the goose that may have laid the golden egg: Only the data know whether sabotage pays. *Journal of Regulatory Economics*, 17(2):157–72.

Martin, B. L. and Jayakar, K. (2013). Moving beyond dichotomy: Comparing composite telecommunications regulatory governance indices. *Telecommunications Policy*, 37(9):691–701.

Marty, F. (2012). Une régulation du secteur de lâénergie au travers des procédures d'engagements? réflexions sur le contentieux concurrentiel européen. *Économie publique/Public economics*, (26-27):93–128.

Meeusen, W. and Van Den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18(2):435–444.

Merkert, R. and Nash, C. A. (2013). Investigating european railway managersâ perception of transaction costs at the train operation/infrastructure interface. *Transportation Research Part A: Policy and Practice*, 54:14–25.

Mizutani, F., Kozumi, H., and Matsushima, N. (2009). Does yardstick regulation really work? empirical evidence from japans rail industry. *Journal of regulatory economics*, 36(3):308–323.

Mizutani, F. and Uranishi, S. (2013). Does vertical separation reduce cost? An empirical analysis of the rail industry in European and East Asian OECD countries. *Journal of Regulatory Economics*, 43(1):31–59.

Nash, C. (1997). The separation of operation from infrastructure in the provision of railway services: The british experience. *The Separation of Operation from Infrastructure in the Provision of Railway Service*, pages 53–89.

Nash, C. (1999). Capri, concerted action on transport pricing research integration, project funded by the european commission. *Institute for Transport Studies, University of Leeds, Leeds*.

Newbery, D. M. (2002). Regulatory challenges to european electricity liberalisation. *Swedish Economic Policy Review*, 9(2):9–44.

Normann, H.-T., Möllers, C., Snyder, C. M., et al. (2015). Communication in vertically related markets: Experimental evidence. In *Annual Conference 2015 (Muenster): Economic Development-Theory and Policy*, number 112842. Verein für Socialpolitik/German Economic Association.

Novshek, W. and Sonnenschein, H. (1982). Fulfilled expectations cournot duopoly with information acquisition and release. *The Bell Journal of Economics*, 13(1):214–218.

Perennes, P. (2014). Les économistes et le secteur ferroviaire: deux siècles d'influence réciproque. *L'Économie politique*, (2):101–112.

Piacenza, M. (2006). Regulatory contracts and cost efficiency: Stochastic frontier evidence from the Italian local public transport. *Journal of Productivity Analysis*, 25(3):257–277.

Pittman, R. (2005). Structural separation to create competition? The case of freight railways. *Review of Network Economics*, 3(4):1–16.

Preston, J. (2002). The transaction cost economics of railways. *TSU REF 941*.

Quinet, E. (2003). Short term adjustments in rail activity: the limited role of infrastructure charges. *Transport Policy*, 10(1):73–79.

Reiffen, D., Schumann, L., and Ward, M. R. (2000). Discriminatory dealing with downstream competitors: Evidence from the cellular industry. *Journal of Industrial Economics*, 48(3):253–86.

Reiffen, D. and Ward, M. R. (2002). Recent empirical evidence on discrimination by regulated firms. *Review of Network Economics*, 1(1):1–15.

Rey, P. and Tirole, J. (2007). A primer on foreclosure. *Handbook of industrial organization*, 3:2145–2220.

Riordan, M. H. and Williamson, O. E. (1985). Asset specificity and economic organization. *International Journal of Industrial Organization*, 3(4):365–378.

Rogerson, W. P. (1992). Overhead allocation and incentives for cost minimization in defense procurement. *The Accounting Review*, 67(4):671–690.

Roy, W. and Yvrande-Billon, A. (2007). Ownership, contractual practices and technical efficiency: The case of urban public transport in France. *Journal of Transport Economics and Policy*, 41(2):257–282.

Sappington, D. (2006). On the merits of vertical divestiture. *Review of Industrial Organization*, 29(3):171–191.

Sappington, D. E. (2002). Price regulation and incentives. *Handbook of telecommunications economics*, 1:225–293.

Sappington, D. E. M. and Weisman, D. (2005). Self-sabotage. *Journal of Regulatory Economics*, 27(2):155–175.

Saussier, S. and Tirole, J. (2015). Strengthening the efficiency of public procurement. *Note du CAE*, (22).

Seabright, P. (1996). Accountability and decentralisation in government: An incomplete contracts model. *European economic review*, 40(1):61–89.

Shleifer, A. (1985). A theory of yardstick competition. *RAND Journal of Economics*, 16(3):319–327.

Smith, A. S. and Wheat, P. (2012). Evaluating alternative policy responses to franchise failure: Evidence from the passenger rail sector in britain. *Journal of Transport Economics and Policy (JTEP)*, 46(1):25–49.

Sobel, J. (1999). A reexamination of yardstick competition. *Journal of Economics and Management Strategy*, 8(1):33–60.

Spence, A. M. (1975). Monopoly, quality, and regulation. *The Bell Journal of Economics*, 6(2):417–429.

Staropoli, C. and Yvrande-Billon, A. (2009). La dynamique des réformes d'industries de réseaux : Le cas de l'électricité et des chemins de fer en grande-bretagne. *Annals of Public and Cooperative Economics*, 80(1):89–117.

Stern, J. (1997). What makes an independent regulator independent? *Business Strategy Review*, 8(2):67–74.

Stern, J. (2013). Econometric benchmarking ant its uses by orr: A review. *CCRP Working Paper*, (21).

Stigler, G. J. (1971). The theory of economic regulation. *The Bell journal of economics and management science*, 2(1):3–21.

Stock, J. H. and Watson, M. W. (2003). *Introduction to econometrics*, volume 104. Addison Wesley Boston.

Stoft, S. (2002). Power system economics. *Journal of Energy Literature*, 8:94–99.

Thatcher, M. (2002). Delegation to independent regulatory agencies: Pressures, functions and contextual mediation. *West European Politics*, 25(1):125–147.

Urdanoz, M. and Vibes, C. (2013). Regulation and cost efficiency in the european railways industry. *Journal of productivity analysis*, 39(3):217–230.

Van de Velde, D., Nash, C., Smith, A., Mizutani, F., Uranishi, S., Lijesen, M., and Zschoche, F. (2012). Eves-rail-economic effects of vertical separation in the railway sector. *Report for CER-Community of European Railway and Infrastructure Companies, Amsterdam/Brussels*.

Vickers, J. (1995). Competition and regulation in vertically related markets. *Review of Economic Studies*, 62(210):1–17.

Walras, L. et al. (1897). L'état et les chemins de fer. *Revue du droit public et de la science politique*, 7:417–438.

Weitzman, M. L. (1980). The" ratchet principle" and performance incentives. *The Bell Journal of Economics*, 11(1):302–308.

Wheat, P. and Smith, A. (2008). Assessing the marginal infrastructure maintenance wear and tear costs for britain's railway network. *Journal of Transport Economics and Policy*, 42(2):189–224.

Wheat, P. and Smith, A. S. (2015). Do the usual results of railway returns to scale and density hold in the case of heterogeneity in outputs? a hedonic cost function approach. *Journal of Transport Economics and Policy*, 49(1):35–57.

Williamson, O. E. (1971). The vertical integration of production: Market failure considerations. *American Economic Review*, 61(2):112–123.

Williamson, O. E. (1975). *Markets and Hierarchies, Analysis and Antitrust implications: A Study in the Economics of Internal Organization*. The Free Press, New York, NY, USA.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism*. The Free Press, New York.

List of Tables

List of Figures