FACULTEIT ECONOMIE
EN BEDRIJFSWETENSCHAPPEN

KATHOLIEKE
UNIVERSITEIT
LEUVEN

ESSAYS ON RISK IN ENERGY ECONOMICS

Proefschrift voorgedragen tot
het behalen van de graad van
Doctor in de Economische
Wetenschappen

door

Joris Morbee

Nummer 382 2012

Contents

Doctoral Committee

Advisor:

Prof. dr. Stef Proost (KULeuven)

Chair:

Prof. dr. Erik Buyst (KULeuven)

Members:

Prof. dr. Jo Van Biesebroeck (KULeuven)
Prof. dr. Guido Pepermans (KULeuven, HUB)
Prof. dr. ir. William D'haeseleer (KULeuven)
Prof. dr. ir. Bert Willems (Tilburg University)
Prof. Dr. Christian von Hirschhausen (TUBerlin)
Dr. Stathis Peteves (European Commission)

8

Acknowledgements

I am very grateful to ...

> *Stef, for his incredible patience, unconditional support, and great economic insight;*
>
> *Bert, for generously sharing his exceptional brainpower and ideas;*
>
> *Pieter, for his much appreciated work on nuclear taxation;*
>
> *the members of the Doctoral Committee, for their thorough review of the manuscript and their constructive comments;*
>
> *the KULeuven Energy Institute and the TUMATIM project of the Belgian Science Policy Office, for supporting me financially;*
>
> *the Joint Research Centre of the European Commission, in particular the Energy Systems Evaluation Unit, for providing the opportunity to complete this dissertation;*
>
> *my fellow PhD students and other KULeuven colleagues, for giving me an unforgettable stay in Leuven;*
>
> *my family and friends, for their encouragement and support throughout this journey.*

Thank you so much!

Joris Morbée
February 2012

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Energy markets are characterized by large uncertainties and risks. The annual volatility of the Brent oil price is 28%, meaning that there is a 1-in-3 chance that next year's oil price will be more than 28% higher or lower than this year's price.[1] Similarly, the annual volatility of gas prices for domestic consumers in Belgium/Brussels is 14%. The uncertainty is much larger than in many other goods and services, such as cars, housing, or travel, to name but a few household spending categories. This phenomenon is all the more important since energy is an essential input to many production processes and consumption patterns.

The risk in energy markets has several underlying causes: *technical*, such as the recent application of new techniques that allow for the extraction of 'shale gas', which has depressed gas prices in the US; *macroeconomic*, such as the drop in oil demand following the 2008/2009 global economic crisis, which roughly halved oil prices; and *political*, such as the Russian-Ukrainian gas crisis in 2006 and 2009, or the first oil shock in the 1970s.

Part I of this thesis deals with *political* risk, and analyzes decisions of resource-rich countries that affect the allocation of energy-related rents. Chapter 2 studies the Russian-Ukrainian gas crisis and how it impacts European import strategies. Chapter 3 investigates the taxation of resource extraction in petroleum-producing countries. Chapter 4 also studies taxation, but focuses on a resource that is mostly exploited in Western countries, namely nuclear power. Chapter 5 also deals with Western countries and explores the possible outcome of potential international negotiations on the distribution of rents

[1] The oil price volatility is based on the annual time series 1976-2010 of Brent oil prices in US dollar per barrel, according to BP (2011). The gas price volatility is based on Eurostat (2011) semi-annual prices for households in Belgium/Brussels – before taxes – in the period 1985-2011, excluding seasonal effects.

arising from a trans-European CO_2 pipeline network for Carbon Capture and Storage (CCS).

Part II of this thesis investigates how firms can protect themselves against the risks in energy supply, by *hedging* their exposure. The main challenge in hedging is that energy markets are typically very *incomplete*, in that not enough different contracts (such as options) exist to enable firms to hedge their exposure completely. Chapter 6 analyzes the effect of market incompleteness on welfare and investment incentives in the specific case of an electricity market with demand uncertainty. Chapter 7 provides a generalization of the theory for a generic market structure with non-specified uncertainty.

Table 1.1 compares the chapters of this dissertation in terms of the energy commodities they deal with and the methodological emphasis. The chapters in Part I each deal with a different commodity: gas, oil, electricity and CO_2, respectively. In Part II, Chapter 6 is applied to the case of electricity, while Chapter 7 has generic validity. The methodological emphasis in Chapters 2, 4 and 7 is on theoretical modeling. Chapters 5 and 6 have a less theoretical focus and mainly feature a numerical simulation. Chapter 3 has a substantial empirical section.

Table 1.1: Comparison of chapters in terms of commodity and methodology ("+" means moderate emphasis. "++" means strong emphasis.)

Chapter	**Commodity**	**Model**	**Simulation**	**Econometrics**
2	Gas	++	+	
3	Oil	+		++
4	Electricity	++	+	
5	CO_2		++	
6	Electricity	+	++	
7	*(Generic)*	++		

1.2 Summary of chapters

Chapter 2. Europe's dependence on Russian gas imports has been the subject of increasing political concern after gas conflicts between Russia and Ukraine in 2006 and 2009. This paper assesses the potential impact of Russian unreliability on the European gas market, and how it affects European gas import strategy. We also study to what extent Europe should invest in strategic gas storage capacity to mitigate the effects of possible Russian unreliability. The European gas import market is described by differentiated competition between Russia and a – more reliable – competitive fringe of other exporters. The results show that

Russian contract volumes and prices decline significantly as a function of unreliability, so that not only Europe but also Russia suffers if Russia's unreliability increases. For Europe, buying gas from more reliable suppliers at a price premium turns out to be generally more attractive than building strategic gas storage capacity.

This chapter is joint work with Stef Proost, and has been published in *The Energy Journal*, 2010, Volume 31, Issue 4.

Chapter 3. Tax rates on resource extraction vary widely between different countries. This paper develops a political economy model of resource taxation, in order to explain the differences in *government take*, i.e. the share of resource extraction profits appropriated by the governments of resource-rich countries. The theoretical model features taxes on resource extraction, endogenous exploration and a formalized classification of government types. The model predicts that, all else equal, the government take of resource extraction profits would increase with government *autocracy* (as opposed to democracy) and government *benevolence*, and with higher existing reserves base, while the government take would decrease with higher undiscovered reserves or higher import dependence. Nearly all of these effects are confirmed empirically for the case of oil, using OLS on a cross-section of 77 countries. The statistically significant positive effect of government autocracy (i.e. dictatorship) on government take is shown to be robust in an IV analysis, and, to a lesser extent, a panel data regression.

This chapter has been presented at *ETE day* in Leuven on March 3rd 2008, *BEED 2009* at UCL-CORE on February 2nd 2009, the workshop on *Political Economy and Institutions in Transport, Energy and Environment* at KULeuven-CES on February 26th 2009, the *Spring Meeting of Young Economists 2009* in Istanbul on April 23rd-25th 2009, the *International Energy Workshop 2009* in Venice on June 17th-19th, *EAERE 2009* in Amsterdam on June 24th-27th 2009, and a workshop on *Political Economy and the Environment* at UCL-CORE on October 22nd-23rd 2009.

Chapter 4. The taxation of nuclear energy is studied using a stylized model of the electricity sector, with one dominant nuclear producer and a competitive fringe of fossil-fuel plants. We show that an unanticipated tax on nuclear production can generate significant government revenue in the short run without disturbing the market, but will harm investment incentives in the long run, especially if the government cannot credibly commit to a future tax rate. Even if the government is capable of credibly committing to an optimal long-run tax, government revenues from the long-run tax will be very low due to the market power of the incumbent. Lifetime extension agreements negotiated with multiple potential players, and competitive auctioning of new nuclear licenses are shown to be the most attractive policies. The analytical results are illustrated with a numerical simulation for the case of Belgium.

This chapter is joint work with Pieter Himpens and Stef Proost and has been presented at the conference on *The Economics of Energy Markets* at the Toulouse School of Economics in January 2010, an *ETE seminar* in Leuven in March 2010, and the conference in honor of Yves Smeers at CORE-UCL in June 2010. The text has been submitted to *Energy Economics*.

Chapter 5. If CO_2 Capture and Storage (CCS) is to become a viable option for low-carbon power generation, its deployment will require the construction of dedicated CO_2 transport infrastructure. In a scenario of large-scale deployment of CCS in Europe by 2050, the optimal (cost-minimising) CO_2 transport network would consist of large international bulk pipelines from the main CO_2 source regions to the CO_2 sinks in hydrocarbon fields and aquifers, which are mostly located in the North Sea. In this paper, we use a Shapley value approach to analyse the multilateral negotiation process that would be required to develop such jointly optimised CO_2 infrastructure. Using the *InfraCCS* CO_2 pipeline network optimisation tool, we perform numerical simulations on the cost burden allocation of a 28.0 billion euro CO_2 pipeline network, which would be required to reach the EU's 2050 climate goals in the PRIMES-based *Power Choices* scenario. We analyse two cases: one with national pipeline monopolies and one with liberalised pipeline construction. We find that countries with excess storage capacity capture 38% to 45% of the benefits of multilateral coordination, with the higher number corresponding to the case with liberalised pipeline construction. Countries with a strategic transit location capture significant rent in the case of national pipeline monopolies. Finally, the liberalisation of CO_2 pipeline construction reduces by two-thirds the differences between countries in terms of cost per tonne of CO_2 exported. As a side result of the analysis, we find that the resource rent of a depleted hydrocarbon field (when used for CO_2 storage) is roughly \$1 per barrel of original recoverable oil reserves, or 1 euro per MWh of original recoverable gas reserves.

This chapter has been presented at the *International Energy Workshop 2011* in Stanford, US, in July 2011.

Chapter 6. The high volatility of electricity markets gives producers and retailers an incentive to hedge their exposure to electricity prices by buying and selling derivatives. This paper studies how welfare and investment incentives are affected when an increasing number of derivatives are introduced. It develops an equilibrium model of the electricity market with risk averse firms and a set of traded financial products, more specifically:

a forward contract and an increasing number of options. We first show that aggregate welfare (the sum of individual firms' utility) increases with the number of derivatives offered, although most of the benefits are captured with one to three options. Secondly, power plant investments typically increase because additional derivatives enable better hedging of investments. However, the availability of derivatives sometimes leads to 'crowding-out' of physical investments because firms' limited risk-taking capabilities are being used to speculate on financial markets. Finally, we illustrate that players basing their investment decisions on risk-free probabilities inferred from market prices, may significantly overinvest when markets are not sufficiently complete.

This chapter is joint work with Bert Willems and has been published in *Energy Economics*, 2010, Volume 32, pp. 786-795.

Chapter 7. In this paper we show that free entry decisions may be socially inefficient, even in a perfectly competitive homogeneous goods market with non-lumpy investments. In our model, inefficient entry decisions are the result of risk-aversion of incumbent producers and consumers, combined with incomplete financial markets which limit risk-sharing between market actors. Investments in productive assets affect the distribution of equilibrium prices and quantities, and create risk spillovers. From a societal perspective, entrants underinvest in technologies that would reduce systemic sector risk, and may overinvest in risk-increasing technologies. The inefficiency is shown to disappear when a complete financial market of tradable risk-sharing instruments is available, although the introduction of any individual tradable instrument may actually decrease efficiency.

This chapter is joint work with Bert Willems. It is No. 11.17 in the Center for Economic Studies Discussions Paper Series (DPS), CentER Discussion Paper No. 2011-057, and TILEC Discussion Paper No. 2011-029.

Part I

Resource Policy

Chapter 2

Russian Gas Imports in Europe: How Does Gazprom Reliability Change the Game?

This chapter has been published in *The Energy Journal*.

2.1 Introduction

In recent years, security of gas supply has been high on the political agenda in Europe. Gas import dependence of the European OECD bloc will increase from 45% in 2006 to 69% in 2030, according to the IEA (2008) Reference Scenario. Russia plays a crucial role, given that it already supplies more than half of Europe's gas imports and that it has the largest proven natural gas reserves in the world (BP, 2006-2008).[1] This has been a source of increasing political concern, especially since 2006, when Russian gas export monopolist Gazprom launched an effort to increase the gas prices paid by Russia's neighboring states, as shown in Table 2.1. The price conflict in Ukraine sparked strong political reactions in Europe, because it led to interruptions of gas supplies to Europe in the beginning of 2006 and 2009. Energy supply security and in particular the potential unreliability of Russian gas imports became an important topic at EU summits and G8 meetings, and in bilateral discussions with Russia. After the second conflict in January 2009, Czech Prime Minister Topolanek – then President of the European Council – even stated explicitly

[1] In this paper, the terms *Europe* and *European* refer to the EU-27 plus Norway, Switzerland and Iceland, unless indicated otherwise.

Table 2.1: Gas prices for Russia's neighboring states, in USD per tcm

Country	**Price on Dec 31, 2005**	**Increased price demanded by Gazprom**	**Price on Jan 1, 2007**
Ukraine	50	230	130
Belarus	46	200	100
Georgia	100	235	235
Moldova	80	(unknown)	170

Note: tcm = thousand cubic meters. Source: Press sources (2006)

that "the EU must weaken its dependence on Russian gas imports" (IHT, Jan 28, 2009).[2]

This paper provides an economic perspective on Russia's strategic position in the European gas market by answering the following two research questions:

1. What is the potential impact of Russian unreliability on the European gas market, and how does this affect European gas import decisions?

2. To what extent should Europe invest in strategic gas storage capacity to mitigate the effects of potential Russian unreliability?[3]

We study long-term gas contracting in a non-cooperative setting, using a partial equilibrium model of the European gas market, with differentiated competition between one potentially unreliable 'dominant firm' (Russia) and a reliable 'competitive fringe' of other non-European import suppliers. Russia's potential unreliability is modeled by assuming that there is a probability δ that Russia does not comply with the long-term contracts it has signed: with probability δ, Russia 'defaults' and withholds supply to increase its price to monopolistic levels for a duration of 4 months.

[2]It should be noted that the relation between Russia and Europe is very different from the relation between Russia and its neighboring states, which, before the price increase, were receiving gas from Russia at prices below netback parity. In addition, it is suspected that Russia's price increases in neighboring states are a prelude to deregulation of Russia's domestic gas market, which currently also has below-market prices. Both considerations imply Russia had understandable reasons for raising prices to its neighbors. On the other hand, Russian gas prices for Europe in 2006 were already in line with the prices in the middle column of Table 2.1, which made Europe a profitable and important customer for Russia. Given that, in addition, Gazprom was trying to enter the downstream European gas market, Russia was unlikely to act in the same way towards Europe as it did towards Ukraine. Nevertheless, as a result of the Ukrainian gas crises, European politicians and gas consumers clearly started questioning the reliability of Russia as a gas supplier.

[3]Hence, the focus of the paper is on *strategic* storage. This storage is in addition to any storage that is necessary for *technical* reasons such as seasonal and daily fluctuations. Storage for technical reasons is not considered in this paper and is assumed to be sized independently of the supply security considerations raised in this paper.

The numerical analysis in this paper shows that it is not optimal for Russia to cut gas supplies to Europe completely during a crisis: rather, one can expect Russia to reduce its gas supplies by roughly 40% during the 4 months, thereby temporarily increasing gas prices by roughly 40%. More importantly, the analysis shows that not only Europe but also Russia suffers when Russia's probability of default δ increases. Indeed, as Russia becomes – or is perceived as becoming – more unreliable, Europe procures a larger volume of long-term gas import contracts from the competitive fringe. This way, Europe makes itself less dependent on Russia and therefore less vulnerable in the event of Russian withholding. With increasing Russian unreliability, the volume of long-term gas import contracts with Russia decreases while Russia has to grant an ever higher discount in its contracts. The resulting negative impact on Russia's profits is not sufficiently counterbalanced by the gains it makes in case it does not comply with its contracts. As a result, Russia's expected profits are found to decrease as Russia's unreliability increases. As mentioned before, investments in strategic gas storage capacity can reduce Europe's vulnerability. However, the numerical simulations show that strategic storage capacity is only attractive for Europe if Russian unreliability is high (δ of more than 30%) and storage capacity costs are reduced by a factor 3 to 4 compared to typical current cost levels.

Earlier papers have studied Russian gas imports into Europe from different perspectives. Hirschhausen et al. (2005) focus on the strategic interaction between Russia and transit countries such as Ukraine and Belarus. Grais and Zheng (1996) analyze the quantity, price and transit fee of gas contracts between Russia and Europe, in a hierarchical three-stage Stackelberg game in which Russia is the leader, followed by the transit country, which in turn is followed by the response of European demand (factoring in a potential alternative gas supplier). They study the impact of exogenous shocks, e.g. an exogenous change in the preference for Russian gas over other gas, and they mention reliability as a potential cause of such a shock. Our model has a non-cooperative multi-stage structure similar to Grais and Zheng (1996), but goes a step further by explicitly examining how reliability affects the demand for Russian gas compared to gas from other suppliers: a demand shift resulting from a change in (un-)reliability is an endogenous effect in our model. In addition, our paper investigates investment in strategic gas storage capacity. On the flip side, to keep the paper focused, we do not model the strategic behavior of transit countries.

The effect of a Russian supply interruption has recently been examined by Hartley and Medlock (2009): as part of their analysis of potential futures for Russian gas exports, they use a comprehensive numerical dynamic spatial equilibrium model to study the global supply chain repercussions of a scenario in which Russia withholds roughly one third of its gas supplies to Europe during a 4-month period. Hartley and Medlock (2009) model the interruption as a deterministic shock with exogenous size. In contrast, in our model, the size of the shock is endogenous, and more importantly, there is uncertainty as to whether the shock will occur. Our model provides an analytical study of how the anticipation of a possible shock – in other words, the perception of unreliability – alters strategic decisions. Our methodology for modeling unreliability is taken from the pioneering paper by Nordhaus (1974), who analyzes oil supply interruptions using a model with two regimes: a normal regime and a supply interruption regime, each with its probability. Like Nordhaus (1974), we investigate the option of investing in storage capacity.[4] However, in addition, our model analyzes the contrast between an unreliable supplier and a reliable competitive fringe. In this setting, gas import contracts with the reliable competitive fringe and investments in storage capacity are (imperfect) substitutes.

Since our paper studies long-term gas import contracts, there are similarities with the literature that deals with these contracts (such as Boucher et al., 1987, and Neuhoff and Hirschhausen, 2005) and with the 'hold-up' literature, such as Hubert and Ikonnikova (2004). However, an important difference between the approach in this paper and the approach of Hubert and Ikonnikova (2004) or Ikonnikova and Zwart (2009) is that the latter two papers use cooperative game theory and explicitly model the negotiation/bargaining between the various parties. Our paper, in contrast, describes the gas market in a non-cooperative setting with quantity competition, following the seminal work of Mathiesen et al. (1987), several well-known analyses such as Golombek et al. (1995, 1998), Boots et al. (2004) and more recent work such as Holz et al. (2008) and Lise et al. (2008).[5] Most of this literature considers European consumers as price-takers with linear demand, which is also the approach taken in this paper.[6]

On a broader microeconomic level, the analysis of this paper fits into the literature on differentiated competition. Indeed, as will be shown in Section 2.3, the contrast of a potentially unreliable gas import supplier (in this case: Russia) and a set of reliable import suppliers (in this case: the competitive fringe of other non-European import suppliers), results in a market structure similar to differentiated competition. Singh and Vives (1984) for example, compare Cournot and Bertrand competition in differentiated duopoly, while

[4]Nordhaus (1974) also investigates import taxes, but it turns out that storage is the most specific response to supply security concerns. In this paper, storage shall therefore be used as the exemplification of a broader range of policy measures (e.g. import taxes, rationing, subsidies for renewable energy, etc.).

[5]Note in particular that an analysis of long-term contracts is not considered inconsistent with non-cooperative modeling. On the contrary, Boots et al. (2004) also use (non-cooperative) Cournot-Nash modeling, which they justify by writing "competition can be expected to take place through quantities, since long-term take-or-pay contracts still prevail in the natural gas market" (Boots et al., 2004, p.74).

[6]However, unlike this paper, the models of Boots et al. (2004) and Golombek et al. (1995, 1998) analyze a segmentation of the European market, based on country and/or type of consumer and/or season. Our paper has only one aggregate demand curve. Note that Holz et al. (2008), as an exception, use non-linear demand curves.

Gaudet and Moreaux (1990) do the same for the particular case of nonrenewable natural resources. The main contribution of our paper is that it introduces the notion of unreliability directly into the market structure of the European gas market.

2.2 Model of the European gas market

2.2.1 European demand, domestic supply, and objective function

Europe is modeled as a large number of uncoordinated gas consumers and domestic gas producers, with an overarching government that can decide to invest public funds in gas storage capacity. We assume Europe is a price-taker with a linear *long-run* inverse demand curve for gas:

$$p(q) = \alpha + \beta q \tag{2.1}$$

European domestic producers supply an exogenous and fixed[7] quantity q_D, and the remaining *excess demand* $q - q_D$ needs to be satisfied by non-European imports. *Short-run* demand is also linear, but with a steeper slope β_{SR}:

$$p_{SR}(q) = p^* + \beta_{SR} \cdot (q - q^*) \tag{2.2}$$

with (p^*, q^*) representing the long-run equilibrium.

We assume that decisions on long-term gas import contracts and publicly financed strategic storage capacity investments are based on a combination of the interests of importers, end-consumers, domestic producers and taxpayers. We therefore assume that Europe maximizes the expected total 'European surplus' $E[S]$:

$$\max E[S] \quad \text{with} \quad S = CS + \Pi_D - G \tag{2.3}$$

where CS is the consumer surplus,[8] Π_D represents the profits of domestic producers, and G is the public expenditure on gas storage capacity investments. $-G$ represents the interests of the recipients of marginal expenditures out of general government revenue. Note that equation (2.3) assumes risk-neutrality. Annex 2.D deals with the case of European risk aversion.

[7]A case of elastic domestic supply is discussed in Annex 2.E.

[8]For the sake of simplicity, our model does not go to the level of individual end-consumers such as households, industrial users and power generators. Therefore, the term CS, as we will compute it based on demand curve (2.1), is in fact the *importer* surplus. In practice, this surplus is somehow divided into importers' profits on the one hand and end-consumer surplus on the other hand. We will not make that distinction, since it depends on market power and regulation in individual countries. We will simply refer to CS as 'consumer surplus'.

2.2.2 Non-European gas import suppliers

Excess demand needs to be satisfied by signing long-term import contracts with non-European import suppliers. We assume that the non-European import suppliers have a *dominant firm – competitive fringe* structure.[9] Russia is the 'dominant firm' and the other non-European gas import suppliers are grouped together as the 'competitive fringe'.

Russia is modeled as a monolithic entity, i.e. the Russian state is not distinguished from the gas exporter Gazprom. Russia is assumed to be a risk-neutral profit maximizer. Russia is modeled to be *unreliable*: once the long-term contracts have been signed, there is a probability δ that Russia temporarily does not comply with its previous supply commitments, i.e. Russia 'defaults'. Conversely, there is a probability $(1-\delta)$ that Russia complies with its long-term contracts during the entire period. All participants know the parameter δ upfront.[10] Russia's long-run marginal costs of production are assumed constant at c_R.

The *competitive fringe* is a diversified set of current or potential future non-European gas import suppliers, including both pipeline and LNG supplies. Therefore, we assume that – *as a group* – the competitive fringe is *reliable*: even if Russia defaults, the competitive fringe delivers the originally promised contract quantity q_0 at the originally promised contract price p_0. This requires two assumptions. First, we assume that the long-term gas import contracts between Europe and the competitive fringe are not indexed on any gas spot market price, which would rise sharply in the event of Russian default. In practice, this condition is fulfilled since most current long-term gas import contracts contain little or no indexation on gas spot market prices. Second, we assume that the competitive fringe players do not deviate from their contracts. This is a major assumption, which can be justified by the difference in scale between Russia and each of the other non-European import suppliers. Each of the other non-European import suppliers has much less incentive to be unreliable because the market impact of each of them is much smaller. In addition, a supplier who is perceived as unreliable could face the threat of being replaced by another supplier in the long term. Russia, on the other hand, is hard to replace completely in the long term, even if it behaves unreliably.

[9] This fairly standard model of industrial organization is described in multiple textbooks, e.g. Carlton and Perloff (2000, Chapter 4).

[10] Hence, δ is exogenous, and there is perfect and complete information about it. The rationale for exogeneity of δ is that Russia's decision-makers are also aware of the potential unreliability of the Russian state, and that they do not have full control over Russia's image of unreliability, nor over Russia's actual behavior over the entire period for which gas contracts are signed. For instance, although Russia never cut gas supplies to Europe during the Ukrainian gas conflicts, the conflict nevertheless led to an increased perception of unreliability in Europe. As we will see later, our model shows that if Russia had full control over its unreliability, it would be optimal for Russia to be perfectly reliable ($\delta = 0$). Since we want to study the effects of increased unreliability (whether it is pursued deliberately or not), we make δ exogenous. In Section 2.5, we mention a different approach which could lead to an endogenous δ.

As we will see below in Section 2.2.3, the reliability of the competitive fringe does not mean that – in the event of Russian default – there would be price discrimination between end-consumers of Russian gas and end-consumers of gas from the competitive fringe. There will be only one single end-consumer price.[11] However, the rents that result from the compliance of the competitive fringe in the event of Russian default accrue to European importers. Therefore, the most important implication of our assumption is that these rents are part of the European surplus function S in equation (2.3), and are not part of the profits of the competitive fringe. As for costs, we assume that the long-run marginal cost curve of the competitive fringe is linearly increasing: $c_0 + d_0 q_0$ (with q_0 the volume of long-term gas import contracts supplied by the competitive fringe, and c_0, d_0 positive constants).

The above-mentioned long-run marginal cost functions (i.e. c_R for Russia, and $c_0 + d_0 q_0$ for the competitive fringe) include not only production costs, but also transportation costs. The calibration for the numerical simulations of Section 2.4 will take this into account.[12] Finally, since this paper analyzes the gas market on an aggregated European level and does not model gas delivery to end-consumers, distribution costs are irrelevant.

2.2.3 Structure of the game

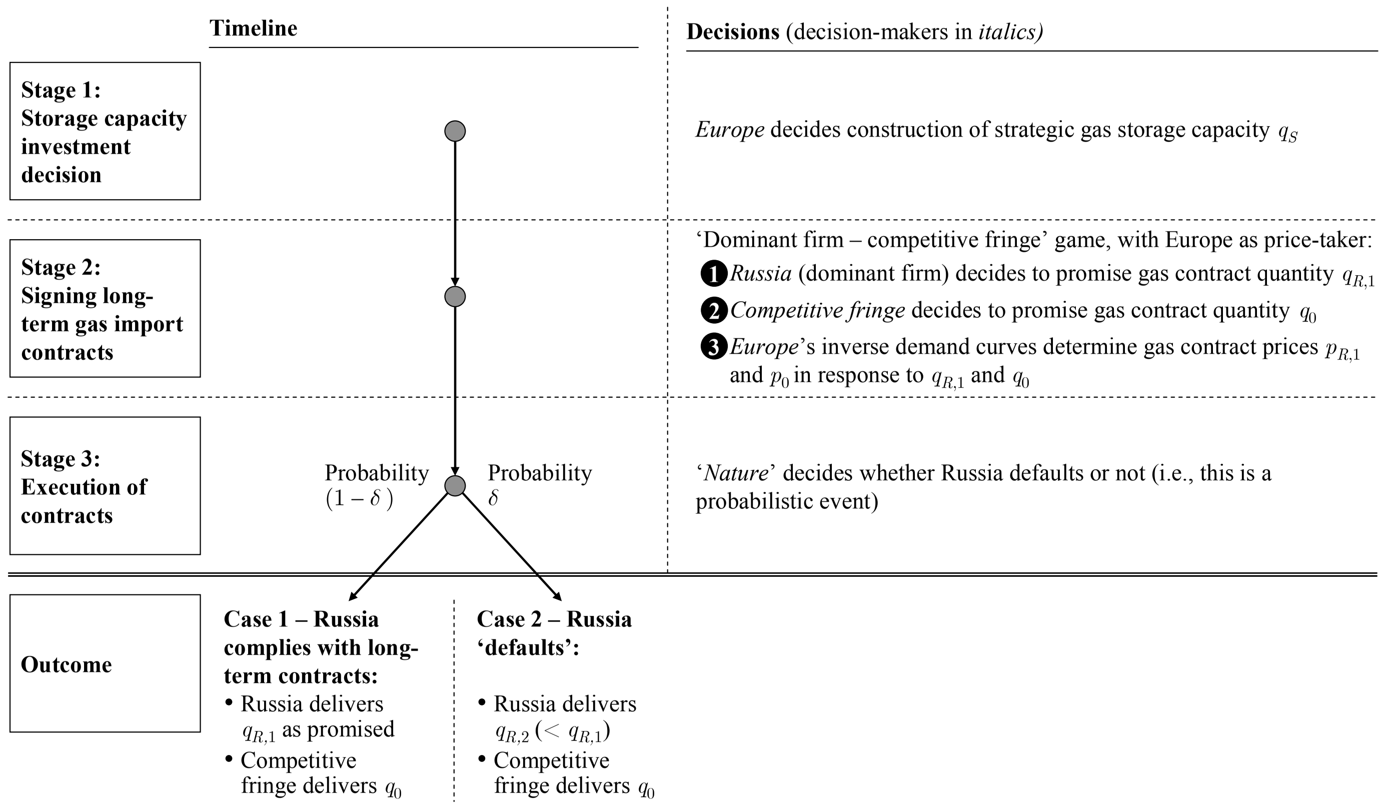
The interaction between Europe, Russia and the competitive fringe, is modeled as a game in three stages. Figure 2.1 explains the different stages of the game. In a nutshell: in Stage 1 Europe decides how much to invest in strategic gas storage capacity; Stage 2 is the stage in which Europe signs long-term gas import contracts with Russia and the competitive fringe; Stage 3 consists in the execution of the long-term gas import contracts, in which Russia may or may not comply with the long-term contracts it has signed. We represent the imported gas quantities by $q_{R,1}$ (Russia complies with long-term contracts), $q_{R,2}$ (Russia defaults) and q_0 (competitive fringe). The corresponding prices are denoted $p_{R,1}$, $p_{R,2}$ and p_0.

Before describing each of the stages in detail, it is important to note that the stochastic outcome of Stage 3 influences the strategic interaction in Stage 2, because Europe and Russia factor the expected value of Stage 3 pay-offs into their decisions in Stage 2. In Stage 3, European surplus is either $S = S_1$ or $S = S_2$ depending on whether Russia complies with its long-term contracts

[11] One could imagine offering interruptible contracts to industrial consumers at a discount. We will not consider that option in this paper.

[12] Note that, while certain parts of the transportation cost can be estimated reasonably well (e.g. LNG shipping from overseas suppliers to Europe), transportation sometimes relies on transit countries (e.g. Ukraine), which leads to additional complexity. For example, Hirschhausen et al. (2005) explicitly study the strategic considerations involved in gas transport from Russia to Europe via transit countries Ukraine and Belarus. While these considerations are important, our paper focuses on the strategic interaction between Europe and its import suppliers. We use OME (2002) estimates of the transit fees.

Figure 2.1: Timeline and decisions in the model proposed in this paper



or not. In Stage 2, Europe therefore tries to maximize the expected surplus $E[S]$. This maximization problem can be translated into demand functions for Russian and other long-term gas import contracts by finding – for given long-term gas contract prices – the optimal long-term gas contract quantities that maximize Europe's expected surplus $E[S]$ in Stage 3. As for Russia, its expected profits in Stage 3 are either $\Pi_R = \Pi_{R,1}$ or $\Pi_R = \Pi_{R,2}$, depending on whether Russia complies with its long-term contracts or not. Therefore, in the 'dominant firm – competitive fringe' game in Stage 2, dominant firm Russia sets the optimal gas contract quantity to maximize its expected profits $E[\Pi_R]$ in Stage 3, taking into account the long-term gas import contract supply curve of the competitive fringe and Europe's above-mentioned demand functions for Russian and other long-term gas import contracts. European demand for long-term gas import contracts will turn out to be *differentiated* between gas import contracts from Russia and gas import contracts from the competitive fringe, because their effect in Stage 3 is different. The rest of this section describes the three stages in more detail.

In *Stage 1*, Europe decides to foresee a quantity q_S (in bcm, i.e. billion cubic meters) of strategic gas storage capacity, to be used as a buffer in case of withholding of gas supply by Russia. Given the long lead times involved in the development of storage sites, this decision cannot be postponed until it is known whether Russia will comply with its contracts or not (i.e. it cannot wait until Stage 3). Furthermore, in our model, the storage capacity investment decision takes place before decisions are made regarding the amounts of long-term gas imports that are contracted from Russia and the competitive fringe (i.e. before Stage 2). The reason is that investment in storage capacity is a decision that Europe can make unilaterally. By making the storage capacity investment decision in a separate stage upfront (Stage 1), Europe gives its storage capacity investment decision an advantageous Stackelberg leadership position in the strategic game with its gas import suppliers. In making the decision about storage capacity investment, Europe takes into account the strategic behavior of Stage 2, and it has perfect and complete information to do so.

In *Stage 2* Europe signs long-term gas import contracts with Russia and with the competitive fringe. Our approach is non-cooperative, with Europe as a price-taker in a 'dominant firm – competitive fringe' model of the long-term gas import contract market. Russia, as the dominant firm, puts a quantity $q_{R,1}$ (in bcm per year) on the European market, for which it receives a price $p_{R,1}$ (in EUR per tcm, i.e. EUR per thousand cubic meters).[13] In making its decision, Russia already takes into account the subsequent decision of the competitive fringe, who put a quantity q_0 (in bcm per year) on the market, for which they receive a price p_0 (in EUR per tcm). The prices $p_{R,1}$ and p_0 are the response

[13]Note that *bcm* per year is consistently used for quantities, while EUR / *tcm* is consistently used for price. The alternative use of bcm and tcm makes the resulting quantity and price numbers conveniently end up in the 0-200 range.

of the European inverse demand functions to the quantities $q_{R,1}$ and q_0. The quantity-price pairs $(q_{R,1}, p_{R,1})$ and (q_0, p_0) represent the long-term gas import contracts signed between Europe and Russia, and between Europe and the competitive fringe, respectively. Because of Russian unreliability, the prices $p_{R,1}$ and p_0 do not need to be the same. Although there are separate inverse demand functions for Russian and other gas – resulting from the behavior of importers – the end-consumers face a single price for gas and cannot choose their own mix of reliable and non-reliable gas. There is a single end-consumer price in each of the two states of the world in Stage 3.

Stage 3, the final stage of the game, is the execution of the long-term gas import contracts signed in Stage 2. Stage 3 is the stage that results in actual pay-offs for the participants to the game. We study one representative year: although the import contracts and storage capacity investment decisions are long-term decisions that will hold for multiple years, all volumes and monetary pay-offs in Stage 3 are shown as annual amounts. In a representative year, there is a probability $1-\delta$ that Russia honors its commitments, and effectively delivers $q_{R,1}$ at a price $p_{R,1}$. This is 'Case 1' (Russia complies with long-term contracts). Figure 2.2 illustrates Case 1 graphically. q_D is the gas supply from European domestic producers, which is assumed to be exogenous and fixed (inelastic). The shaded area, S_1, is the European surplus according to equation (2.3), but without taking any storage capacity investments into account.[14] End-consumers pay a single price corresponding to $p^* \in [p_{R,1}, p_0]$, such that demand at price p^* is exactly equal to $q_D + q_0 + q_{R,1}$.

In a representative year, there is also a probability δ of default, in which case Russia withholds supply to maximize short-run profits. This is 'Case 2' (Russia defaults), which is depicted in Figure 2.3. Assuming that neither q_D nor q_0 can increase in the short run, Russia can set $q_{R,2} < q_{R,1}$, for which it can command a price $p_{R,2} \gg p_{R,1}$. Note that this price is derived from the short-run demand curve (2.2). Europe responds by cutting consumption and using the maximum amount of stored gas, which is constrained by the storage capacity q_S chosen in Stage 1. The storage capacity investment only covers the cost of the storage facility and the capital cost of the unused gas, but not the purchase price of the stored gas itself. The gas withdrawn from the storage will therefore need to be replaced for future crises, and we assume that this can be done at some point at a price equal to p_0. Effectively, the price of using gas from the storage is therefore p_0 (in addition to storage capacity costs, which are sunk). The competitive fringe always delivers q_0 at price p_0, whatever happens in Stage 3. As before, this does not mean that identical end-consumers would pay different prices in the event of Russian default. Since the marginal unit of gas import supply in the short run in case of Russian default

[14] For the sake of simplicity, we assume that the domestic suppliers have zero cost, hence the shaded area for $q \in [0, q_D]$ in Figure 2.2 extends all the way down to the horizontal axis. A non-zero cost would merely constitute a uniform shift of the European surplus function, which would not affect results.

Figure 2.2: Demand and supply in Case 1 – Russia complies with long-term contracts

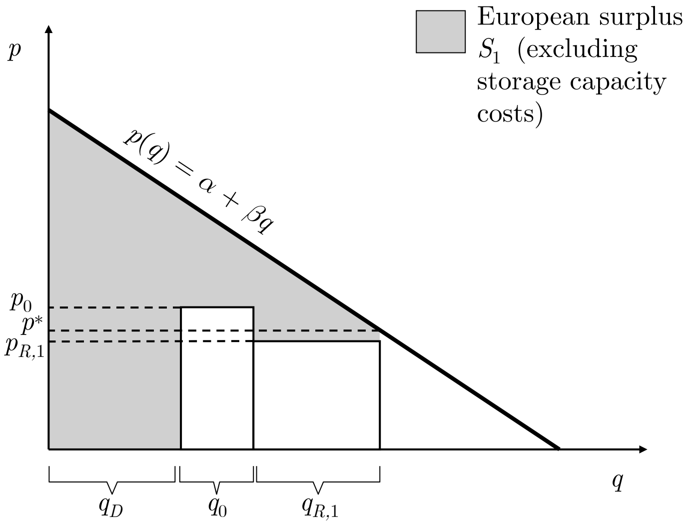
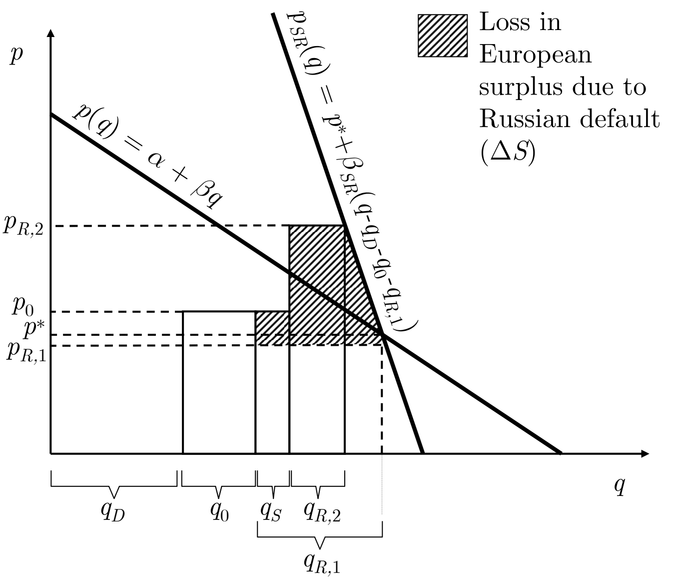


Figure 2.3: Demand and supply in Case 2 – Russia 'defaults'



has a cost $p_{R,2}$ (because only Russia could increase supply), the 'marginal' price for end-consumers should correspond to $p_{R,2}$. While this does create a rent from the fringe supply contracts equal to $(p_{R,2} - p_0)q_0$, the rent is part of the European surplus.[15] In total, the European surplus in case of Russian default is lower than S_1 from Figure 2.2. Figure 2.3 shows ΔS, the loss in European surplus due to Russian default. This loss is discussed in more detail in equation (2.8) in the next section.

The three stages of the game represent three distinct decisions. We assume that this 3-stage game is played once. In practice, the game is obviously repeated after a number of years, but because the lead times for gas projects are very long, we do not consider the repeated game. Finally, if Russia 'defaults' (probability δ), the assumption is that this happens only during a fraction τ of the year. For the remaining fraction $(1-\tau)$ of the year, Russia respects $q_{R,1}$ and $p_{R,1}$. This is comparable with the approach taken by Hartley and Medlock (2009): in their scenario of a Russian supply interruption, they consider a supply reduction which lasts for 4 months in the year 2010. In our model, this corresponds to setting $\tau = 4/12$. If Russia defaults, Figure 2.3 represents the supply situation during a fraction τ of the year, while Figure 2.2 represents the supply situation during the remaining fraction $(1-\tau)$ of the year.[16] The volumes q_D, q_0, q_S, $q_{R,2}$ shown in Figure 2.3 should be interpreted as annualized volumes.[17] This means, in particular, that in order to have access to an annualized storage withdrawal volume of q_S during a Russian default (which

[15]Both in Case 1 and in Case 2, there may be a rent (or loss) for importers, because the price paid by end-consumers (p^* in Case 1, and $p_{R,2}$ in Case 2) does not correspond to the average price paid by the importers (a weighted average of p_0 and $p_{R,1}$ in Case 1, and a weighted average of p_0 and $p_{R,2}$ in Case 2). This positive or negative rent is treated as an integral part of European surplus. In the simplest situation, the rent takes the form of windfall profits (or losses) for gas importers. However, more realistically, we can expect that European governments would intervene and take measures that would redistribute the rents (or losses) to end-consumers, e.g. through non-linear tariffs. One example of non-linear tariffs during a Russian default (Case 2), would be a measure that allows all households a rationed share of q_0+q_S at a price corresponding to p_0, while the remaining gas imports are priced according to $p_{R,2}$. This measure would distribute the rent $(p_{R,2} - p_0)q_0$ to end-consumers whilst ensuring that demand is reduced to the available gas quantity $q_D + q_0 + q_S + q_{R,2}$ because the 'marginal' price perceived by households is still $p_{R,2}$.

[16]The effects of τ and δ on the end result are quite similar. For example, the expression that we will develop for the discount in Russia's contract price (equation (2.11)) will turn out to be a function of the product $\delta\tau$. This would mean that e.g. a 30% chance of a 4-month supply interruption would be to some extent equivalent to a 40% chance of a 3-month supply interruption. In the remainder of the paper, τ is chosen to be 4/12. A different value of τ would roughly result in a horizontal scaling of the results in Figures 2.4 and 2.5.

[17]The notion 'annualized' in this paper means that the quantity is extrapolated to an entire year. For example, suppose Europe has a contract with Russia for $q_{R,1} = 120$ bcm per year, i.e. 10 bcm per month. During a 4-month crisis ($\tau = 4/12$), Russia reduces supply from 10 bcm per month to 6 bcm per month. In that case, we will have $q_{R,2} = 12 \times 6 = 72$ bcm per year. Note however, that the crisis lasts for only 4 months, so the volume supplied by Russia during the crisis is only $4 \times 6 = 24$ bcm. However, to make the magnitude of $q_{R,1}$ and $q_{R,2}$ comparable, we choose to represent annualized amounts in the figures and formulas: everything is expressed per year.

lasts for a fraction τ of the year), a storage capacity of only τq_S is needed. Furthermore, the total European surplus S_2 during a representative year in the event of Russian default is given by:

$$S_2 = (1-\tau)S_1 + \tau(S_1 - \Delta S) = S_1 - \tau\Delta S \tag{2.4}$$

In summary, our model describes Russia's unreliability as a potential 'default' event, with a probability δ of default.[18] The model takes into account two ways for Europe to escape from the unreliability of Russian gas supplies: on the one hand, *diversification* by signing long-term contracts with the competitive fringe, and on the other hand, investments in *strategic storage capacity*.[19] The next section solves the model analytically.

2.3 Analytical solution

We will now analyze the game described in Section 2.2 using backward induction. The three stages of the game will therefore be discussed in reverse order. This section discusses each of the steps in detail.

2.3.1 Stage 3: Execution of contracts

Stage 3 determines the pay-offs for Europe, Russia and the competitive fringe. There are two possible cases: either Russia complies with the long-term contracts it has signed (Case 1) or Russia 'defaults' (Case 2). We will compute the pay-offs of Europe and Russia in each of these two cases.

Case 1: Russia complies with long-term contracts. In this case, the European surplus in a representative year corresponds to the shaded area S_1 in Figure 2.2:

$$S_1 = \alpha\cdot(q_D+q_0+q_{R,1})+\frac{1}{2}\beta\cdot(q_D+q_0+q_{R,1})^2-p_0q_0-p_{R,1}q_{R,1}-c_S\tau q_S \tag{2.5}$$

This result is obtained by applying equation (2.3), or directly graphically from Figure 2.2. The first two terms are simply the integration of the inverse demand curve (2.1) on the interval $[0; q_D + q_0 + q_{R,1}]$. The next two terms represent the expenditure on imported gas, taking into account that both Russia and the competitive fringe comply with their contracts. The last term is the yearly storage capacity cost $G = c_S\tau q_S$. In this expression, c_S is the yearly

[18]The model thus accounts for uncertainty in Russia's behavior (i.e. deliberate supply withholding and price increases), as opposed to *technical* uncertainty. Technical uncertainty is the risk of a sudden supply interruption because of technical failure of e.g. gas pipeline systems. Technical uncertainty is not considered in this paper.

[19]Similarly, the section on supply security in the energy policy communication of European Commission (2007) emphasizes strategic storage and diversification.

constant marginal cost of gas storage capacity, expressed in EUR per tcm per year. One could interpret G as the yearly rent to be paid for the storage site. Note that G has to be paid whether or not the gas is actually withdrawn.

Russia's profits in a representative year in Case 1 are:

$$\Pi_{R,1} = (p_{R,1} - c_R)q_{R,1} \tag{2.6}$$

Case 2: Russia defaults. In this case, Russia does not supply $q_{R,1}$ at $p_{R,1}$, but delivers a lower quantity $q_{R,2}$ at a higher price $p_{R,2}$, for a fraction τ of the representative year. Since Russia's unilateral action comes as a surprise, the relation between $q_{R,2}$ and $p_{R,2}$ is determined by Europe's short-run demand curve (2.2), taking into account the mitigating effect of storage. Annex 2.A derives Russia's optimal quantity and price:

$$\begin{aligned} q_{R,2} &= -\frac{1}{2\beta_{SR}}\left[p^* - \beta_{SR}(q_{R,1} - q_S) - c_R\right] \\ p_{R,2} &= \frac{1}{2}\left[p^* - \beta_{SR}(q_{R,1} - q_S) + c_R\right] \end{aligned} \tag{2.7}$$

Remember that $\beta_{SR} < 0$ and so the second term in the expression for $p_{R,2}$ is a *positive* mark-up. Higher $q_{R,1}$ or lower q_S lead to higher vulnerability of Europe and therefore increase the potential monopoly price $p_{R,2}$. In other words, the impact of Russian unreliability is larger when Russia has a larger market share to begin with (larger $q_{R,1}$), or when Europe has less strategic gas storage capacity (lower q_S). The price $p_{R,2}$ also increases with p^*, because p^* is the starting point of the European price before Russian withholding.

The loss in European surplus during the Russian default can be derived from Figure 2.3:

$$\begin{aligned} \Delta S &= (p_0 - p_{R,1})q_S + (p_{R,2} - p_{R,1})q_{R,2} + \\ &\quad + \left[(p^* - p_{R,1}) + \frac{1}{2}(p_{R,2} - p^*)\right](q_{R,1} - q_S - q_{R,2}) \end{aligned} \tag{2.8}$$

ΔS applies only during the crisis, which lasts for a fraction τ of the year. Note that the value of ΔS is an annualized amount, like $q_{R,2}$. The first term in equation (2.8) is the consumer surplus lost because gas from the storage is more expensive than the original contract with Russia. The second term is the consumer surplus lost because of the Russian price increase from $p_{R,1}$ to $p_{R,2}$. The last line in equation (2.8) is the loss of consumer surplus due to the unserved demand $q_{R,1} - q_S - q_{R,2}$. In Figure 2.3, this corresponds to the part of the striped area above the interval $q \in [q_D + q_0 + q_S + q_{R,2}; q_D + q_0 + q_{R,1}]$. The total European surplus S_2 in a representative year in which Russia defaults can be obtained by substituting equation (2.8) into equation (2.4).

Russia's annualized profits during the 4-month crisis are $(p_{R,2} - c_R)q_{R,2}$. Russia's profits $\Pi_{R,2}$ during an entire representative year in which Russia defaults are simply a weighted average of this amount and $\Pi_{R,1}$, with weights τ and $1-\tau$, respectively.

Since $p^* = \alpha + \beta \cdot (q_D + q_0 + q_{R,1})$, equations (2.5) through (2.8) can be easily expressed as a function of q_S, $q_{R,1}$, q_0, $p_{R,1}$ and p_0, i.e. the parameters of Stages 1 and 2. The results of these equations are taken into account by Europe and Russia when they make strategic decisions in Stage 2.

2.3.2 Stage 2: Signing long-term gas import contracts

In Stage 2, Europe signs long-term gas import contracts with Russia and with the competitive fringe, i.e. the quantities $q_{R,1}$ and q_0 and the prices $p_{R,1}$ and p_0 are determined. In our non-cooperative setting, Russia and the competitive fringe set quantities to maximize profits while taking into account Europe's inverse demand functions. In our solution procedure, we will first determine the European inverse demand functions for long-term gas import contracts, then determine the non-strategic decisions of the competitive fringe, and finally analyze the actions of 'dominant firm' Russia.

European inverse demand functions for long-term gas import contracts. For given long-term gas import contract prices $p_{R,1}$ and p_0, European demand for long-term gas import is derived by finding the optimal quantities $q_{R,1}$ and q_0 that maximize the expected value of European surplus $E[S]$:

$$E[S] = (1-\delta)S_1 + \delta S_2 = S_1 - \delta\tau\Delta S \tag{2.9}$$

with S_1 and ΔS as computed above. The resulting quantities $q_{R,1}(p_{R,1}, p_0, q_S)$ and $q_0(p_{R,1}, p_0, q_S)$ will also be a function of stage-1 decision variable q_S. By inverting the resulting expressions, we obtain the inverse demand functions. For the special case in which $\tau = 1$, $q_S = q_D = 0$ and $c_R = 0$, the inverse demand functions are:[20]

$$\left\{ \begin{array}{rcrcrcrcr} p_{R,1} & = & \alpha\frac{\phi}{1-\delta} & + & \beta\frac{\phi + \frac{3\delta}{4}(k-1)}{1-\delta} & q_{R,1} & + & \beta\frac{\phi}{1-\delta} & q_0 \\ p_0 & = & \alpha\left(\phi + \frac{3\delta}{4}\right) & + & \beta\phi & q_{R,1} & + & \beta\left(\phi + \frac{3\delta}{4}\right) & q_0 \end{array} \right. \tag{2.10}$$

with $k = \beta_{SR}/\beta \gg 1$ and $\phi = 1 - \delta(3 + k^{-1})/4 > 0$. These are the inverse demand functions for *differentiated competition*: because of Russian unreliability in Stage 3, the long-term gas import contract negotiations involve two differentiated goods, namely long-term gas import contracts with Russia on the one hand, and long-term gas import contracts with the competitive fringe on the other hand. The prices of these two goods can be different. The price

[20] This is a special case that leads to insightful analytical expressions. The general case, which is used in the numerical simulations in Section 2.4, can also be expressed analytically, but the resulting expressions are long and not very insightful. The formulas of the general case are available in *Maple* format from the corresponding author upon request.

obtained by Russia depends not only on the quantity set by Russia, but also on the quantity set by the competitive fringe (and vice versa). Note that the differentiation applies only at the contracting stage (Stage 2). Once the gas flows (Stage 3), the gas molecules are identical and there is by assumption no more differentiation in the final consumer market. Since $\beta < 0$, the partial derivatives $\partial p_0/\partial q_0$, $\partial p_0/\partial q_{R,1}$, $\partial p_{R,1}/\partial q_0$ and $\partial p_{R,1}/\partial q_{R,1}$ in equation (2.10) are all negative, as is expected for substitute goods.

A quick check is that for $\delta = 0$ and hence $\phi = 1$, the two suppliers are identical, and equations (2.10) reduce to equation (2.1). When $\delta > 0$, we observe that $\partial p_{R,1}/\partial q_{R,1} < \partial p_0/\partial q_0 < 0$, meaning that Russia faces a more elastic demand curve than the competitive fringe, due to its unreliability. Likewise, $\partial p_{R,1}/\partial q_0 < \partial p_0/\partial q_{R,1}$: Russia's price drops more steeply in response to a quantity increase by the competitive fringe than vice versa. The effect of the asymmetry in Europe's preferences is that Russia's contract price $p_{R,1}$ will be lower than p_0. In other words: when $\delta > 0$, Russia needs to offer Europe a *discount* $\Delta p = p_0 - p_{R,1} > 0$ due to its unreliability. Annex 2.B shows that in general, for small values of δ, the percentage discount is approximately given by:

$$\frac{\Delta p}{p^*} \approx \frac{3}{4}\frac{\delta\tau}{|e_{SR}|}\left(\frac{q_{R,1}}{q^*}\right) \tag{2.11}$$

with e_{SR} the short-run price elasticity of European demand for gas. Russia's discount increases with the probability δ and duration τ of possible interruptions, and with Europe's dependence on Russian long-term gas import contracts as a share of the total gas supply $(q_{R,1}/q^*)$. Russia's discount decreases as Europe's short-run price elasticity of demand $|e_{SR}|$ increases (in absolute terms).

Non-strategic quantity decision by the competitive fringe. By definition, the competitive fringe behaves non-strategically and supplies long-term gas import contracts to Europe up to the point where the contract price equals the marginal cost of additional long-term gas imports. For the special case in which $\tau = 1$, $q_S = q_D = 0$ and $c_R = 0$, we find q_0 by setting p_0 from equation (2.10) equal to the marginal cost $c_0 + d_0 q_0$. We find:

$$q_0 = \frac{\alpha(\phi + \frac{3}{4}\delta) - c_0}{d_0 - \beta(\phi + \frac{3}{4}\delta)} + \frac{\beta\phi}{d_0 - \beta(\phi + \frac{3}{4}\delta)} q_{R,1} \tag{2.12}$$

which provides us with the reaction of the competitive fringe as a function of the decision $q_{R,1}$ by 'dominant firm' Russia. The procedure for the general case is completely analogous.

Quantity decision by 'dominant firm' Russia. The 'dominant firm' Russia faces a residual (inverse) demand function $p_{R,1} = p_{R,1}(q_{R,1})$, which is found by substituting equation (2.12) in the expression for $p_{R,1}$ in equation (2.10). Using the residual demand function, Russia's expected profits $E[\Pi_R] = (1-\delta)\Pi_{R,1} + \delta\Pi_{R,2}$ can be expressed as a function of $q_{R,1}$ (and q_S). Russia chooses a long-term contract quantity $q_{R,1}$ to maximize $E[\Pi_R]$ as a monopolist on the residual demand function.[21] For the special case $\delta = 0$ (and hence also $q_S = 0$) we find the traditional solution of the 'dominant firm – competitive fringe' model:[22]

$$q_{R,1} = -\frac{1}{2\beta d_0}\left[(\alpha + \beta q_D)d_0 - \beta c_0 - c_R(d_0 - \beta)\right] \tag{2.13}$$

Using equation (2.13), we can also assess the impact of the recent 'gas glut' on the European gas market. The gas glut is caused by large-scale extraction of shale gas in the US. The use of shale gas in the US makes more LNG supplies available to Europe, which makes the slope of the cost curve of the competitive fringe less steep, i.e. it lowers d_0. According to equation (2.13) we find:

$$\frac{\partial q_{R,1}}{\partial d_0} = \frac{c_R - c_0}{2d_0^2} > 0 \tag{2.14}$$

hence a decrease in d_0 leads to a reduction of import contracts with Russia. A similar effect applies to discovery and extraction of shale gas in Europe. This would lead to an increase in q_D. Since we have:

$$\frac{\partial q_{R,1}}{\partial q_D} = -\frac{1}{2} \tag{2.15}$$

we find that in the model in this paper, every 1 bcm/y produced from European shale gas leads to a reduction of Russian gas imports by 0.5 bcm/y. When considering the effect of unreliability, it should be noted that the decrease in $q_{R,1}$, whether due to a decrease of d_0 or an increase of q_D, will also *lower* the discount offered by Russia. Indeed, Russia's market share drops, which leads to a smaller discount in equation (2.11).

[21] This is the standard textbook solution to the 'dominant firm – competitive fringe' model (see e.g. Carlton and Perloff, 2000). First of all, note that there is an implicit assumption that Russia is a *Stackelberg price leader* vis-à-vis the competitive fringe. An alternative approach would be to have a *Nash-Cournot* equilibrium between Russia and the competitive fringe. Ulph and Folie (1980) compare the two approaches for the case of oil, and find that the Nash-Cournot approach has the undesirable property that it can lead to an unstable equilibrium in which the dominant firm's profits are lower than under perfect competition. In a slightly different (non-energy) setting, Deneckere and Kovenock (1992) show that in duopolistic price leadership games in which firms have capacity constraints, the smaller firm *strictly prefers* – under a relatively wide range of conditions – to be a follower, as opposed to being the leader or making decisions simultaneously. These results support our assumption that Russia behaves as a Stackelberg leader vis-à-vis the competitive fringe. Secondly, the standard textbook approach mentions an alternative solution, in which a dominant firm with low costs can completely push the competitive fringe out of the market, by setting a price below the 'kink' in the residual demand curve. However, the calibration later in our paper shows that $c_0 < c_R$, so we do not have to consider this alternative solution.

[22] The same comment as in Footnote 20 applies here.

2.3.3 Stage 1: Storage investment decision

Equations (2.10), (2.12) and (2.13) describe special cases in which, among others, $q_S = 0$. In the complete derivation of the model, all these equations are a function of q_S, the storage capacity investment decision that Europe makes in Stage 1. Therefore, also $E[S]$ can be expressed as a function of a single decision variable q_S. In Stage 1, Europe chooses the amount of storage capacity investment q_S that maximizes $E[S]$, obviously subject to the constraint $q_S \geq 0$.[23] Once q_S is determined, $q_{R,1}$ can be computed, followed by q_0, $p_{R,1}$, p_0, $q_{R,2}$ and $p_{R,2}$, according to the generalized versions of the equations above.

The existence of a unique pure-strategy equilibrium is guaranteed because our model consists of a set of sequential decisions, each of which is based on a quadratic (concave) pay-off function.

2.4 Numerical results

The parameters of the model are calibrated on cost data and elasticities from the literature, the 2007 baseline for volume, and the average price 2003-2007. Annex 2.C contains details on the choice of the parameters, while Annex 2.E performs a sensitivity analysis on the elasticities.

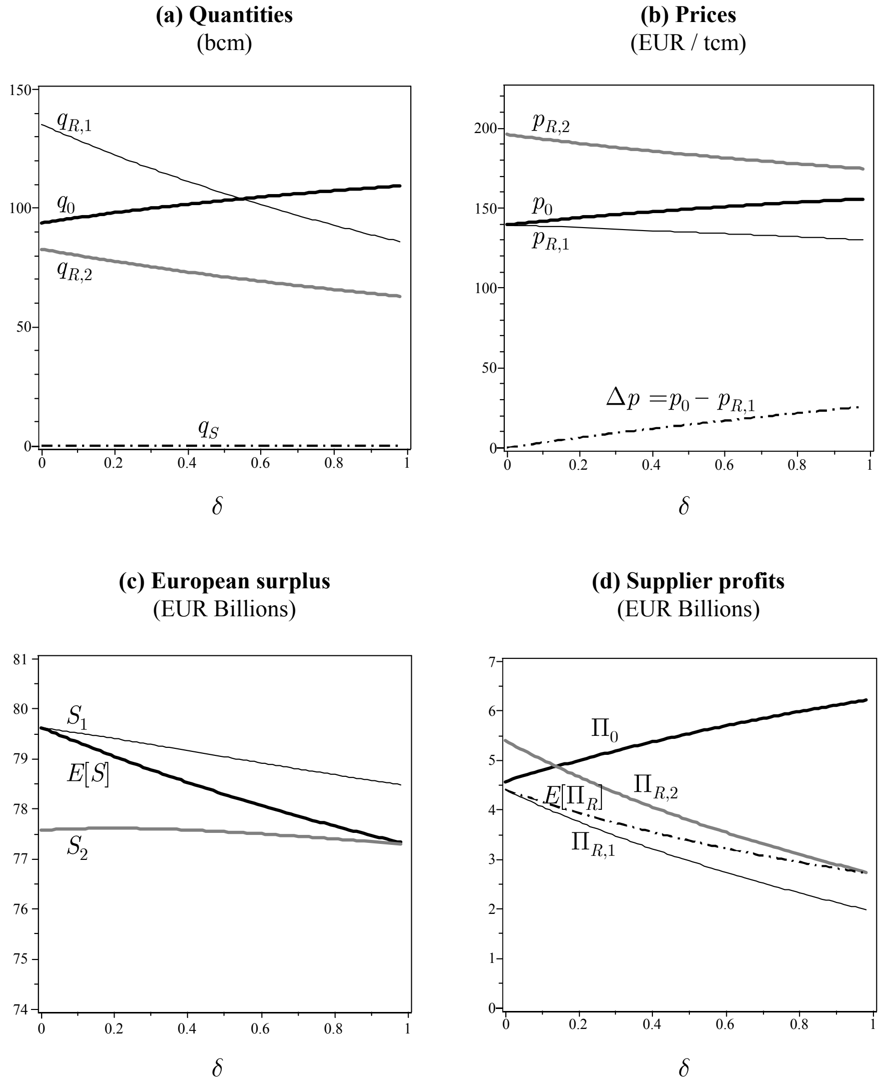
2.4.1 Effect of default probability δ on long-term gas import contracts and pay-offs

The top half of Figure 2.4 shows how quantities and prices vary as δ, the probability of Russian 'default', goes from 0 to 1. The graph also shows the discount $\Delta p = p_0 - p_{R,1}$ of long-term gas import contracts offered by Russia compared to contracts offered by the competitive fringe.

For $\delta = 0$, there is no risk and there is obviously no price difference between the contract with Russia and the contracts with the competitive fringe. The simulation shows that in this case, Europe buys $q_{R,1} = 135$ bcm per year from Russia and $q_0 = 94$ bcm per year from the other suppliers. This is not too far from the actual data in 2007 as cited by BP (2006-2008), which mentions 120 bcm per year from Russia and 95 bcm per year from other non-European import suppliers. Indeed, until recently, Russia was considered a reliable supplier, and so it is not surprising that the currently observed market quantities correspond to the case $\delta = 0$.

[23]The unconstrained optimal value of q_S tends to $-\infty$ as $\delta \to 0^+$. Hence, for sufficiently small δ the constraint is always binding: $q_S = 0$. Under certain conditions, the constraint is binding across the entire interval $\delta \in [0, 1]$. Under other conditions, there is a threshold $\delta = \delta_S$ above which the constraint is not binding. In the latter case the optimal q_S can be expressed analytically. However, the threshold δ_S itself cannot be calculated analytically, since it is the solution of a polynomial of order 7 in δ. Section 2.4 numerically computes the threshold levels δ_S under various conditions.

Figure 2.4: Base scenario: Relation between δ (horizontal axis) and quantities, prices, consumer surplus and supplier profits (vertical axis)*



*Note that $q_{R,2}$ is an 'annualized' amount in the sense of Footnote 17, which means that $q_{R,2}$ is extrapolated as if the crisis lasts the entire year instead of 4 months. On the other hand, S_2 and $\Pi_{R,2}$ do take into account that the crisis is limited to 4 months: they contain 4 months of crisis plus 8 months of non-crisis.

For $\delta > 0$ Russia becomes unreliable. When Russia 'defaults', it delivers only an annualized amount $q_{R,2}$ instead of $q_{R,1}$, at a higher price $p_{R,2}$ instead of the originally agreed long-term gas import contract price $p_{R,1}$. Panel (a) of Figure 2.4 shows that the quantity withheld would be around 40% and panel (b) shows that the resulting price increase would be around 40% as well. Although substantial, such a price increase is only a 2-sigma event over 3 trading days at gas hubs such as NBP (*National Balancing Point*, in the UK) when considering a typical daily volatility of 10%.[24]

As δ increases, Europe increases its volume q_0 of long-term gas import contracts with the competitive fringe, at a slowly increasing contract price p_0. Meanwhile, Europe procures a smaller volume $q_{R,1}$ with long-term contracts from Russia, even though Russia is obliged to give an increasing discount Δp to 'compensate' the risk for Europe. It is obvious why Russia would want to give the discount: as δ increases, there is a higher chance that Russia can charge the monopoly price $p_{R,2}$ in Stage 3 (by supplying only a quantity $q_{R,2}$ of gas). By giving a discount Δp, Russia can induce Europe to sign the long-term gas import contracts $q_{R,1}$ (despite the unreliability), which puts Europe in a vulnerable situation. For example, for $\delta = 20\%$, the Russian contractual discount is 6.3 EUR/tcm or roughly 4.5% of the price, which is consistent with the approximative equation (2.11) which predicts a discount of 4.4%. Despite the discount, Russia loses market share as δ increases and for $\delta > 57\%$ supply from the competitive fringe outstrips Russian supply. Clearly, Europe tries to make itself less dependent on Russia and therefore less vulnerable in the event of Russian withholding.

Panels (c) and (d) of Figure 2.4 show the effect on European surplus and on suppliers' profits, respectively. Recall that S_1 is the European surplus in Case 1 (Russia complies with long-term contracts) while S_2 is the European surplus in Case 2 (Russia defaults). $E[S]$ is the expected value of the European surplus. For $\delta = 0$ and $\delta = 1$, we obviously find $E[S] = S_1$ and $E[S] = S_2$, respectively. As δ increases, $E[S]$ decreases: despite the Russian discount and shifting supply mix, Russian unreliability causes a loss of expected European surplus. Panel (d) shows Russia's profits in Case 1 ($\Pi_{R,1}$), Case 2 ($\Pi_{R,2}$) and the expected value $E[\Pi_R]$, as well as the profits Π_0 obtained by the competitive fringe. Clearly, Russia's expected profits decrease monotonically with increasing δ: the negative impact of the Russian contract discount and loss of Russian market share is not sufficiently counterbalanced by Russia's increased likelihood of benefiting from a crisis. The only party gaining from increased unreliability is the competitive fringe. The competitive fringe profits Π_0 increase with increasing δ, because increased Russian unreliability allows them to sell a larger volume at a higher price.

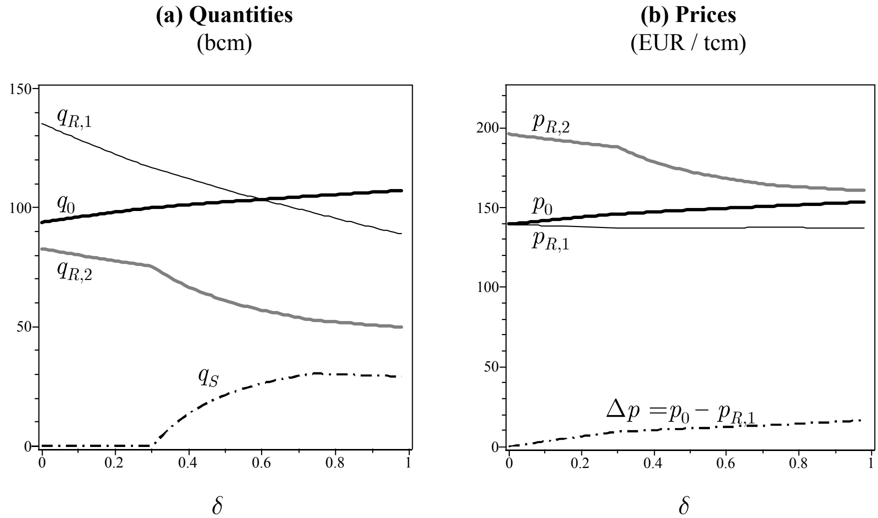
[24]In fact over the period 2003-2007 there have been 13 instances of (day-ahead) spot price increases of 40% or more between the closing prices of two consecutive trading days. However, we need to mention that only a very small share of gas volumes is traded on the gas hub spot markets, and liquidity is particularly low on the days with large swings.

The most important observation is that both Russia and Europe suffer when δ increases. Although δ is exogenous in our model, the results show that it would be attractive for both Europe and Russia to invest in a more reliable relationship, i.e. lower δ.

2.4.2 Conditions for strategic gas storage capacity investment $q_S > 0$

In the simulations of Figure 2.4, the value of q_S is always 0, meaning that it is never interesting for Europe to build any strategic gas storage capacity whatever the value of δ. The annual cost of storage capacity, $c_S = 50$ EUR per tcm per year, is too high compared to the potential gains. Figure 2.5 repeats the simulations with $c_S = 15$ EUR per tcm per year.[25] The result is identical to Figure 2.4 for $\delta < 30\%$. For $\delta \geq 30\%$, Russian unreliability is high enough to make investments in strategic gas storage capacity q_S competitive. As of that point, $p_{R,2}$ (Russia's potential 'monopoly price') drops significantly. As a result, Russia's market share loss compared to the competitive fringe slows down slightly, while its discount Δp flattens out.

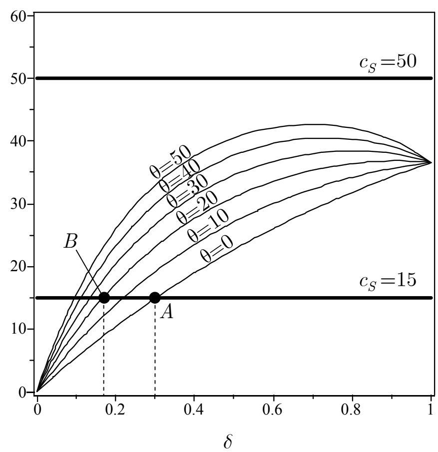
Figure 2.5: Scenario with reduced storage costs: Relation between δ (horizontal axis) and quantities and prices (vertical axis)



[25]In particular, if the storage site could be set up so that it can be used for seasonal arbitrage while the cushion gas serves as strategic storage, the cost of strategic storage would be significantly reduced. Typical ratios of total gas (working gas plus cushion gas) to working gas are 3-4 for aquifers and depleted reservoirs, hence our choice $c_S = 15$ instead of 50 EUR per tcm per year.

Besides lower storage capacity costs c_S, another factor that can encourage investments in strategic gas storage capacity, is risk aversion. Annex 2.D explains how our model can take into account European risk aversion, as measured by θ, the *coefficient of relative risk aversion*. Typical values of θ are 2 to 4 for financial assets and 10 to 15 when real assets are also included (Palsson, 1996). $\theta = 0$ corresponds to the risk-neutral case which we have been studying in this paper so far. Figure 2.6 covers different values of θ. For each value of θ, the graph contains a curve (as a function of δ) that shows the maximum value of c_S for which $q_S > 0$. There is storage investment $q_S > 0$ in the region of the (δ, c_S) space below each curve. One can see that for $\theta = 0$ and $c_S = 15$ EUR per tcm per year, the storage option becomes interesting for $\delta \geq 30\%$ (point A), which is obviously identical to what has been observed in Figure 2.5. If θ goes up to 20, then the threshold level comes down to $\delta = 17\%$ (point B). However, for $c_S = 50$ EUR per tcm per year, storage remains unattractive, unless $\theta \gg 50$, which is highly unrealistic.

Figure 2.6: Relation between δ (horizontal axis) and maximum value of c_S (in EUR per tcm per year) for which $q_S > 0$ (vertical axis), for different levels of risk aversion θ



2.5 Conclusions

The first research question of this paper is how Russian unreliability may impact the European gas market and how this affects European gas import decisions. Our numerical simulations show that it is not optimal for Russia to cut gas supplies to Europe completely during a crisis: rather, one can expect Russia to reduce its gas supplies by roughly 40%, thereby temporarily increasing gas prices by roughly 40%. More importantly, the analysis shows that not only Europe but also Russia suffers when Russia's probability of default δ increases, due to erosion of its price and market share. These results add weight to the conclusion that the Ukraine incidents probably were not aimed at exploiting monopoly profits from Europe. As observed in Footnote 2, more plausible explanations are a desire to obtain netback parity from neighboring countries, perhaps a prelude to raising prices closer to market levels in Russia itself. Quite possibly, the European perception of these crises as expressions of Russian market power have been harmful to the interests of both Europe and Russia.

The second research question of this paper is to what extent Europe should invest in strategic gas storage capacity to mitigate the effects of possible supply withholding by Russia. We find that strategic storage capacity is attractive for Europe only if Russian unreliability is high (δ of more than 30%) and storage capacity costs are reduced by a factor 3 to 4 compared to typical current cost levels. The threshold of 30% default probability is lowered when Europe is assumed to be risk averse.

The results of this paper are obtained using a partial equilibrium model of the market for long-term gas import contracts, with differentiated competition between one potentially unreliable 'dominant firm' (Russia) and a reliable 'competitive fringe' of other non-European import suppliers. Future research could examine the impact of the other suppliers becoming unreliable as well. Another possible extension is to turn our model into a repeated game. In such a game, δ could become endogenous as part of a mixed Russian strategy. Finally, the topic of this paper could be placed in a broader comparison of policy measures (import taxes, rationing, interruptible consumer contracts, etc.) that can be used to address gas import challenges.

2.6 Acknowledgements

The authors thank the participants at the 2nd Enerday conference in Dresden in April 2007, the 9th IAEE European Energy Conference in Florence in June 2007, a discussion session at the European Commission – DG TREN in December 2008, a CEPE seminar at ETH Zürich in May 2009, and an ESE Meeting at the Institute for Energy in Petten in December 2009, for their comments. We are also indebted to the editor and four anonymous referees for their constructive feedback.

2.A Annex: Optimal quantity and price for Russia in case of default

This annex explains equation (2.7). Russia's annualized profits during the crisis are:

$$\Pi_{R,crisis} = (p_{R,2} - c_R)q_{R,2} \tag{2.16}$$

where $p_{R,2}$ and $q_{R,2}$ follow the short-run demand curve (2.2) around the point $(q_D + q_0 + q_{R,1}; p^*)$:

$$p_{R,2} = p_{SR}(q_D + q_0 + q_S + q_{R,2}) \quad \text{with} \quad p_{SR}(q) = p^* + \beta_{SR} \cdot (q - q_{R,1} - q_0 - q_D) \tag{2.17}$$

To determine the optimal $q_{R,2}$ (and consequently, the optimal $p_{R,2}$) the derivative of (2.16) is used:

$$\begin{aligned} \frac{d\Pi_{R,crisis}}{dq_{R,2}} &= \frac{\partial \Pi_{R,crisis}}{\partial q_{R,2}} + \frac{\partial \Pi_{R,crisis}}{\partial p_{R,2}} \cdot \frac{dp_{R,2}}{dq_{R,2}} \\ &= \left[p_{SR}(q_D + q_0 + q_S + q_{R,2}) - c_R\right] \\ &\quad + \left[q_{R,2} \cdot p'_{SR}(q_D + q_0 + q_S + q_{R,2})\right] \end{aligned} \tag{2.18}$$

Setting (2.18)$= 0$ and solving together with (2.17), yields the monopoly quantity and price shown in equations (2.7). Strictly speaking, the Kuhn-Tucker conditions for constrained optimization with constraint $q_{R,2} \geq 0$ should be used. This constraint is ignored in the analytical presentation of Section 2.3, but in the numerical simulations of Section 2.4 it is taken into account, not only for $q_{R,2}$, but also for q_0, $q_{R,1}$ and q_S. Except for q_S, the constraint is never binding. Note that (2.7) is the well-known textbook expression for monopoly pricing with linear demand and constant marginal cost.

2.B Annex: Russian discount $\Delta p = p_0 - p_{R,1}$

This annex explains equation (2.11). We develop a first-order approximation for Δp around $\delta = 0$. In first order, the inverse demand functions (2.10) reduce to:

$$\begin{cases} p_{R,1} \approx \alpha(\phi + \delta) + \beta\left(\phi + \frac{3\delta}{4}(k-1) + \delta\right) q_{R,1} + \beta(\phi + \delta)\ q_0 \\ p_0 \approx \alpha\left(\phi + \frac{3\delta}{4}\right) + \beta\phi\ q_{R,1} + \beta\left(\phi + \frac{3\delta}{4}\right) q_0 \end{cases} \tag{2.19}$$

using the fact that $\phi/(1-\delta) \approx \phi + \delta$ in first order. Subtracting the first equation from the second, we obtain:

$$\Delta p \approx \frac{\delta}{4}\left[-3\beta k q_{R,1} - \alpha - \beta q_{R,1} - \beta q_0\right] \tag{2.20}$$

The inverse demand functions (2.10) describe the special case in which $\tau = 1$, $q_S = q_D = 0$ and $c_R = 0$. When we redo the exercise with those parameters included, equation (2.20) modifies only slightly:[26]

$$\Delta p \approx \frac{\delta\tau}{4}\left[-3\beta k q_{R,1} - \alpha - \beta q_{R,1} - \beta q_0 - \beta q_D + c_R\right] \tag{2.21}$$

which can be further simplified, using $p^* = \alpha + \beta \cdot (q_D + q_0 + q_{R,1})$ and the short-run equivalent of equation (2.26) – assuming $q^* \approx q_{cal}$ and $p^* \approx p_{cal}$. We obtain:

$$\Delta p \approx \frac{\delta\tau}{4}\left[-3\beta_{SR} q_{R,1} - p^* + c_R\right] \tag{2.22}$$

$$\approx \frac{\delta\tau}{4}\left[3\left(\frac{1}{|e_{SR}|}\frac{p^*}{q^*}\right) q_{R,1} - (p^* - c_R)\right] \tag{2.23}$$

After dividing by p^* we obtain:[27]

$$\frac{\Delta p}{p^*} \approx \frac{\delta\tau}{4}\left[\frac{3}{|e_{SR}|}\left(\frac{q_{R,1}}{q^*}\right) - \left(1 - \frac{c_R}{p^*}\right)\right] \tag{2.24}$$

Since we typically have:

$$\frac{3}{|e_{SR}|}\left(\frac{q_{R,1}}{q^*}\right) \gg 1 - \frac{c_R}{p^*} \tag{2.25}$$

equation (2.24) can be further approximated by equation (2.11).[28]

2.C Annex: Calibration of the model parameters for the numerical simulations

This section describes the numerical assumptions for the parameters used in our model, which are based on estimates from the literature.

Demand. The parameters α, β and β_{SR} are determined using elasticities from the literature, the 2007 baseline for volume, and the average price 2003-2007. β can be easily derived from a calibration point (p_{cal}, q_{cal}) and an elasticity value e:

$$\beta = \frac{1}{e}\frac{p_{cal}}{q_{cal}} \tag{2.26}$$

[26]Since Section 2.3.3 shows that $q_S = 0$ for small enough values of δ, we keep $q_S = 0$ in the derivation of equation (2.21).

[27]Note that Russia's discount decreases when its *rent margin* $(1 - c_R/p^*)$ increases. Naively, one might think that Russia's discount should *increase* in case of a higher rent margin, since a higher rent margin offers more financial room for discounting. However, a higher rent margin provides Russia with an incentive to withhold less in the event of a default, thus leading to a lower discount on the long-term contract price.

[28]With the calibration parameters used in the rest of the paper, the left-hand side of equation (2.25) is roughly 10 times larger than the right-hand side.

Long-run price elasticity of demand is taken equal to -0.93, following Golombek et al. (1998). Short-run price elasticity of demand is determined based on the very comprehensive literature survey of Dahl (1993). In Dahl (1993), the average short-run elasticity over the 15 studies that compute both short-run and long-run elasticities is -0.27.[29] In case this value seems large (in absolute terms), one should consider that we are ignoring any elasticity of European domestic supply: q_D is exogenous and fixed. The number -0.27 should therefore also include the effect of a non-zero elasticity of domestic supply. In addition, our model allows Russia to withhold supplies for 4 months at a time. This means that our short-run price elasticity of demand relates to a time frame of a few months, which makes the value of -0.27 seem reasonable. For the sake of safety, Annex 2.E performs a sensitivity analysis on the elasticities.

According to BP (2006-2008), total European gas consumption in 2007 was $q_{cal} = 494$ bcm per year. The average German border price as registered by the German government (BAFA, 2009) was $p_{cal} = 144$ EUR per tcm over the period 2003-2007. Based on these numbers and the above-mentioned elasticities, the resulting β, β_{SR} and, finally, α can be determined.

Costs. Total per-unit production and transportation costs are based on OME (2002), which shows costs curves for additional volumes of gas supply to the EU in 2010.[30] For Russia, marginal costs are assumed constant (c_R, see Section 2.2.2).[31] According to OME (2002), the long-run marginal cost for production in the Nadym-Pur-Taz region with transport through Ukraine (a combination which represents a very large share of current Russian exports) is \$2.8/MMBtu. The long-run marginal cost for production on the Yamal peninsula with transport through Belarus (a combination which represents large future potential for Russian exports) is also \$2.8/MMBtu. We there-

[29]Note that the average long-run elasticity of the same 15 studies was -0.99, which is in line with the -0.93 from Golombek et al. (1998).

[30]Supply to the EU-15 is used because this is the most realistic estimate of the cost of supply to an 'average' country in Europe. Using data for EU-27 would understate the costs for Russia.

[31]Our assumptions about marginal production costs are a slightly simplified version of Boots et al. (2004) and Golombek et al. (1995, 1998), who model marginal production costs as:

$$MC_{production} = a + bq + c\ln(1 - q/Q) \tag{2.27}$$

where q is production and Q is capacity (the third term makes production costs go up to infinity as soon as capacity is reached). The constants a, b, c and Q are determined for each supplier country separately. The exact values of a, b, c and Q are listed in Golombek et al. (1995), and are the same in the three studies Boots et al. (2004) and Golombek et al. (1995, 1998). For Russia however, b is 0, which is in line with the assumption of constant marginal cost used in this paper. Compared to Boots et al. (2004) and Golombek et al. (1995, 1998), our main assumptions are that (i) we leave out the non-linear capacity term $c\ln(1 - q/Q)$ for both Russia and the other suppliers (to keep our model analytically solvable), and (ii) we aggregate the production of the other suppliers and assume a linear marginal cost curve for the total.

fore choose c_R =\$2.8/MMBtu, or 107 EUR/tcm.[32] For the other suppliers, we assume that marginal costs are linearly increasing ($c_0 + d_0 q_0$, see Section 2.2.2). For c_0, we choose the cheapest source of gas imports to the EU-15 according to OME (2002): Algerian gas imported via the MedGaz pipeline at \$1.1/MMBtu. So $c_0 = 42$ EUR/tcm. The slope d_0 of the marginal cost curve is determined based on the slope of the cost curve for additional gas imports into the EU-15 on p. 14 of the OME (2002) report, excluding domestic European and Russian supplies. The cost curve for additional gas supplies starts at \$1.1/MMBtu with Algerian gas, and climbs up to \$3.0/MMBtu (Qatar LNG) to reach an additional volume of 70 bcm per year. We therefore choose $d_0 =$ (\$1.9/MMBtu)/70 bcm = 1.04 EUR/tcm/bcm. The annual gas storage capacity costs c_S are taken at the lower end of the range 50-70 EUR per tcm per year, mentioned by Mulder and Zwart (2006).

2.D Annex: Modeling European risk aversion

Equation (2.3) assumes risk-neutrality: Europe maximizes the expected value of European surplus S. The most straightforward way to introduce risk aversion is to assume instead that Europe maximizes the expected value of *a concave transformation of* European surplus. A theoretical justification for this transformation is to model European decision-making using the Stigler-Peltzman model[33] and assume that Europe maximizes a political support function M:

$$M = M(CS, \Pi_D, -G) \tag{2.28}$$

instead of maximizing just S as in equation (2.3). Assuming that different constituencies are treated identically, we can simplify the expression for M:

$$M = f(S) \quad \text{with} \quad S = CS + \Pi_D - G \tag{2.29}$$

Following Peltzman (1976), we have $f' > 0$ and $f'' < 0$. If Europe's objective is to maximize the expected value of M, then the concavity of f leads to risk averse behavior.[34] To make the degree of risk aversion explicit, we choose a particular functional form for $f(\cdot)$, namely a function that yields *constant relative risk aversion (CRRA)*:

$$f(x) = u_\theta(x) = \frac{x^{1-\theta}}{1-\theta} \tag{2.30}$$

[32]USD/EUR conversions are done at the average exchange rate for the year 2002 in which the OME estimates were made (1 USD = 1.06 EUR).

[33]See Stigler (1971) and Peltzman (1976).

[34]In the true sense of the political support function, this would only model the risk aversion of the politicians. However, we shall assume that risk aversion of consumers and domestic producers is also included in f.

θ is the coefficient of relative risk aversion.[35]

To determine the European inverse demand functions for long-term gas import contracts (Section 2.3.2), we need to maximize $E[M] = E[f(S)]$ instead of $E[S]$. Since $f' > 0$, the maximization of $E[f(S)]$ is equivalent to the maximization of $C = f^{-1}(E[f(S)])$. Let us now define $\epsilon = S_1 - S_2 = \tau\Delta S$, i.e. the potential 'downside' of the deal with Russia. We need to maximize:

$$\begin{aligned} C &= u_\theta^{-1}((1-\delta)u_\theta(S_1) + \delta u_\theta(S_2)) \\ &= u_\theta^{-1}((1-\delta)u_\theta(S_1) + \delta u_\theta(S_1 - \epsilon)) \\ &\approx S_1 - \delta\epsilon - \frac{1}{2}\delta(1-\delta)\theta\frac{\epsilon^2}{S_1} + \text{higher-order terms in } \theta \text{ and } \epsilon \end{aligned} \tag{2.31}$$

in which the last step results from a Taylor expansion around $\theta = 0$ and $\epsilon = 0$. By defining σ:

$$\sigma = \delta + \frac{1}{2}\delta(1-\delta)\theta A \quad \text{with} \quad A = \frac{S_1 - S_2}{S_1} \tag{2.32}$$

we can rewrite equation (2.31) as:

$$C \approx S_1 - \sigma\epsilon = S_1 - \sigma\tau\Delta S = (1-\sigma)S_1 + \sigma S_2 \tag{2.33}$$

Equation (2.33) is completely equivalent to equation (2.9) but with δ replaced by σ. The analytical results of Section 2.3 therefore remain valid for a risk averse Europe, provided we replace δ by σ in equations modeling Europe's decisions. This approach has the advantage of having a very intuitive interpretation: in equation (2.33), Europe 'perceives' a Russian default probability σ, which is different from the 'real' default probability δ. Europe's risk aversion is thus modeled as a higher perceived default probability.[36] For the cases $\delta = 0$ and $\delta = 1$, there is no uncertainty, so $\sigma = \delta$. The more uncertainty (i.e. the closer to $\delta = 0.5$), the larger the difference between σ and δ.

2.E Annex: Sensitivity analysis on elasticities

The analyses in Section 2.4 assume that long-run price elasticity of European demand for imported gas is -0.93, and that the short-run elasticity is -0.27.

[35] The precise level of risk aversion θ is a parameter to the simulations. Section 2.4.2 contains simulations for different values of θ.

[36] The caveat is that σ is actually not a constant, but depends on A, which is the proportional 'deviation' between S_1 and S_2. However, in addition to the approximations already made in the derivation of equation (2.33), σ is treated as a constant in subsequent analyses. σ therefore changes with δ as defined in equation (2.32), but does not depend on S_1 and S_2, because a fixed A is taken. The value of A is chosen to be 0.05 in the numerical simulations. This is about twice the value observed in Figure 2.4 (Panel (c)) in order to take into account the fact that the absolute value of European surplus is probably lower due to the costs of domestic production, which are ignored in the computation of $E[S]$.

In other words, the ratio k between long-run and short-run elasticities is 3.4 ($k = e/e_{SR} = \beta_{SR}/\beta$). The empirical evidence on this ratio k, however, is scattered. Neuhoff and Hirschhausen (2005) report k values of 4 to 5 for industrial demand, and 5 up to 10 for residential and commercial demand. In this section we shall briefly review how the results of Section 2.4.1 change when we double k from 3.4 to 6.8.

First, let us double k by reducing the *short-run* elasticity to half its original value, i.e. we set $e_{SR} = -0.14$ instead of -0.27. The slope of the short-run demand curve β_{SR} becomes twice as steep. We keep the *long-run* elasticity e of gas demand constant, meaning that β is held constant. The simulation results are shown in Figure 2.7 and are graphically similar to those in Figure 2.4, but the effects are more pronounced: Russia's monopolistic price $p_{R,2}$ is obviously higher, its market share declines faster as a function of δ, while its discount increases more steeply, as is expected based on equation (2.11). However, still no storage is built when $c_S = 50$ EUR per tcm per year.

Secondly, let us double k by increasing the *long-run* elasticity to twice its original value, while keeping $e_{SR} = -0.27$.[37] The results are shown in Figure 2.8. Because of the very high long-run elasticity, the volume of long-term gas import contracts with Russia for $\delta = 0$ exceeds quite significantly the actual 2007 volume: 197 bcm per year versus the actual 120 bcm per year. However, the large value of k brings the volumes down quite rapidly as δ increases. As of $\delta \approx 0.5$ the results become graphically very similar to the base scenario in Figure 2.4.

All in all, the conclusions are fairly robust vis-à-vis changes in k.

[37]While this seems quite an extreme stress-test, one can argue that it is justified because long-run elasticity of domestic supply q_D is ignored in this paper. Furthermore, note that a change in β requires also a change in α to make sure that the new demand curve still passes through the same calibration point.

Figure 2.7: Scenario with halved short-run elasticity: Relation between δ (horizontal axis) and quantities and prices (vertical axis)

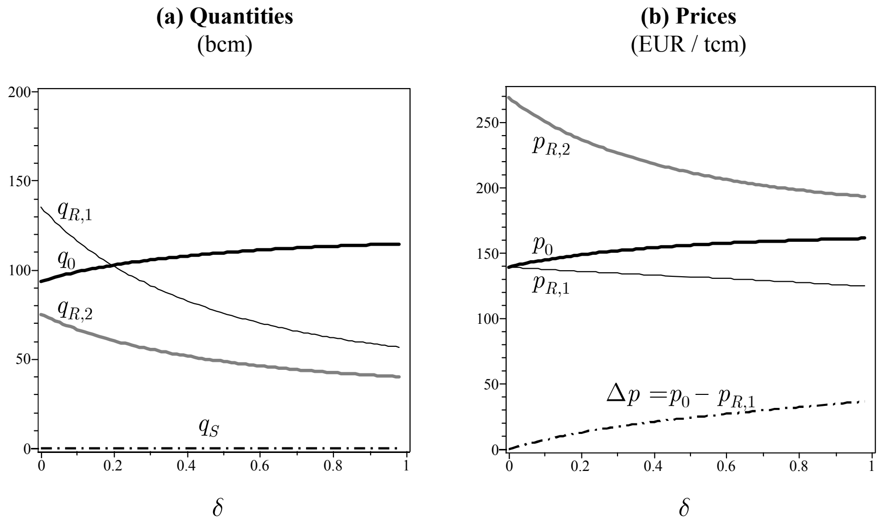
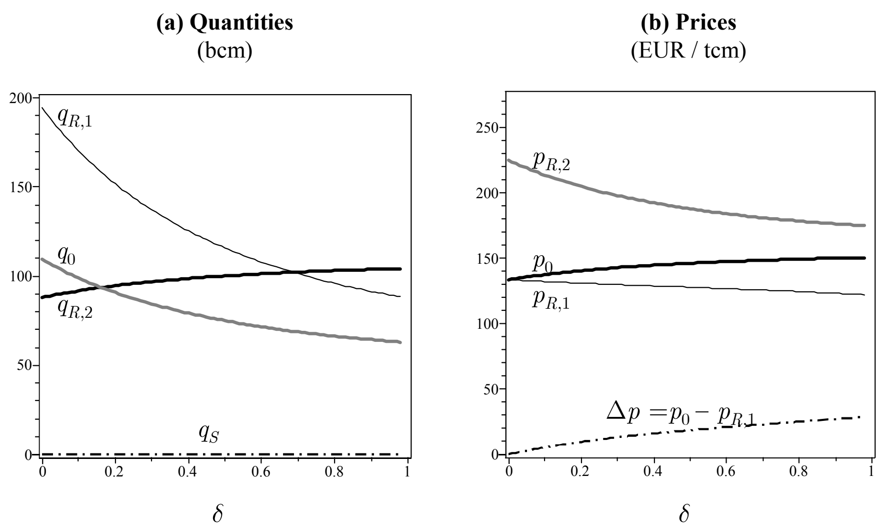


Figure 2.8: Scenario with double long-run elasticity: Relation between δ (horizontal axis) and quantities and prices (vertical axis)



Chapter 3

Political Economy of Resource Taxation: Why Do Dictators Charge More?

3.1 Introduction

There is a wide variation between the fiscal systems for resource extraction in different countries. This becomes clear when looking for example at the government take of petroleum extraction profits. *Government take* is the total share of resource extraction profits that the producing firm needs to transfer – in the form of royalties, production-sharing agreements, corporate taxes, etc. – to the government of the country in which it operates. For example, oil companies that produce oil in the US or the UK typically face a government take of around 40%. Oil companies in Indonesia or Myanmar, by contrast, face a government take of more than 70%. This paper aims to examine the underlying causes of these large differences. We construct a model of a resource-rich country, which incorporates elements from both resource economics and political economy theory. The results of the model are tested empirically on a data set of 77 countries. Although the theoretical model is applicable to a wide range of resources (petroleum, metals, precious minerals, etc.), the empirical work is focused on oil.

There is a large body of literature on resource taxation within the Hotelling (1931) tradition of dynamic partial equilibrium models of the resource sector. Well-known studies such as Burness (1976), Dasgupta and Heal (1979), Dasgupta et al. (1980) and Heaps (1985) analyze the impact of different forms of taxation on intertemporal allocation of a given amount of an exhaustible resource. Other models not dealing with taxation, such as Peterson (1978) and Pindyck (1978) endogenize the amount of reserves by studying the exploration decision of firms. Conrad and Hool (1981), Campbell and Lindner (1985) and others integrate both taxes and exploration in the same model. The model developed in our paper also includes both taxation and exploration, but on a more stylized level. In particular, we do not use a continuous-time model, but rather a two-period model, which is less common in the resource economics literature, although it has been used by e.g. Stiglitz (1976) and Solow and Wan (1976). Furthermore, our model does not distinguish between different types of taxation. Instead we represent the entire fiscal system using a single profit tax, which, according to Dasgupta and Heal (1979), is by far the most widespread type of tax.

The above-mentioned literature on resource extraction and taxation generally takes the perspective of the producing firm, and is less concerned with the economics of resource-rich countries. This is the topic of a separate strand of literature, which mainly took off after the oil price rises of the 1970s and subsequent drops in the oil price during the 1980s. Most of the literature focuses on the development aspect and relates to what Auty (1993) calls the *resource curse*: although natural resources seem desirable at first sight, abundance of natural resources can distort a country's economy and make the country on balance worse off. A recent study by van der Ploeg (2011) provides a survey of research results. In many analyses, it turns out that economies with abundant natural resources have tended to grow less rapidly than natural-resource-scarce economies (see e.g. Sachs and Warner, 1995). One of the most frequently cited underlying mechanisms is the well-known *Dutch disease*, as described for example by van Wijnbergen (1984) and Krugman (1987) and many other authors. The Dutch disease refers to the effect that abundance of natural resources may shift production factors away from other sectors (in particular: non-resource-based traded goods), thereby reducing "learning-by-doing" in those sectors, which ultimately leads to lower overall productivity growth and reduced long-term welfare. Our paper does not consider these macroeconomic effects but instead is focused on the microeconomic aspects of a resource-rich country.

Indeed, besides the Dutch disease, other factors contribute to the resource curse (Davis, 1995): misguided protection of industry, overly optimistic resource price and export revenue forecasts from governments (leading to sticky overspending, and recession when prices drop), and rent-seeking. Logically, several authors therefore studied the political economy of the resource-abundant states, which will also be our approach. Karl (1997) argues that *"dependence on petroleum revenues produces a distinctive type of institutional setting, the petro-state, which encourages the political distribution of rents. Such a state is characterized by fiscal reliance on petrodollars, which expands state jurisdiction and weakens authority as other extractive capabilities wither. As a result,* *when faced with competing pressures, state officials become habituated to relying on the progressive substitution of public spending for statecraft, thereby further weakening state capacity".* In a similar vein, Auty and Gelb (2001) use an extension of Lal's (1995) typology of governments and argue that resource-abundant countries are more likely to have factional or predatory governments, which tend to distort the economy in the pursuit of rents. As a result, the economy does not follow a competitive industrialization model, but falls into a "staple trap".

The political economy models of Karl (1997) and Auty and Gelb (2001) remain qualitative (unformalized), unlike models about rent seeking, such as the model of Torvik (2002). The latter develops a simple model in which resource-abundance increases the number of entrepreneurs engaged in rent seeking and reduces the number of entrepreneurs running productive firms. Torvik (2002) assumes a 'demand externality': in his model, the reduction in the number of productive firms leads to less income and demand and, thus, lower profits for the remaining productive firms. The resulting drop in income is higher than the increase in income from the natural resource. Robinson et al. (2006) build the first explicitly political, formal model to explain the resource curse. In a two-period probabilistic voting model, they explain public sector clientelism, paid from natural resource rents, as a way to influence election outcomes. The model integrates an endogenous choice of natural resource extraction rate, and studies the effects of resource booms on resource extraction, and on factor misallocation. Deviation from the optimal extraction path as a result of political economy effects is also studied by van der Ploeg (2008), who argues that the fact that resource-rich countries often save less than the marginal resource rents can be due to either anticipation of more favorable conditions (lower extraction costs, higher resource prices, higher interest income) or political distortions in fractionalized societies with imperfect property rights.

Although our model is less concerned with the resource curse, we use a similar formal approach like Robinson et al. (2006). Our model is slightly more sophisticated on the resource economics side, given that it includes resource taxation and endogenizes firms' exploration and production decisions, but less explicit on the political economy side, as will be explained in the next section. The reason for this is that the inclusion of non-democracies in our model is essential (since many resource-rich states are non-democratic) but, as pointed out by Acemoglu and Robinson (2006, p.89), the literature on decision-making in non-democracies is much smaller than the vast literature on democracies. For instance, standard political economy literature such as Persson and Tabellini (2000) and Besley (2006) is focused exclusively on democratic decision-making. Our model is inspired by the stylized models developed by Niskanen (1997), who compares fiscal decisions of autocratic, democratic, and optimal governments. To model the government characteristics of resource-abundant countries, we use and formalize the above-mentioned government typology developed by Lal (1995).

One of the main contributions of our paper is that it complements the theoretical model with an empirical analysis of government take in 77 oil-producing countries. The empirical methodology draws from Persson and Tabellini (2003), who compare fiscal and other policies across democracies, and Mulligan et al. (2004), who compare policies of democracies and non-democracies. Unlike Mulligan et al. (2004), our empirical investigation checks the robustness of OLS estimates through IV and panel data estimation.

This paper is structured as follows. First, Section 3.2 develops a stylized theoretical model of the political economy of taxation of resource extraction. Next, Section 3.3 describes our data set, defines our regression analysis, and presents the regression results. Finally, Section 3.4 summarizes the main conclusions from the analysis, and suggests potential directions for further research.

3.2 A political economy model of resource taxation

3.2.1 Structure of the game

We study a model with two periods, $t = 1, 2$. As we will see below, this allows for smooth integration with median-voter or probabilistic voting equilibria from political economy theory. The length of each period can be thought of as 5 to 10 years.

In each period t, the government first sets the tax rate τ_t ($0 \leq \tau_t \leq 1$), after which a competitive firm decides on the level of exploration activity and production of the resource. The tax τ_t is assumed to be implemented as a simple proportional profit tax on resource extraction profits. Initial reserves of the resource are denoted with R_0. Exploration activity in each period t adds ΔR_t to the reserves base due to new discoveries. Production from reserves in each period t is denoted with q_t. Production is sold under a price-taking assumption on the world market, at price p_t per unit, with $p_2 > p_1$. Hence, the price increases between period 1 and period 2, which is consistent with the Hotelling (1931) concept.

Between period 1 and period 2, there is a stochastic event that may result in a change of government. In the case of a *democratic* government, this event can simply be thought of as an election. Assuming that the election leads to a typical median-voter or probabilistic voting equilibrium as in Downs (1957) or Hotelling (1929), where one votes in only one dimension (here the tax rate) and voter preferences are single-peaked, the policies proposed by competing politicians before the election will converge and chances of getting elected will become identical for all candidates. E.g. in a typical median-voter equilibrium, the two competing candidates will propose identical policies and each have a 50% chance of getting elected. Rather than explicitly modelling the election game, we will simply assume that there is an exogenous probability P_e that the government in office at $t = 1$ gets reelected to period $t = 2$.[1] In the case of a *non-democratic* government, there is no election, but we assume that there is a chance that a revolution takes place between period 1 and period 2. The probability that the non-democratic government in office at $t = 1$ is still in office at $t = 2$ is exogenously given by P_{nr}, the probability that no revolution takes place. In this way, democracy and non-democracy can be treated within the same conceptual framework. This is similar to some of the results of Robinson et al. (2006). In their formal model of political economy of resource-abundant states, they derive explicit analytical expressions for reelection and non-revolution probabilities and find that there are striking similarities between the two. Interestingly, as a corollary, Robinson et al. (2006) make an extrapolation to the case of endogenous democratic institutions. They conclude that politicians will prefer democracy if, in our notation, $P_e > P_{nr}$, i.e. if the reelection probability in case the country is democratic, is larger than the non-revolution probability in case the country is non-democratic. This may be the case for instance if society is structured in such a way that a very large winning coalition would be needed in order to sustain a dictatorship. The point about endogenous democratic institutions is explored in detail by Acemoglu and Robinson (2006). For instance, in their static model of democratization (Acemoglu and Robinson, 2006, Section 5.6) the non-democratic elites choose whether or not to establish democracy by weighing the possibility of a revolution against the likely income redistribution in case of democratization. To keep our argument focused, we do not explicitly model the democratization game in this paper. However, based on the argument by Robinson et al. (2006) we assume that in any given non-democratic country, P_{nr} is larger than the reelection probability P_e that would apply if the same country was a democracy. Indeed, in a dictatorship, if the reelection probability P_e after democratization were larger than the non-revolution probability P_{nr} during the dictatorship, then the country's dictator would already have chosen for a transition to democracy, hence P_{nr} would not be applicable anymore. So, whenever P_{nr} is applicable, we have $P_{nr} > P_e$. One could argue that a symmetric reasoning would hold for a democracy, which would therefore only exist if $P_e > P_{nr}$. However, it seems more difficult for a democratic government to decide to make the transition to a dictatorship, than it is for a dictator to facilitate democratization. In this paper, we therefore use the assumption that, all else equal, we have $P_{nr} > P_e$.

[1]Note that this also holds for median-voter models. Indeed, in typical median-voter models, it is assumed that if the policies of the two competing politicians are identical, then they each have a 50% probability of winning. See e.g. Persson and Tabellini (2000).

3.2.2 Firm behavior

In each period t, firms respond to the tax rate τ_t by choosing their activity level of exploration and production. Let us first study the production decision regarding existing reserves R_0, i.e. reserves for which no further exploration is needed. We assume these reserves are in the hands of a competitive firm, which acts as a price-taker on the world market. Marginal production costs are assumed to be constant at a level of c per unit. In a Hotelling (1931) world, and assuming no taxation, prices p_1 and p_2 would be such that the firm would be indifferent between producing its reserves at $t = 1$ or leaving them in the ground for production at $t = 2$. The condition for this is $p_1 - c = \delta_f(p_2 - c)$, with δ_f the appropriate discount factor for the firm. In such a world, first the countries with the cheapest production costs (lowest c) would produce their reserves, and more expensive countries would only produce after the cheapest resources have been depleted. Looking for example at the large disparities in oil production costs of currently producing countries[2], this is clearly not a realistic view: many inframarginal low-cost countries are producing simultaneously with marginal high-cost countries. For the purpose of our analysis, we shall assume that the country under consideration is an inframarginal producer, such that:

$$p_1 - c > \delta(p_2 - c) \tag{3.1}$$

In the absence of taxes, the competitive firm would extract all of R_0 in period 1. With profit taxes τ_t, the firm's profits per unit of production change from $(p_t - c)$ to $(1 - \tau_t)(p_t - c)$. If taxes were constant over time ($\tau_1 = \tau_2$), this would not affect the production path, a result also obtained in a more general context by Dasgupta and Heal (1979). We shall assume that p_1, p_2 and δ are such that production of R_0 will remain unaffected even if $\tau_1 \neq \tau_2$. This could be the case e.g. if p_1 is not very different from p_2, if the country is a highly inframarginal producer, or if the uncertainty caused by the potential change of government between periods 1 and 2 leads to a very high corporate discount rate (i.e. very low discount factor δ_f). Hence, the existing reserves R_0 will be fully extracted in period 1 and sold on the world market at p_1.[3]

Let us now turn to the exploration decision. As in Pindyck (1978), we assume that additions to reserves, for given effort level, decrease with cumulative

[2]See e.g. Aguilera et al. (2009).

[3]Obviously, complete extraction of R_0 within one period may actually be unrealistic from a technical point of view, at least for the case of oil. From a technical point of view, if fast extraction is desired, one would rather expect extraction rates to increase quickly (within the first few years of operation of the field) up to a peak production level, which is maintained for a relatively short period and after which a more or less exponential decline follows. Clear evidence of this can be found by looking at historic production profiles of already depleted fields, such as the production profiles provided by the Norwegian Petroleum Directorate (2011). Such a highly non-linear production profile obviously cannot be modelled in a two-period model. The most important feature of this non-linear production profile, however, is that it is strongly weighted towards the first few years. In our two-period model, this essential feature is best captured by assuming that all of R_0 is produced in period 1.

discoveries. Put otherwise, the marginal exploration effort required for a given amount of additional discovery increases with cumulative discoveries. For the sake of analytical convenience, we assume that marginal exploration effort is a linearly increasing function of cumulative discoveries. If we identify exploration effort with exploration cost, then we have the following expression for marginal exploration cost w (i.e. exploration cost required for an additional unit of discovered reserves) as a function of cumulative discovered reserves:

$$w = w_0 + \phi \sum \Delta R_t \tag{3.2}$$

with $\phi > 0$ the slope of the curve, and $\sum \Delta R_t$ the sum of discoveries ΔR_t up to and including the period under consideration. We assume $w_0 < p_2 - c$, i.e. there is a net societal benefit in doing some exploration, maybe not in the near future (period 1) but at least in the more distant future (period 2). It is convenient to express ϕ as a function of *undiscovered reserves* $\bar{R}$. Undiscovered reserves, for instance in the case of petroleum, are *"quantities of undiscovered conventional oil, gas, and natural-gas liquids that have the potential to be added to reserves (proved and inferred) in some specified future time span"* (Schmoker and Klett, 2000). In our case therefore, undiscovered reserves $\bar{R}$ correspond to those reserves that can be profitably discovered and produced by the end of period 2, hence $\bar{R}$ is such that:[4]

$$w_0 + \phi \bar{R} = p_2 - c \tag{3.3}$$

Equation (3.2) can then be rewritten as a function of $\bar{R}$ instead of ϕ:

$$w = w_0 + \frac{p_2 - c - w_0}{\bar{R}} \sum \Delta R_t \tag{3.4}$$

How much of the undiscovered reservers $\bar{R}$ will be explored for and discovered in periods 1 and 2? We denote these discoveries by $\Delta R_t, t = 1, 2$. Let us first consider period 1. The government sets a tax rate τ_1 and a competitive firm responds by exploring for and discovering ΔR_1. Following the above-mentioned argument for R_0, all of the newly discovered reserves ΔR_1 in period 1 will be extracted in period 1, leading to a taxable production profit of $(p_1 - c)\Delta R_1$.[5] The competitive firm would set its exploration decision so as

[4]Note that the assessments of undiscovered reserves made by e.g. USGS (2000) are based on geology and do not consider full economics of exploration and production, see Schmoker and Klett (2000). Inclusion of petroleum volumes in the undiscovered reserves is based on a minimum 'cut-off' field size, which is chosen by taking into account the 'forecast span' of 30 years. In particular, this means that undiscovered reserves $\bar{R}$ can be safely assumed to be independent of tax rates, a point that will become important in the empirical analysis in Section 3.3. This also explains why we do not include a tax rate in equation (3.3).

[5]Note that we do not deduct exploration costs from the taxable profit. This is in line with typical exploration and production contracts. More precisely, only the successful exploration efforts can typically be deducted from profits, as in a case examined by Campbell and Lindner (1985). In our model, this could be implemented as a fixed tax deduction per unit

to equate marginal exploration cost and benefit:

$$(1-\tau_1)(p_1-c) = w_0 + \frac{p_2-c-w_0}{\bar{R}}\Delta R_1 \tag{3.5}$$

hence:

$$\Delta R_1 = \frac{(1-\tau_1)(p_1-c)-w_0}{p_2-c-w_0}\bar{R} \tag{3.6}$$

The decision in period 2 is analogous, but we need to take into account that an amount ΔR_1 has already been discovered and produced. Equating marginal exploration cost and benefit, we get

$$(1-\tau_2)(p_2-c) = w_0 + \frac{p_2-c-w_0}{\bar{R}}(\Delta R_1 + \Delta R_2) \tag{3.7}$$

hence

$$\Delta R_2 = \left(1-\tau_2\frac{p_2-c}{p_2-c-w_0}\right)\bar{R} - \Delta R_1 \tag{3.8}$$

All of ΔR_2 will obviously be extracted in period 2. The total quantities produced in periods 1 and 2 are given by:

$$q_1 = R_0 + \Delta R_1 \tag{3.9}$$
$$q_2 = \Delta R_2 \tag{3.10}$$

3.2.3 Government objective

Government revenues per period are given by $(p_t - c)\tau_t q_t$. Besides direct revenues, we assume that the government derives a strategic benefit of σ per unit of resource production, resulting from such factors as reduced import dependence or diplomatic leverage. Total government benefits per period are thus given by:

$$G_t = [(p_t-c)\tau_t + \sigma]q_t \tag{3.11}$$

with $\sigma \geq 0$. We will refer to G_t simply as 'government revenues'.

As mentioned in the introduction, we use the government typology developed by Lal (1995) and applied to resource-abundant states by Auty and Gelb (2001). Governments are classified along two dimensions: autonomy and benevolence. *Autonomy* refers to whether a government decides on its own (i.e. a dictatorship) as opposed to having to take into account other parties. For the sake of clarity, we will call this dimension *autocracy* instead of *autonomy*. In our formal model, we describe this dimension using the variable α, with $\alpha = 1$ implying dictatorship and $\alpha = 0$ a multiparty (or democratic)

of production. Indeed, while marginal exploration costs increase with cumulative discoveries, the amount of successful – i.e. tax-deductible – exploration per unit of production is likely to stay more or less the same. Since the inclusion of such a fixed tax deduction per unit of production would make the model mathematically more complicated without changing any insights, we do not include it.

government. *Benevolence* refers to whether a government maximizes social welfare, as opposed to maximizing its own private objective function. In our formal model, we describe this dimension using the variable β, with $\beta = 1$ referring to a benevolent government and $\beta = 0$ referring to a predatory government, which maximizes its own private objective function. In particular, the typology of Lal (1995) implies that there are multiparty (seemingly democratic) governments that are maximizing the welfare of a small fraction of the population ($\alpha = 0$, $\beta = 0$), while there exist autocratic or oligarchic governments that may be relatively benevolent. An example of the former is India, an example of the latter is Singapore.[6]

Assuming that the resource firm is foreign and that resource taxes τ_t do not introduce any distortions elsewhere in the economy, a *benevolent* government ($\beta = 1$) – i.e. a government that seeks to maximize social welfare – would maximize in period 1 the discounted sum of government revenues:

$$W_{\text{benevolent}} = G_1 + \delta_g G_2 \tag{3.12}$$

with δ_g the appropriate discount factor of the government. A *predatory* government ($\beta = 0$), i.e. a government in which officials seek to maximize their own profit, would have a different objective function. Assuming that government officials are able to (mis)appropriate a fixed share of government revenues, they would however still try to maximize government revenues, just like the benevolent government.[7] However, the officials will take into account that they only get access to the tax revenues of period 2 if they are still in office at that time. Assuming risk-neutral officials, the predatory government would therefore maximize $W_{\text{predatory}}$:

$$W_{\text{predatory}} = G_1 + P\delta_g G_2 \tag{3.13}$$

with P the probability that the period-1 government is still in office in period 2. Combining (3.12) and (3.13), a generic government maximizes:

$$\begin{aligned} W &= \beta W_{\text{benevolent}} + (1-\beta) W_{\text{predatory}} &\quad (3.14)\\ &= G_1 + [\beta + (1-\beta)P]\delta_g G_2 &\quad (3.15) \end{aligned}$$

As discussed in the Section 3.2.1, a dictatorship ($\alpha = 1$) would have $P = P_{nr}$, while a democracy would have $P = P_e$. Combining both, we find:

$$P = \alpha P_{nr} + (1-\alpha) P_e \tag{3.16}$$

and hence

$$\begin{aligned} W &= G_1 + [\beta + (1-\beta)(\alpha P_{nr} + (1-\alpha)P_e)]\delta_g G_2 &\quad (3.17)\\ &\equiv G_1 + \Psi G_2 &\quad (3.18) \end{aligned}$$

[6] Our typology is thus different from Niskanen (1997), who distinguishes only three types of government: autocratic, democratic, and optimal. In our model, optimality (a proxy for benevolence) is orthogonal to the democracy dimension.

[7] Maximization of tax revenues is also assumed by Niskanen (1997), for the case of autocratic governments.

with Ψ defined as

$$\Psi = [\beta + (1-\beta)(\alpha P_{nr} + (1-\alpha)P_e)]\delta_g \tag{3.19}$$

In reality, the government of the country may not be fully benevolent or fully predatory, but somewhere in between the two extremes, and likewise for the level of autocracy. In the remainder of the paper we assume that government decisions in period 1 are determined by equation (3.17), even in the case of continuously measured α and β.

As a final note on government incentives, it should be pointed out that in an ideal case the government would attempt to distinguish its tax rate for existing versus newly discovered reserves. However, such distinction may be difficult due to government credibility issues, so we do not consider it here.

3.2.4 Theoretical determinants of government take

We now derive the optimal government tax decisions τ_t using backward induction. In period 2, government maximizes G_2, given by equation (3.11). Using (3.10) and (3.8), G_2 can be expressed as a concave quadratic function of τ_2:

$$G_2 = [(p_2 - c)\tau_2 + \sigma]\left[\left(1 - \tau_2 \frac{p_2 - c}{p_2 - c - w_0}\right)\bar{R} - \Delta R_1\right] \tag{3.20}$$

Let $\hat{\tau}_2$ be the value of τ_2 that corresponds to the peak in G_2. Assuming $0 \leq \hat{\tau}_2 \leq 1$,[8] the optimal period-2 tax rate τ_2 is given by:

$$\hat{\tau}_2 = \frac{1}{2}\frac{(p_2 - c - w_0)(\bar{R} - \Delta R_1) - \sigma\bar{R}}{(p_2 - c)\bar{R}} \tag{3.21}$$

and the corresponding government revenues:

$$\hat{G}_2 = \frac{1}{4}\frac{[(p_2 - c - w_0)(\bar{R} - \Delta R_1) + \sigma\bar{R}]^2}{(p_2 - c - w_0)\bar{R}} \tag{3.22}$$

with ΔR_1 from equation (3.6). Substituting (3.22) and (3.11) into (3.18), we can find W as a concave quadratic function of τ_1.[9] Let $\hat{\tau}_1$ be the value of τ_1

[8] This will be the case for typical values of the input parameters. However, there may be realistic parameter configurations which would cause the peak value of the quadratic function G_2 to be outside of the natural interval for tax rates $[0, 1]$. Depending on assumptions, this would most likely result in τ_2 being bounded at 0 or 1. A detailed analysis of the different cases would complicate our explanation, and is not really needed for the purpose of this paper. Indeed, we are mainly interested in understanding the general direction of the impact of various parameters on the tax rates in order to have a theoretical intuition on which to build the empirical testing in Section 3.3. In this context it therefore seems acceptable to simply analyze the peak value of G_2, without applying the constraints on τ_2. The same principle will be applied to τ_1.

[9] We do not show W here, because the expression is long and not very insightful.

that corresponds to the peak in W. Again assuming $0 \leq \hat{\tau}_1 \leq 1$, the optimal period-1 tax rate τ_1 is given by:

$$\hat{\tau}_1 = \frac{2(p_1 - c - w_0 - \sigma) + (p_2 - p_1 + \sigma)\Psi + 2(p_2 - c - w_0)\frac{R_0}{\bar{R}}}{(p_1 - c)(4 - \Psi)} \tag{3.23}$$

Within this model, $\hat{\tau}_1$ from equation (3.23) is the most appropriate measure of current taxation on resource extraction. In the remainder of this section we will analyze the determinants of $\hat{\tau}_1$ in more detail. Taking into account that $p_2 > p_1 > c$; $p_2 > c + w_0$; $\sigma \geq 0$ and $0 \leq \Psi \leq 1$, we can easily see that:

$$\begin{aligned}
\frac{\partial \hat{\tau}_1}{\partial R_0} &= \frac{2(p_2 - c - w_0)}{(p_1 - c)(4 - \Psi)\bar{R}} && > 0 \\
\frac{\partial \hat{\tau}_1}{\partial \bar{R}} &= -\frac{2(p_2 - c - w_0)R_0}{(p_1 - c)(4 - \Psi)\bar{R}^2} && < 0 \\
\frac{\partial \hat{\tau}_1}{\partial \sigma} &= -\frac{2 - \Psi}{(p_1 - c)(4 - \Psi)} && < 0 \\
\frac{\partial \hat{\tau}_1}{\partial w_0} &= -\frac{2\left(1 + \frac{R_0}{\bar{R}}\right)}{(p_1 - c)(4 - \Psi)} && < 0
\end{aligned} \tag{3.24}$$

So, all else equal, government take of resource extraction profits could be expected to *increase* with (existing) reserves R_0, but *decrease* with higher undiscovered reserves $\bar{R}$, with higher strategic value of resource production σ – e.g. because of higher import dependence – or with higher exploration costs w_0. The impact of production cost is slightly more complicated. The partial derivative of $\hat{\tau}_1$ with respect to c is:

$$\frac{\partial \hat{\tau}_1}{\partial c} = \frac{2(-w_0 - \sigma) + (p_2 - p_1 + \sigma)\Psi + 2(p_2 - p_1 - w_0)\frac{R_0}{\bar{R}}}{(p_1 - c)^2(4 - \Psi)} \tag{3.25}$$

which could be either positive or negative, depending on the parameter configuration. The impact of production cost on government take is therefore ambiguous in this model. Finally, we study the impact of government typology, as measured by autocracy α and benevolence β. Since the numerator of expression (3.23) is increasing in Ψ while the denominator is decreasing in Ψ, it is clear that $\partial \hat{\tau}_1 / \partial \Psi > 0$. Hence,

$$\frac{\partial \hat{\tau}_1}{\partial \alpha} = \frac{\partial \hat{\tau}_1}{\partial \Psi}\frac{\partial \Psi}{\partial \alpha} = \frac{\partial \hat{\tau}_1}{\partial \Psi}\delta_g(1 - \beta)(P_{nr} - P_e) \geq 0 \tag{3.26}$$

where we have used the fact that $P_{nr} > P_e$, as discussed in Section 3.2.1.

Likewise for $\partial\hat{\tau}_1/\partial\beta$:

$$\begin{aligned}
\frac{\partial\hat{\tau}_1}{\partial\beta} &= \frac{\partial\hat{\tau}_1}{\partial\Psi}\frac{\partial\Psi}{\partial\beta} \\
&= \frac{\partial\hat{\tau}_1}{\partial\Psi}\delta_g(1-\alpha P_{nr}-(1-\alpha)P_e) \\
&\geq \frac{\partial\hat{\tau}_1}{\partial\Psi}\delta_g(P_{nr}-\alpha P_{nr}-(1-\alpha)P_e) \\
&= \frac{\partial\hat{\tau}_1}{\partial\Psi}\delta_g(1-\alpha)(P_{nr}-P_e) \qquad \geq \quad 0
\end{aligned} \tag{3.27}$$

All else equal, according to this model the government take of resource extraction profits would *increase* with higher autocracy or higher benevolence. Simply put: under comparable circumstances, a dictator would impose a higher tax than a non-dictatorial government, and a benevolent government would impose a higher tax than a predatory or corrupt government.

To provide some intuition for these results, we need to distinguish three revenue streams for the government:

(i) Revenues from production of existing reserves in period 1

(ii) Revenues from exploration and subsequent production of undiscovered reserves in period 1

(iii) Revenues from exploration and subsequent production of undiscovered reserves in period 2

A higher tax rate in period 1 will *increase* revenue stream (i), but *decrease* revenue stream (ii) due to reduced exploration. Since the latter effect leaves more undiscovered reserves for period 2, a higher period-1 tax rate will also *increase* revenue stream (iii). The period-1 tax rate in equilibrium depends on the trade-off between these three revenue streams. At the equilibrium point, the total revenues in period 1, i.e. the sum of revenue streams (i) and (ii), would *decrease* with higher period-1 tax rate. Indeed, if the total period-1 revenue were to increase with higher period-1 tax rate, then the government would increase the period-1 tax rate, since also the period-2 revenues (revenue stream (iii)) would increase with higher period-1 tax rate. Given that we consider an equilibrium, we can therefore say that total period-1 revenues *decrease* with higher period-1 tax rate, while total period-2 revenues *increase* with higher period-1 tax rate.

As a result of these effects, the government faces a trade-off when deciding on its government take: a higher current tax rate increases long-term revenues, but decreases short-term revenues. The resulting choice of current tax rate will depend on the attitude of the government towards the future. Assuming that autocratic governments have a higher probability of staying in power than democratic governments, they will assign a higher weight to future revenues, hence choose a higher current tax rate. Benevolent governments consider societal benefits rather than their own benefits while in power, so they would also put a higher weight on future revenues, hence also choose a higher current tax rate. A non-benevolent government, on the other hand, especially a non-autocratic non-benevolent government, has an incentive to lower tax rates in the short term, thereby encouraging earlier exploration so that it can get access to the resource revenues within its time in office. This is analogous to the dynamic common-pool problem in public debt theory,[10] with debt being equivalent to a negative resource.

3.3 Empirical testing

3.3.1 Cross-sectional estimates using OLS

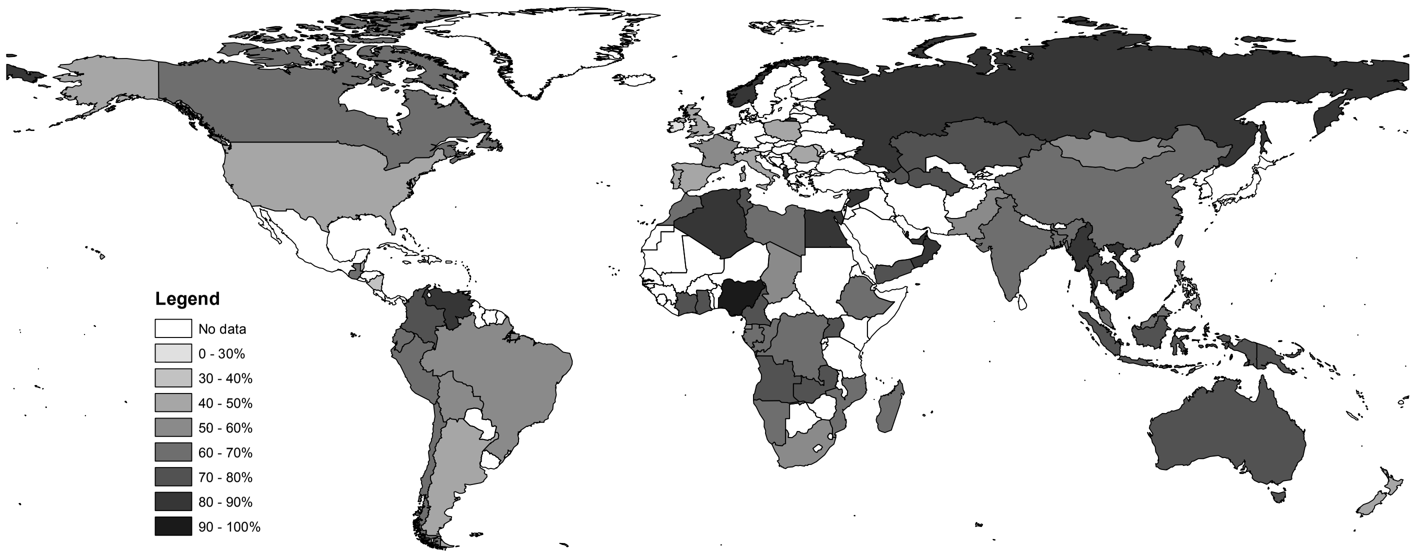
In this section, we test the theoretical conclusions from Section 3.2 on a cross-section of 77 oil-producing countries, using OLS. Our objective is to perform an empirical analysis of the main determinants of government take in the oil sector, and to verify our theoretical conclusions about the impact of government autocracy, government benevolence and geological characteristics, on government take.

Our dependent variable is *GovTake*: government take expressed in percent. Data on government take is scarce. For the cross-sectional analysis in this paper, we compile data points from six different sources: Johnston (1994, 1999, 2003), Van Meurs and Seck (1997), Himona (2005) and Tordo (2006). In cases when multiple sources provide an estimate of the government take of the same country, we give preference to the most recent source. When multiple values are given for the same country in a single source (e.g. varying by region within the country), we use the average of all values. In total, the data set contains 77 countries. This is comparable to the sample size of the cross-sectional analyses of Mulligan et al. (2004, Table 2) and Persson and Tabellini (2003, Table 3.1) on corporate tax rates and central government revenues, respectively. Figure 3.1 shows the data.

Based on the model of Section 3.2 the main independent variables of interest are the degree of autocracy, the government's benevolence, discovered and undiscovered reserves, whether a country is an importer or exporter, and its oil exploration and production costs. The degree of autocracy, denoted with variable name *Aut_Polity2*, is measured using the *polity2* variable from the Polity IV project (Marshall and Jaggers, 2010), inverted and scaled to the 0-1 range with 0 denoting full democracies and 1 denoting absolute autocracies. We take the average value of the period to which the *GovTake* estimates apply

[10]See e.g. Persson and Tabellini (2000, Chapter 13).

Figure 3.1: Government take in the oil sector in 77 countries. Source: Johnston (1994, 1999, 2003), Van Meurs and Seck (1997), Himona (2005) and Tordo (2006)



Note: The largest oil producer, Saudi Arabia, is missing from the analysis. The reason is fairly simple: oil production in Saudi Arabia is fully in the hands of the state-owned national oil company, hence it is impossible to distinguish between government take and contractor take.

(1994-2006). For Belize and Brunei, which are not assessed in the Polity IV project, we impute values based on data from Freedom House (2011). The degree of benevolence, denoted with variable name *Ben_WGI*, is measured using the 2006 value of the *Control of Corruption* variable from the Worldwide Governance Indicators (Kaufman et al., 2010), also scaled to the 0-1 range. Countries' already discovered reserves, denoted with variable name *Reserves*, are based on the 2006 estimate of oil reserves made by BP (2011). Reserves are assumed to be 0 for those countries that are not mentioned by BP (2011). Undiscovered reserves, variable name *Undiscovered*, are based on the country estimates made by USGS (2000)[11]. *OilImporter* is a dummy indicating whether a country was an oil importer in 2006, as can be inferred from the production and consumption data of BP (2011). Data on production costs, variable name *ProdCost*, is drawn from Aguilera et al. (2009), who provide production costs for all major oil and gas producing basins. Production costs for a country are computed as the average of all basins that are wholly or partly in that country, weighted by future oil reserves in that basin.[12] No suitable data set on exploration costs (w_0) was found, so exploration costs are not part of the analysis.

Besides the above-mentioned independent variables of interest, we include a number of covariates, in order to control for generic country characteristics. We include all covariates used by Mulligan et al. (2004, Table 2) in their regression of corporate tax rates across democratic and non-democratic countries. We also include the covariates used by Persson and Tabellini (2003, Table 3.1) in their cross-country regression of central government revenue.[13] Combining all of these, we have three groups of exogenous covariates. The first group consists of *economic characteristics*: *LYP*, the log of PPP converted GDP per capita (chain series) according to the Penn World Tables (Heston et al., 2011); *OECD*, a variable defined by Persson and Tabellini (2003) indicating whether a country was an OECD member in the early 90s; *Trade*, the sum of exports and imports of goods and services measured as a share of gross domestic product according to the World Development Indicators – WDI (World Bank, 2011); and *TransEcon*, a dummy for former communist countries, taken from the Global Development Network Growth Database (Easterly, 2001). The second

[11]Data on cooperative jurisdictions (e.g. Australia/Indonesia Zone of Cooperation) is distributed 50/50 between the two cooperating countries.

[12]For countries that do not contain any part of any of the specified basins, we use the estimates for 'Remaining provinces' which are provided per subcontinent, unless the country contains no undiscovered reserves according to USGS, in which case we use the estimates for 'Areas not assessed in USGS', which are also provided per subcontinent.

[13]A few of the covariates used by Persson and Tabellini (2003) are not included in our regression: *FEDERAL* (which measures whether a country has a federal government structure or not) because this is mostly relevant for the specific case of Persson and Tabellini (2003) since they look only at central government revenue, whereas our data on government take includes all taxes imposed; *GASTIL* because it measures the same thing as our variable *Aut_Polity2*; and the dummies indicating colonial origin, because most of this effect is already captured by the variable *Leg_British* used by Mulligan et al. (2004).

group of exogenous covariates consists of *population characteristics*: *LPop*, the log of population according to the Penn World Tables; *Prop1564*, the proportion of people aged 15-64 in the population according to the WDI; and *Prop65*, the proportion of people aged 65+ in the population, again from the WDI. Such population characteristics may be important for oil taxation because they influence the government's budget constraints: *Prop1564* because it is linked to government revenue potential, and *Prop65* because it may be related to expenditures. The third group of exogenous covariates. consists of *regional/colonial dummies*: *AFRICA*, *ASIAE* and *LAAM*, indicating whether a country is in Africa, eastern and southern Asia (except Japan, which is included in *OECD*), or southern and central America including the Caribbean, respectively; and *Leg_British*, indicating whether the origins of a country's legal system are British, as defined by Easterly (2001).

Table 3.1 shows the results of ordinary least squares (OLS) regression of *GovTake* on the political and geological characteristics, first without any controls, then progressively including the three groups of controls: economic characteristics, population characteristics, and regional/colonial dummies.[14] The results are broadly in line with the theoretical predictions of Section 3.2. First of all, it is indeed shown that dictatorship, i.e. a higher value of *Aut_Polity2*, has a positive effect on government take. The effect remains significant also when additional controls are added. The results suggest that, all else equal, a dictator would have a 15% higher tax rate on oil extraction than a comparable democracy. Secondly, government benevolence, i.e. a higher value of *Ben_WGI*, also has a significant positive effect on government take, except in the most parsimonious specification without any controls. All else equal, a benevolent government's oil extraction tax rate may be 24% (or more) higher than a comparable predatory government's tax rate. The effect of *Reserves* is also positive and significant in two specifications, although less pronounced than the other effects. *Undiscovered* reserves do not seem to have any significant effect. This may be due to the omittance of exploration costs from the regression. Indeed, the size of the undiscovered reserves may be negatively correlated with exploration costs. Hence, the negative direct impact of undiscovered reserves on government take – as predicted by the theoretical model – may be masked by the negative impact of exploration costs on government take. Since exploration data is unavailable however, this cannot be confirmed based on this analysis. *ProdCost* does not have a significant coefficient either, but this was to be expected since the model does not predict an unambiguous direction of the impact. *OilImporter*, an indicator of strategic value of oil production, has a significant negative effect as expected, except in the simplest

[14] Note that we do not use robust standard errors, because in several cases these turn out to be slightly smaller than the conventional standard errors, which may hint at finite sample bias. Unfortunately, increasing the sample is not an option since there are only a limited number of oil-producing countries in the world. As mentioned before, sample size is already nearly as large as for Persson and Tabellini (2003, Table 3.1) and Mulligan et al. (2004, Table 2).

Table 3.1: OLS regression results

	(1) GovTake	(2) GovTake	(3) GovTake	(4) GovTake
Aut_Polity2	21.57***	20.50***	16.75***	14.73**
	(5.276)	(5.087)	(5.667)	(7.290)
Ben_WGI	-8.299	23.76*	24.40*	33.00**
	(7.670)	(12.88)	(13.37)	(14.70)
Reserves	0.158	0.167*	0.136	0.188*
	(0.0954)	(0.0900)	(0.0963)	(0.0984)
Undiscovered	0.0404	0.0683	0.0852	0.125
	(0.134)	(0.134)	(0.143)	(0.142)
OilImporter	-3.555	-7.778**	-7.734**	-8.764**
	(3.482)	(3.510)	(3.596)	(3.666)
ProdCost	0.0757	-0.0357	-0.106	-0.253
	(0.249)	(0.250)	(0.262)	(0.264)
LYP		-3.409*	-3.032	-4.105
		(1.838)	(2.598)	(2.775)
OECD		-10.34*	-3.415	-9.506
		(5.316)	(7.165)	(7.765)
Trade		-0.0521	-0.0649	-0.107**
		(0.0372)	(0.0419)	(0.0456)
TransEcon		-3.079	2.081	-0.535
		(4.351)	(6.210)	(7.358)
LPop			0.0850	-1.412
			(1.170)	(1.303)
Prop1564			0.0752	-0.264
			(0.359)	(0.388)
Prop65			-0.992	-0.728
			(0.684)	(0.724)
AFRICA				-6.065
				(5.803)
ASIAE				5.396
				(5.307)
LAAM				-6.922
				(6.692)
Leg_British				-2.166
				(3.326)
N	77	74	74	74
adj. R^2	0.405	0.484	0.479	0.497

Standard errors in parentheses.

All regressions include a constant term (not shown).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

specification without any controls. Among the controls, it seems that most of the explanatory power is in the economic characteristics. The inclusion of population characteristics and regional/colonial dummies does not surface any significant coefficients. In fact, the inclusion of population characteristics actually lowers the adjusted R^2 slightly.

3.3.2 IV estimates

The causal effects of interest in the regression of the previous section may be obscured by endogeneity bias. In this section we therefore use an instrumental variables (IV) approach to verify our findings from OLS.

Our first concern about the OLS specifications of Table 3.1 is possible *omitted variable bias* with respect to the political characteristics *Aut_Polity2* and *Ben_WGI*, meaning that there would be unobserved country characteristics influencing *GovTake*, which, through their correlation with *Aut_Polity2* and *Ben_WGI* would bias the coefficient estimates of these two variables in Table 3.1. This may be an important issue, although e.g. Mulligan et al. (2004) do not address this in their OLS regression of tax rates across democracies and non-democracies. Omitted variable bias in our paper has already been partly addressed by including a large set of covariates in the OLS regression. To address the problem in a more structural way, we use exogenous instruments for *Aut_Polity2* and *Ben_WGI*. We take three instruments from the well-known paper by Hall and Jones (1999): *EurFrac*, the fraction of the population speaking one of the five primary European languages (including English) as their mother tongue; *EngFrac*, the fraction of the population speaking English as their mother tongue; and *Lat01*, the distance from the equator, measured as the absolute value of latitude and scaled to lie between 0 and 1.[15] An explanation of the mechanism through which these instruments have an effect on political institutions is given by e.g. Acemoglu et al. (2001). Simply put, these instruments are related to whether European colonists used a particular country as a place to settle, or primarily for exploitation of resources. The variable *Lat01*, in particular, is related to tropical diseases, which create an inhospitable environment that encourages exploitation rather than settlement. Persson and Tabellini (2003) use the same three instruments in their analysis of the size of government. In addition, Persson and Tabellini (2003) use three discretely measured indicators of the date of origin of each country's constitution. This is meaningful for Persson and Tabellini (2003) because they consider only democracies, but not in our case, so we limit ourselves to the three instruments *EurFrac*, *EngFrac* and *Lat01*. Anyhow, the three indicators

[15] Unlike Hall and Jones (1999), we do not use the 'Frankel-Romer predicted trade share' as an instrument because this is specific to Hall and Jones's analysis of country productivity. Likewise, Persson and Tabellini (2003) use this instrument only in their analysis of productivity, and not in their analysis of the size of government (which is closest to our analysis of government take).

of constitutional origin used by Persson and Tabellini are very weak instruments, as also highlighted by Acemoglu (2005) in his critique of their work, so including them would increase the risk of weak instrument bias.[16]

Using *EurFrac*, *EngFrac* and *Lat01* as instruments for *Aut_Polity2* and *Ben_WGI*, we run a two-stage least squares (2SLS) regression, including as always the geological variables and all controls related to economic and population characteristics. The regional/colonial dummies are not used, like in Persson and Tabellini (2003). Columns (1) and (2) in Table 3.2 show the first-stage results, i.e. the regression of *Aut_Polity2* and *Ben_WGI* on the instruments and the controls. Column (1a) in Table 3.3 shows the second-stage results using 2SLS. Columns (1) and (2) of Table 3.2 indicate that there is a weak instrument problem in this first specification. Persson and Tabellini (2003) face the same issue and solve this by not including any controls (except one) in the first stage. This non-standard approach may lead to biased results, and is also one of the main points of critique in the review by Acemoglu (2005). If in our case we removed all controls from the first stage, the first-stage results would look like columns ($1'$) and ($2'$) of Table 3.2. This would solve the weak instrument problem: in both columns the F statistic is larger than the recommended value of 10, suggested by Stock et al. (2002).[17] Nevertheless, we do not use ($1'$) and ($2'$), and instead follow a different approach to address the weak instrument problem. Our approach is two-fold. First, in column (1b) of Table 3.3, we show the limited information maximum likelihood (LIML) estimation results, which are less sensitive to the weak instrument problem. Note that the estimates are very similar to those of column (1a), which does give a bit more confidence in the results. Second, as an alternative, we study the effects of *Aut_Polity2* and *Ben_WGI* separately and use the single best instrument for each. In the case of *Aut_Polity2* the best instrument is *EurFrac*, while in the case of *Ben_WGI* it is *Lat01*. All above-mentioned controls are included in the first stage. The first-stage results when focusing on one endogenous variable at a time and using single instruments are shown in columns (3) and (4) of Table 3.2. For *Aut_Polity2* the first-stage F-statistic of the excluded instrument increases to 13.72, which suggests the instrument *EurFrac* is strong enough. For *Ben_WGI*, the F-statistic improves but remains below 10. However, just-identified IV is median-unbiased, hence the weak instrument problem is anyhow likely to be less of an issue in this case. Columns (3) and

[16]It should be noted that said critique also questions the validity of the Hall and Jones (1999) instruments *EurFrac*, *EngFrac* and *Lat01* in the analysis of Persson and Tabellini (2003). The main part of that critique, however, is on the use of these instruments for the explanation of specific institutional features (e.g. whether a country has a majoritarian or proportional electoral system) as opposed to institutional quality in general. Since our instrumented variables *Aut_Polity2* and *Ben_WGI* are more similar to the latter, the use of the Hall and Jones (1999) instruments seems at least more justified in our case than in the case of Persson and Tabellini (2003).

[17]It should be noted that the recommendation made by Stock et al. (2002) was for the case of a single endogenous variable, while our case has two.

Table 3.2: First-stage IV regression results for autocracy (*Aut_Polity2*) and benevolence (*Ben_WGI*)

	(1) *Aut_Polity2*	(2) *Ben_WGI*	(1′) *Aut_Polity2*	(2′) *Ben_WGI*	(3) *Aut_Polity2*	(4) *Ben_WGI*
EurFrac	-0.334***	-0.0114	-0.437***	0.127**	-0.343***	
	(0.110)	(0.0472)	(0.100)	(0.0534)	(0.0926)	
EngFrac	0.0294	0.0861	-0.0177	0.199**		
	(0.130)	(0.0560)	(0.144)	(0.0767)		
Lat01	0.158	0.210*	-0.298	0.494***		0.207**
	(0.276)	(0.119)	(0.179)	(0.0954)		(0.112)
Controls†	Yes	Yes	No	No	Yes	Yes
N	68	68	68	68	68	68
adj. R^2	0.536	0.781	0.338	0.522	0.549	0.779
F (excluded instruments)	4.58	2.00	12.41	25.36	13.72	4.64

Standard errors in parentheses.

All regressions include a constant term (not shown).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

† *Reserves*, *Undiscovered*, *OilImporter*, *ProdCost*, *LYP*, *OECD*, *Trade*, *TransEcon*, *LPop*, *Prop1564* and *Prop65*.

Table 3.3: Second-stage IV regression results for autocracy (*Aut_Polity2*) and benevolence (*Ben_WGI*)

	(1a) *GovTake*	(1b) *GovTake*	(3) *GovTake*	(4) *GovTake*
Aut_Polity2	41.72**	41.80**	40.26***	
	(16.42)	(16.47)	(12.88)	
Ben_WGI	-67.24	-67.65		31.31
	(57.73)	(57.97)		(56.20)
Estimator	2SLS	LIML	2SLS	2SLS
Number of instruments	3	3	1	1
Variables instrumented	2	2	1	1
N	68	68	68	68
adj. R^2	0.014	0.010	0.390	0.447

Standard errors in parentheses.

All regressions include *Reserves*, *Undiscovered*, *OilImporter*, *ProdCost*, *LYP*, *OECD*, *Trade*, *TransEcon*, *LPop*, *Prop1564*, *Prop65* and a constant term (not shown).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(4) of Table 3.3 show the second-stage results corresponding to the first stage results in columns (3) and (4) of Table 3.2, respectively.

Looking at the results, the most striking feature is that the coefficient for *Aut_Polity2* is remarkably consistent across the different specifications in Table 3.3. The coefficient is actually larger than the OLS estimates shown in Table 3.1. It seems therefore that the positive effect of autocracy on government take is confirmed in this IV analysis. The results for *Ben_WGI* are less clear. In none of the IV specifications its effect on government take proves to be statistically significant. In specification (4) – the specification least subject to weak instrument bias – the coefficient turns positive and is similar in size to the OLS estimate. The error margin remains large however.

Besides the endogeneity of *Aut_Polity2* and *Ben_WGI*, a second concern about the results of the OLS analysis in Table 3.1 is possible *reverse causation* with respect to *Reserves*. Indeed, one could imagine that a high tax rate may already have led to prior underinvestment in exploration, and hence lower current reserves. The direct causal effect of reserves on government take may therefore be underestimated by the opposite reverse causal effect of government take on reserves. We again use an IV approach. The ideal instrument in this case would be a measure of resource-richness that would be independent of past exploration efforts, and hence unaffected by government take. We choose two instruments and use each of these in turn. The first instrument is *Res80*, the reserves per country as they were in 1980, according to BP (2011). These

Table 3.4: Second-stage IV regression results for reserves

	(1) *GovTake*	(2) *GovTake*	(3) *GovTake*	(4) *GovTake*
Aut_Polity2	17.45***	18.85***		
	(5.202)	(5.937)		
Ben_WGI	20.29	13.65		
	(12.41)	(13.40)		
Reserves	0.0136	-0.155	0.0577	-0.290
	(0.121)	(0.396)	(0.133)	(0.479)
Estimator	2SLS	LIML	2SLS	LIML
Instrument	*Res80*	*Prod65*	*Res80*	*Prod65*
N	66	62	66	62
adj. R^2	0.516	0.454	0.451	0.274
First-stage F (excluded instrument)	56.21	2.80	55.72	2.75

Standard errors in parentheses.

All regressions include *Undiscovered*, *OilImporter*, *ProdCost*, *LYP*, *OECD*, *Trade*, *TransEcon*, *LPop*, *Prop1564*, *Prop65* and a constant term (not shown).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

are probably much less affected by the current government take since they were discovered before 1980 under a fiscal regime that was probably different. 1980 is the earliest year available in the reserves time series of BP (2011). The second instrument is *Prod65*, oil production per country in 1965. While this is a less adequate measure for reserves, BP (2011) provides data back to 1965, which makes a reverse causal relation with current government take even less likely.

Table 3.4 shows the second-stage IV regression results under four different specifications, and also mentions first-stage F statistics in each case. Since *Prod65* proves to be a weak instrument (as expected), we use it only in combination with LIML estimation. Columns (1) and (2) include *Aut_Polity2* and *Ben_WGI* as covariates. Since these variables are also endogenous – as mentioned above – it may be safer to leave them out of this regression, as is done in columns (3) and (4). Overall, the effect of reserves on government take, which was already relatively weak in the OLS result, disappears in this IV analysis. The next section may cast some light on a possible underlying mechanism for this.

Table 3.5: Explanatory power of Fraser Institute indices

	(1) *GovTake*	(2) *GovTake*	(3) *GovTake*
FiscTerms_0711	0.228**		0.304
	(0.0911)		(0.207)
TaxReg_0711		0.191**	-0.0859
		(0.0939)	(0.210)
Constant	58.85***	60.06***	59.26***
	(3.594)	(3.739)	(3.747)
N	69	69	69
adj. R^2	0.072	0.044	0.060

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.3.3 Panel data estimates

An alternative approach to address the omitted variable bias mentioned in the previous section, is panel data analysis with country fixed effects. Since time series data on government take is not available, we use a proxy. From 2007 to 2011, the Fraser Institute has conducted an annual *Global Petroleum Survey*, which identifies barriers to investment in the upstream petroleum industry across countries/jurisdictions, based on questionnaire responses from firms in the industry (Angevine et al., 2007-2011). In particular, two questions in the survey ask to what extent petroleum investment in a given jurisdiction is deterred by 'fiscal terms' and by 'taxation regime', respectively. The responses to these questions result in a score per jurisdiction on the dimensions 'fiscal terms' and 'taxation regime'. These scores, or a combination thereof, could be viewed as a proxy for government take. Despite the short time span of the time series, we will use these data for our panel data estimation. Table 3.5 shows the results of a regression of *GovTake*, as defined earlier, on *FiscTerms07_11* and/or *TaxReg_0711*. The latter two variables represent the average of the scores on the dimensions 'fiscal terms' and 'taxation regime', respectively, across all editions of the *Global Petroleum Survey* from 2007 to 2011. The table shows that both *FiscTerms07_11* and *TaxReg_0711* each individually constitute a statistically significant explanatory variable for *GovTake*. When combined, the significance disappears, and the coefficient of *TaxReg_0711* becomes negative.[18] Based on these results, we choose to use the annual scores on 'fiscal terms' from the survey as a proxy for government take. We scale the scores by 0.23 to make the resulting estimates comparable with previous sections. The

[18]Clearly, the regression suffers from collinearity, given the correlation of 0.84 between *FiscTerms07_11* and *TaxReg_0711*.

resulting annual variable is called *FiscTerms*. Since the composition of the survey has changed over the years, we have an unbalanced panel.

Data for *Aut_Polity2*, *Ben_WGI* and *Reserves* is also available as a time series, from the same sources as the static data used in previous sections. A potential issue, however, is that there is only limited variation in *Aut_Polity2* for any given country over the short time frame under consideration. As an alternative to *Aut_Polity2*, we therefore use the *Voice & Accountability* variable from the Worldwide Governance Indicators (Kaufman et al., 2010). We invert and scale the latter variable to the 0-1 range, as before, and denote it with *Aut_WGI*. Unlike *Aut_Polity2*, which is measured using a discrete set of possible values, the variable *Aut_WGI* is continuous, and hence may be able to record smaller variations in democratic characteristics. The correlation between *Aut_Polity2* and *Aut_WGI* is 0.84. The data on *ProdCost* and *Undiscovered* is static, hence we cannot use it in the panel data analysis. Data on *OilImporter*, i.e. the dummy indicating whether a country's oil consumption is larger than its production, hardly shows any variation during the period under observation, so we do not use it. Among the controls for economic, population and regional/colonial characteristics, only *LYP*, *Trade*, *LPop*, *Prop1564* and *Prop65* have variation over time, and are included in the regression. We include country fixed effects, to account for country-specific features not accounted for by the controls, and time effects, to account for global changes affecting all countries, such as oil market dynamics.[19]

The results are shown in Table 3.6. Column (1) is the most straightforward specification. Column (2) includes lags for all variables of interest. Columns (3) and (4) are the same as columns (1) and (2), respectively, but with *Aut_Polity2* replaced by *Aut_WGI*. The coefficients of autocracy (*Aut_Polity2* or *Aut_WGI*) and its lag are positive everywhere, which is in line with our OLS and IV estimates. Although nearly all these coefficients are larger than the standard error, only the coefficient of the lag of *Aut_Polity2* is significant. The latter may however be an artefact caused by the infrequent occurrence of variation in *Aut_Polity2*. The coefficients of *Ben_WGI* and its lag are negative everywhere, and but since nearly all of them are quite a bit smaller than the standard error in absolute terms, this does not necessarily need to be inconsistent with previous results. The results for *Reserves*, although again only marginally significant, provide some insight regarding our earlier contradictory findings. Based on Table 3.6, it seems that the link between current reserves and fiscal regime is negative, which may be an indication of the reverse causation mechanism described earlier. By contrast, the coefficient of the lag of *Reserves* is positive, which may be an indication of the direct effect of reserves

[19] Moreover, the reporting of the Fraser Institute's *Global Petroleum Survey* was changed between the 2008 and 2009 editions. Until 2008, jurisdictions were ranked based on the percentage of respondents answering that fiscal terms were a 'Strong deterrent to investment' or that they 'Would not invest'. Starting in 2009, the score also includes respondents answering that they consider fiscal terms a 'Mild deterrent to investment'. This effect should be captured by the year dummies.

Table 3.6: Fixed effects panel regression results

	(1) *FiscTerms*	(2) *FiscTerms*	(3) *FiscTerms*	(4) *FiscTerms*
Aut_Polity2	5.112	4.100		
	(3.478)	(3.544)		
Lagged *Aut_Polity2*		9.099***		
		(3.394)		
Aut_WGI			31.17	34.04
			(24.75)	(26.09)
Lagged *Aut_WGI*				14.30
				(16.90)
Ben_WGI	-9.390	-16.80	-5.520	-3.891
	(13.74)	(14.09)	(13.85)	(14.36)
Lagged *Ben_WGI*		-2.198		-5.184
		(13.26)		(14.32)
Reserves	-0.0572*	-0.0698	-0.0515	-0.0653
	(0.0328)	(0.0591)	(0.0321)	(0.0605)
Lagged *Reserves*		0.0162		0.0220
		(0.0833)		(0.0841)
N	235	235	236	236
adj. R^2	0.702	0.711	0.702	0.698

Standard errors in parentheses.

All regressions include *LYP*, *Trade*, *LPop*, *Prop1564*, *Prop65*, country and year dummies, and a constant (not shown).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

on government take as described in our model. Overall however, it seems the panel data analysis is suffering from the short time span of the data, which impedes the derivation of statistically significant conclusions.

3.4 Conclusions

This paper develops a theoretical model of the political economy of resource-rich countries, featuring taxes on resource extraction, endogenous exploration and a formalized classification of government types. The model predicts that, all else equal, the government take of resource extraction profits would increase with government autocracy and/or benevolence and with higher existing reserves base, while the government take would decrease with higher undiscovered reserves or higher import dependence. Nearly all of these effects are confirmed empirically for the case of oil, using OLS on a cross-section of 77 countries. The statistically significant positive effect of government autocracy (i.e. dictatorship) on government take is shown to be robust in an IV analysis, and, to a lesser extent, a panel data regression.

The model of this paper can be further improved by modelling the behavior of citizens in a more detailed way, thereby endogenizing the election probability or non-revolution probability. Furthermore, an extension to more than two periods or to continuous time would enable a more refined modelling of the complex trade-off between current and future earnings. Finally, the use of longer time series of resource taxation data would allow for more conclusive panel data estimates.

3.5 Acknowledgements

The author would like to thank the participants of the ETE day in Leuven on March 3rd 2008, BEED 2009 at UCL-Core on February 2nd 2009 (in particular Tom Verbeke and Laurent Franckx), the workshop on “Political Economy and Institutions in Transport, Energy and Environment” at KULeuven-CES on February 26th 2009 (in particular Ami Glazer and Stef Proost), the Spring Meeting of Young Economists 2009 in Istanbul on April 23rd-25th 2009 (in particular the discussant Michael Funk), the International Energy Workshop in Venice on June 17th-19th, EAERE 2009 in Amsterdam on June 24th-27th (in particular Jana Stoever and the discussant Elissaios Papyrakis), and a workshop on “Political Economy and the Environment” at UCL-Core on October 22-23rd 2009, for the insightful discussions on this topic. The author is also grateful to Roberto F. Aguilera for kindly providing his database on oil production costs. The usual disclaimers apply. Finally, the author gratefully acknowledges financial support from the KULeuven Energy Institute.

Chapter 4

Taxation of Nuclear Rents: Benefits, Drawbacks, and Alternatives

This chapter has been submitted to *Energy Economics.*

4.1 Introduction

Towards the end of 2008, the government of Belgium attempted to reduce its budget deficit by imposing a tax of 250 million euro on nuclear power producers. Despite appeals to Belgium's Constitutional Court, the tax was upheld and repeated as an annual tax in 2009 and 2010. The tax burden is allocated to nuclear producers in proportion to their capacities, and amounts to around 5 euros per MWh produced. Meanwhile in Germany, in the second half of 2010, the center-right government and the nuclear power producers agreed to extend the planned lifetime of Germany's nuclear power plants in return for a new fuel-rod tax and a compulsory contribution to a renewable energy fund, later to be replaced by a renewables levy per MWh produced. The total resulting charge to the nuclear producers in Germany amounted to slightly over 15 euros per MWh.[1]

This paper analyzes the short-run and long-run economic implications of the introduction of such taxes on nuclear production, and compares taxation with alternative policy measures.[2] The economic model is applicable to various

[1]The information in this paragraph is obtained from articles in De Tijd (March 3 2010 and December 8 2010), Financial Times (Sep 7 2010) and The Economist (June 2 2011). Note that as a result of the Fukushima accident, however, the German government eventually annulled the lifetime extension and maintained the plans for a phase-out by 2022.

[2]Note that our paper is focused on taxation of economic rents. The Fukushima accident

countries. Table 4.1 provides an overview of European countries that have nuclear energy in their generation mix. The table also shows that nuclear

Table 4.1: Overview of nuclear electricity generation capacity in Europe

Country	**Nuclear capacity** (Dec 31, 2010) MW	**Share of nuclear in total gross electricity generation** (2008) %	**Share of nuclear capacity owned by largest 2 players** (Dec 31, 2010) %	**Policy outlook** (Dec 31, 2010)
France	62950	76%	99.7%	Expansion ongoing
Germany	19895	23%	50.1%	Decommissioning by 2022
Sweden	9248	43%	75.0%	Stable
United Kingdom	9218	13%	100.0%	Stable
Spain	7409	19%	61.2%	Stable
Belgium	5839	54%	97.3%	Decommissioning 2015-2030
Czech Republic	3775	32%	100.0%	Expansion proposals
Switzerland	3220	40%	66.3%	Expansion proposals
Finland	2721	30%	100.0%	Expansion ongoing
Hungary	1946	37%	100.0%	Expansion proposals
Slovakia	1940	58%	100.0%	Expansion proposals
Bulgaria	1906	35%	100.0%	Expansion proposals
Romania	1310	17%	100.0%	Expansion proposals
Slovenia	664	38%	100.0%	Expansion proposals
Netherlands	479	4%	100.0%	Expansion proposals

Source: Platts (2010), Eurostat (2011).

firms often have large, concentrated market shares and may therefore be able to exercise market power on national or regional electricity markets. This is an important feature of the nuclear sector, which we will explicitly include in this paper using a ‘dominant firm – competitive fringe’ model. The model allows us to analyze the different types of nuclear rents, the potential for short-run taxation (as in the above example of Belgium), the possible long-run effects of such taxes, and alternative policy measures such as comprehensive lifetime extension deals (as in the German example), auctioning of licenses, and renewables investment quota.

The issue of nuclear rents has been studied from a policy perspective by e.g. CREG (2010a) and Matthes (2010). CREG (2010b) also study the German taxation solution. Such studies focus on detailed quantified analysis of specific

has illustrated the external cost of nuclear power, which could be addressed separately through a Pigouvian tax. The latter is not the subject of this paper.

cases. Our paper aims to complement this literature by developing a formal model that allows for graphical and analytical demonstration of the underlying economic principles.

The oligopolistic nature of our proposed model embodies a number of issues studied in the vast literature on electricity market deregulation, such as Newberry's (2002) analysis of the effectiveness of the deregulation process and the accompanying problems. Indeed, the ongoing policy challenges regarding taxation of nuclear rents are a consequence of the liberalization of formerly regulated electricity monopolies. A large number of papers analyze the electricity market as an oligopoly and many studies like Borenstein et al. (1999), Bushnell et al. (2004), Cardell et al. (1997) and Wolfram (1999) analyze market power of incumbent firms, which may transform their previous regulated monopoly rights into substantial unregulated market power.

Our paper fits into the broader literature on optimal regulation of oligopolies and the electricity sector in particular. Well-known literature such as Demsetz (1968) and Stigler (1971) questions the benefits of such regulation for society as a whole, as opposed to the special interest group of the electricity companies. Models that do advocate regulation of monopolies, typically have as a result that the first-best solution in the case of a technology with low marginal costs – such as nuclear power – is to subsidize the monopolist in order to allow him to set prices at marginal cost without going out of business (see e.g. Train, 1991). The same principle is valid in the 'dominant firm – competitive fringe' model used in this paper: we will find that the optimal tax rate of nuclear power in the long run may be negative. The key difference is that in our model there is a severe capacity constraint on nuclear production, which is especially relevant in the short run. The first-best solution is then to charge the price that reduces demand to the level of available capacity. Riordan (1984) proposes a subsidy scheme to create incentives for correct pricing. However, in the current case, this would result in a negative subsidy, i.e. a lump sum tax on the nuclear firm, which may be difficult to realize for legal reasons. The question still arises why there should be any taxation at all in this case: it may seem that nuclear firms earn a justifiable rent on a scarce resource that they acquired. The resource was however acquired in a regulated setting, in which costs were usually passed on directly to consumers through cost-of-service regulation. For that reason, governments may find it appropriate to tax nuclear rents. Our paper will therefore evaluate the short-run and long-run effects of a nuclear production tax, and study a number of alternative policy instruments.

Among the alternative policy instruments, we find that lifetime extension agreements negotiated with multiple potential players, and competitive auctioning of new nuclear licenses are the most attractive policies. The bargaining and auctioning aspect is stylized, in order to permit transparent comparison with the outcomes of taxation. A more elaborate treatment of nuclear capacity auctions is provided by Fridolfsson and Tangerås (2011).

Our monopolistic 'dominant firm – competitive fringe' equilibrium concept and its oligopolistic extensions are more similar to the typical Cournot models than to the more sophisticated Supply Function Equilibria (SFE). The advantage of the former is clearly computational convenience, as acknowledged by e.g. Ventosa et al. (2005), Borenstein et al. (1999), Hobbs & Pang (2007) and Wei and Smeers (1999). Willems et al. (2009) confront Cournot models and SFE with data of the German electricity market. The authors conclude that SFE models do not significantly outperform the Cournot approach when studying the German electricity market but that they rely on fewer calibration parameters and may therefore be more robust. Willems et al. (2009) suggest that Cournot models are "*...aptly suited for the study of market rules...*", while SFE are suited to study e.g. long-term effects of mergers. In our setting, the Cournot-style approach seems therefore justified.

The paper is structured as follows. Section 4.2 develops our analytical model of an electricity market with a significant share of nuclear production, and identifies the different types of nuclear rents. Section 4.3 investigates the potential magnitude and impact of a nuclear tax in the short run. Next, Section 4.4 demonstrates the long-run commitment disadvantages of a nuclear tax. Section 4.5 studies alternative policy measures. Section 4.6 illustrates the results numerically for the case of Belgium. Section 4.7 concludes the paper.

4.2 Model set-up

4.2.1 Demand and supply

We study a stationary electricity market with linear demand:

$$q(p) = q_0 - \beta p \tag{4.1}$$

with constant parameters q_0 and $\beta \geq 0$. To focus our thoughts, we assume that this is the electricity market of one single country. The demand $q(p)$ represents an 'average' demand in the course of a year, and we do not consider any demand variations within the year, or even within the days of the year.[3]

Two supply technologies are available: (i) nuclear generation with short-run marginal production cost c_n, and (ii) fossil-fuel-based generation with linearly increasing short-run marginal production costs $c_0 + \alpha q$. Because of the low short-run marginal production costs of nuclear power, we assume $c_n \ll c_0$. Nuclear generation q_n is limited by the installed nuclear capacity $q_{n,inst}$. Furthermore, we assume that there is a limit $q_{n,max}$ to the amount of nuclear capacity that can be installed: $q_{n,inst} \leq q_{n,max}$. The constraint may be due to various reasons: technical (e.g. limited access to cooling, insufficient ability of nuclear power to cope with demand variability), political, ethical, etc. For

[3]As we will see below, supply and demand are assumed to always intersect in the same linear part of the supply curve, so the 'average' demand corresponds to the 'average' price. In this case, the fact that we analyze only the average demand is therefore not a restriction.

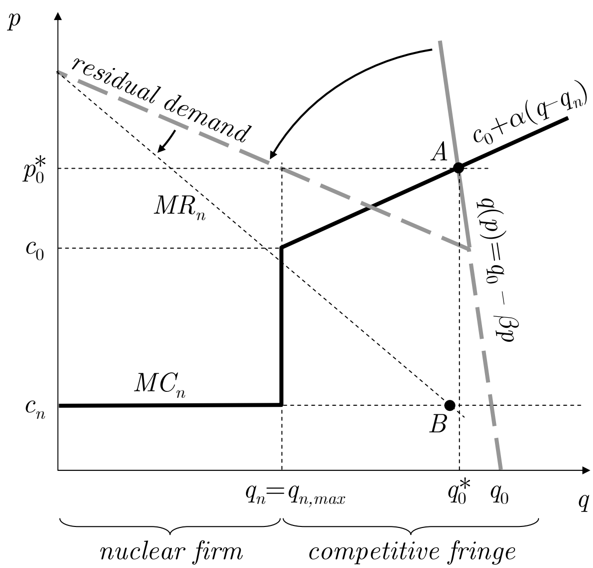
the sake of simplicity, we assume $q_{n,inst} = q_{n,max}$, i.e. nuclear capacity has already been installed up to maximum allowable level. The cost curve $c_0 + \alpha q$ aggregates different fossil-fuel power plants and also includes any relevant foreign production capacity that can be imported. Additional transmission costs for foreign production are included in $c_0 + \alpha q$.

4.2.2 Equilibrium concept

We consider a 'dominant firm – competitive fringe' game, which is fairly standard in microeconomic analysis.[4] We assume that all nuclear production capacity is controlled by one 'dominant firm'. Considering the strong concentration in nuclear power as shown in Table 4.1, this assumption is not unrealistic, and serves as a 'worst case' baseline for the cases with more than one producer. The fossil-fuel and foreign capacity represented by $c_0 + \alpha q$ is assumed to be operated by a 'competitive fringe'. In this model, the dominant firm behaves strategically, while the competitive fringe always behaves fully competitively and hence always produces up to the point where price equals marginal cost.

Figure 4.1 illustrates our equilibrium concept. The thick black line is the

Figure 4.1: Equilibrium in the 'dominant firm – competitive fringe' game described in this paper



industry marginal cost curve. Under perfect competition, the market equil-

[4]See e.g. Carlton and Perloff (2000).

brium would be at point A, the intersection of the industry marginal cost curve and the demand curve. Since it is in practice technically impossible to serve an entire electricity system with nuclear power, we assume that the intersection of supply and demand is always in the upward sloping part of the supply curve (i.e. the part provided by the competitive fringe). The equilibrium price and quantity under perfect competition are p_0^* and q_0^*:

$$p_0^* = \frac{c_0 + \alpha(q_0 - q_{n,max})}{1 + \alpha\beta} \tag{4.2}$$

$$q_0^* = \frac{q_0 - \beta(c_0 - \alpha q_{n,max})}{1 + \alpha\beta} \tag{4.3}$$

In our 'dominant firm – competitive fringe' model, however, the equilibrium can be determined by considering the residual demand curve (i.e. demand minus the quantities supplied by the competitive fringe) and analyzing the dominant firm as a monopolist on this residual demand curve.[5] As shown in Figure 4.1, the residual demand curve translates into a marginal revenue curve (MR_n) of the dominant firm. If nuclear capacity was not constrained by $q_{n,inst}$ then the 'dominant firm – competitive fringe' equilibrium would be at point B, i.e. the intersection of MR_n and the dominant firm's (constant) marginal cost curve. However, given the nuclear capacity constraint ($q_n \leq q_{n,inst}$), the dominant firm's quantity decision q_n becomes equal to $q_{n,inst} = q_{n,max}$. As a result, the equilibrium prices and quantities are the same as under perfect competition, at least in the case shown in Figure 4.1. Only with high operating costs and large capacity would the dominant nuclear have an interest in exploiting his market power by not fully using the available capacity. However, due to the relatively low short-run marginal cost of nuclear production, the constraint $q_n \leq q_{n,inst}$ is generally binding. As a result, in this model, even a single dominant nuclear firm typically does not have an incentive to withhold production. In order to keep our argument focused, we will assume that – in the absence of taxes on nuclear production – the constraint $q_n \leq q_{n,inst}$ is indeed binding.

This assumption can be expressed mathematically. From equation (4.1) and the cost curve of the competitive fringe, we can easily derive the inverse residual demand curve $p_R(q)$, and subsequently MR_n:

$$p_R(q) = \frac{c_0 + \alpha(q_0 - q)}{1 + \alpha\beta} \tag{4.4}$$

$$MR_n = \frac{c_0 + \alpha(q_0 - 2q)}{1 + \alpha\beta} \tag{4.5}$$

[5] Note that our model description implicitly provides the dominant firm with a Stackelberg leadership position vis-à-vis the competitive fringe. Ulph and Folie (1980) analyze an alternative configuration, in which a Nash equilibrium is reached between the dominant firm and the competitive fringe, and find that this leads to a number of undesirable properties, hence we do not consider this option here.

In the absence of the constraint $q_n \leq q_{n,inst}$, the nuclear firm's unconstrained production quantity $q_{n,uncon}$ can be computed by setting $MR_n = MC_n = c_n$:

$$q_{n,uncon} = \frac{1}{2}(q_0 + \frac{c_0 - c_n}{\alpha} - \beta c_n) \tag{4.6}$$

The constraint $q_n \leq q_{n,inst}$ will be binding iff $q_{n,inst} \leq q_{n,uncon}$. Since we assume $q_{n,inst} = q_{n,max}$, the latter condition can be expressed as:

$$2q_{n,max} \leq q_0 + \frac{c_0 - c_n}{\alpha} - \beta c_n \tag{4.7}$$

The first two terms of the right-hand side together $(q_0 - \beta c_n)$ are slightly larger than total electricity demand, hence it is reasonable to assume $q_{n,max} \ll q_0 - \beta c_n$. The last term $(c_0 - c_n)/\alpha$ is also large, because the difference in short-run marginal production cost between nuclear power and the fossil-fuel fired power plants of the competitive fringe is large. With the calibration parameters of Section 4.6, condition (4.7) is satisfied for all European countries in Table 4.1.

Until now, we have assumed that the dominant nuclear firm does not own any of the competitive fringe capacity. In reality, the firm may have diversified its asset portfolio by investing also in fossil-fuel capacity, which in our model is part of the competitive fringe. Let us therefore consider a case in which the nuclear firm owns not only all nuclear capacity, but also an amount q_f of fringe capacity. This set-up may have an effect on its nuclear production decisions. Note that we do not consider any withholding of q_f. The purpose of this paper is to study taxes on nuclear production; any withholding of q_f is a separate issue that is not directly affected by taxes on nuclear production.

The nuclear firm's ownership of q_f adds a term $q_f dp_R(q)/dq$ to MR_n in equation (4.5). We find:

$$MR_n = \frac{c_0 + \alpha(q_0 - 2q - q_f)}{1 + \alpha\beta} \tag{4.8}$$

Hence, the MR_n curve in Figure 4.1 is shifted downward by an amount $q_f \alpha/(1 + \alpha\beta)$, which moves the intersection point B with c_n to the left. The same effect would be accomplished by increasing c_n by an amount $q_f \alpha/(1+\alpha\beta)$. Hence, the key results regarding optimal tax rates in Sections 4.3 and 4.4 of this paper remain valid, provided that c_n is replaced with $c_n + q_f \alpha/(1 + \alpha\beta)$.[6] The intuition behind this is obviously that ownership of fringe capacity by the nuclear firm creates an additional opportunity cost when activating nuclear capacity: in addition to the direct nuclear production cost c_n and changes in nuclear revenues, the firm incurs lost revenues on the capacity q_f due to lower electricity prices.

[6]Note in particular that this requires $\beta \ll \alpha^{-1}$ for Propositions 4.1 and 4.2, and $\sigma_n = 0$ for Proposition 4.2.

4.2.3 Rents

Figure 4.1 demonstrates that nuclear producers obtain a rent $(p_0^* - c_n)q_n$. Three types of rent can be distinguished.

First, a large part is the perfectly competitive *inframarginal rent* required to cover the fixed costs (mostly investment costs) incurred by the nuclear firm. Indeed, while short-run marginal production costs are relatively low, the upfront capital investment of nuclear power is comparatively high. The discounted sum of inframarginal rents needs to cover that investment cost in order to provide sufficient investment incentives. We assume that the sum of the annuity of investment costs plus any relevant fixed operating costs is given by $f_n q_{n,inst}$.

Secondly, there may be a *scarcity rent.* Indeed, without the constraint $q_{n,max}$, the long-run equilibrium installed capacity of nuclear power would evolve such that the electricity price equals $f_n + c_n$. As a result of the constraint, the nuclear firm in Figure 4.1 obtains a scarcity rent $(p_0^* - c_n - f_n)q_n$.

Thirdly, nuclear firms may obtain rents due to *market power.* In Figure 4.1, no withholding takes place, hence there is no market power rent. As mentioned before, the low level of short-run marginal costs makes short-run withholding of nuclear capacity unlikely in general. However, as we will see below, there may be circumstances (e.g. in the event of nuclear taxation) that can lead to the long-run variety of withholding, namely underinvestment.

4.2.4 Government

In this paper, we study the policy options of the national government of the country under consideration. We assume that the government maximizes national welfare W:

$$W = CS + \sigma_n \pi_n + \lambda G \tag{4.9}$$

with CS the consumer surplus, π_n the profits of the nuclear firm, σ_n ($0 \leq \sigma_n \leq 1$) the fraction of the nuclear firm's shares owned by nationals of the country, G the government revenues from the nuclear sector (taxes, licenses, etc.), and λ the marginal cost of public funds[7]. We ignore the potential contribution of the competitive fringe's profits to national welfare. The choice of λ depends on the view one takes regarding government revenues. An assumption $\lambda = 0$ implies that government revenues are wasted. On the other hand, assuming $\lambda > 1$ implies that nuclear revenues are used productively to reduce other taxes, thereby eliminating distortions elsewhere in the economy. For the sake of simplicity, we assume $\lambda = 1$ in the remainder of this paper. Note that if, for ideological or other reasons, the government does not attach any importance to the income of the nuclear firm's national shareholders, this is equivalent to setting $\sigma_n = 0$.

[7] See e.g. Browning (1976).

Assuming a reservation price $\bar{p}$, we can conveniently write CS as:

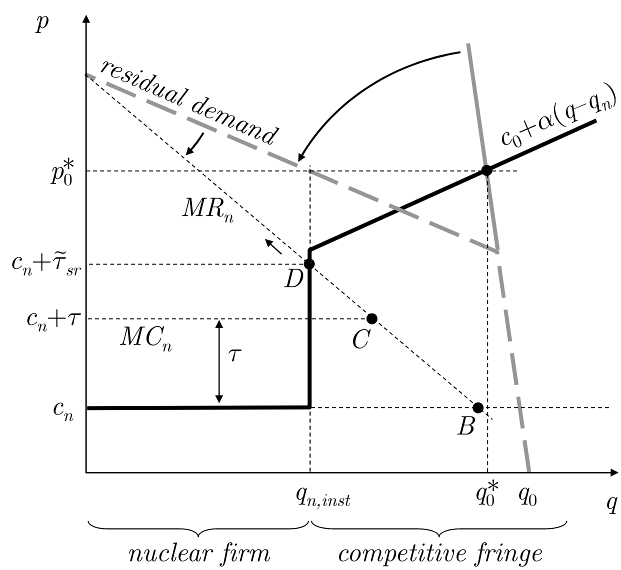
$$CS = \int_{p^*}^{\bar{p}} q(p)dp = (\bar{p} - p^*)q_0 - \frac{\beta}{2}(\bar{p}^2 - {p^*}^2) \tag{4.10}$$

with p^* the market price of electricity in the equilibrium under consideration.

4.3 Short-run taxation potential

We now consider the introduction of an unexpected tax on nuclear production, as it happened in the case of Belgium. 'Unexpected' means that the tax was not anticipated at the time of investment in the nuclear capacity. We consider a tax of τ per unit of nuclear electricity produced. The effect is illustrated in Figure 4.2. For the nuclear firm, the tax increases the short-run marginal cost of production from c_n to $c_n + \tau$. As a result, the intersection point of MR_n and MC_n moves from point B to point C. If τ is high enough, the intersection point may shift to point D or beyond. When the intersection point is beyond point D (indicated with a little arrow in Figure 4.2) the constraint $q_n \leq q_{n,inst}$ is no longer binding and the nuclear firm withholds some of its capacity.

Figure 4.2: Short-run impact of an unexpected tax on nuclear production



Proposition 4.1. *Under the assumptions of Section 4.2 and the additional conditions that $\beta \ll \alpha^{-1}$ (highly inelastic demand) and that:*

$$2q_{n,max} \leq q_0 + \frac{c_0 - c_n}{2\alpha} - \beta \frac{5c_n + 3c_0 + 2\alpha q_0}{8} \tag{4.11}$$

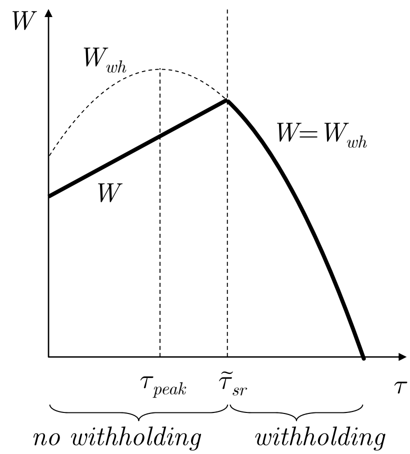
the government's optimal choice of the short-run nuclear tax rate is:

$$\tilde{\tau}_{sr} = \frac{c_0 + \alpha(q_0 - 2q_{n,max})}{1 + \alpha\beta} - c_n \tag{4.12}$$

Proof. First, observe that $\tilde{\tau}_{sr}$ is the maximum tax rate the government can impose without causing the nuclear firm to withhold. This can be easily seen by replacing c_n with $c_n + \tau$ in equation (4.7) and solving for τ.

Lowering the tax rate below $\tilde{\tau}_{sr}$ keeps the constraint $q_n \leq q_{n,inst}$ binding, hence it does not change the equilibrium. The only effect of lowering the tax rate below $\tilde{\tau}_{sr}$ is a linear transfer from government revenue to nuclear profits. Since $\sigma_n \leq 1$ this cannot increase the government's objective function W in equation (4.9). This is shown in Figure 4.3, for $\tau \leq \tilde{\tau}_{sr}$.

Figure 4.3: Shape of W as a function of τ



Raising the tax rate τ above $\tilde{\tau}_{sr}$ leads to withholding. The constraint $q_n \leq q_{n,inst}$ is not binding anymore and nuclear production $q_{n,1}$ is given by equation (4.6) with c_n replaced by $c_n + \tau$. The equilibrium price can then be easily derived:

$$p_1^* = \frac{1}{2}\left[c_n + \tau + \frac{c_0 + \alpha q_0}{1 + \alpha\beta}\right] \tag{4.13}$$

which can be substituted into equation (4.10) for CS. Using also $G = \tau q_{n,1}$, one can derive an analytical expression for the government's objective function

W_{wh} in case of withholding, which is a concave quadratic function of τ. As mentioned before, we assume $\lambda = 1$. For $\sigma_n = 0$, the maximum value of W_{wh} is reached for:

$$\tau_{peak} \approx \frac{c_0 - c_n}{2} + \frac{\alpha\beta}{4}\left(\alpha q_0 - \frac{c_0 - c_n}{2}\right) \tag{4.14}$$

in which we have approximated τ_{peak} using its first-order Taylor polynomial around $\alpha\beta = 0$. Due to condition (4.11), this value τ_{peak} is lower than the first-order polynomial of $\tilde{\tau}_{sr}$ around $\alpha\beta = 0$, hence τ_{peak} is outside the domain where withholding takes place. As a result, W declines with τ when $\tau \geq \tilde{\tau}_{sr}$, as is illustrated in Figure 4.3. This result is also valid if $\sigma_n \geq 0$, because then τ_{peak} will be even smaller since in this case a larger τ contributes an additional negative term to W through π.

Since it is impossible to increase W – neither by lowering τ below $\tilde{\tau}_{sr}$ nor by raising it above $\tilde{\tau}_{sr}$ – the choice $\tau = \tilde{\tau}_{sr}$ must be an optimal value. □

Condition (4.11) is similar to condition (4.7), but slightly more restrictive. With the calibration parameters of Section 4.6 the condition is satisfied when the maximum nuclear capacity represents less than 60% of total demand. Except for France, this is satisfied for all countries in Table 4.1.

Proposition 4.1 demonstrates that it is possible (and optimal in the short run) for governments to impose a substantial[8] unexpected tax on nuclear production, without even affecting the equilibrium on the electricity market. An immediate side-effect of such a tax is that incentives for efficiency improvements at nuclear power plants vanish. Indeed, according to equation (4.12) any investment that reduces c_n would lead to an increase of $\tilde{\tau}_{sr}$ by the same amount, leading eventually to the same total costs for the nuclear firm. The situation of the nuclear firm becomes similar to the situation of a regulated firm under cost-of-service regulation. Fabrizio et al. (2007) analyze efficiency improvements induced by a transition from cost-of-service regulation to market-oriented environments for a large set of US electric generating plants. Their estimates suggest that investor-owned utility plants in restructuring regimes reduced their labor and non-fuel operating expenses by 3 to 5 percent in anticipation of increased competition in electricity generation, relative to plants in states that did not restructure their markets. The reverse effects might be expected in the taxation case analyzed in this paper. In addition to this immediate impact, the tax may have an effect on long-run reinvestment. This will be studied in the next section.

[8] An indication of the likely size of such a tax is provided in Section 4.6.

4.4 Long-run effects of taxation on reinvestment

The government's ability and incentives to introduce ex-post taxes on nuclear production in the short run, as in the previous section, may hamper investments in new nuclear capacity – or any other generating capacity with high capital costs – because firms may fear expropriation of profits. This is a well-known problem in the tax competition literature. Janeba (2000) uses a model in which one firm produces for the world market and its profits are taxed by the government. When the production has to be local, there is a standard time inconsistency problem. The government has an interest to raise the tax after the investment, the firm understands this ex ante and the firm would never invest. In Janeba's model this problem can be avoided by the firm at a high cost by building extra capacity in another country and to shift production to the country where the net profit is the highest. According to Janeba, when capacity costs are sufficiently low, this is the equilibrium. For higher capacity costs, there is an equilibrium without investment. In our case, even if there is enough international transmission capacity, high (nuclear) capacity costs would rule out an equilibrium with excess capacities. A second way to avoid the time inconsistency problem is to rely on other constraints on the behavior of governments. One legal option is to have constitutions that rule out governments changing the tax rules. This type of rule exists in order to limit government debt but is rather naive as governments have by definition the power to tax profits more and lower other taxes. A second constraint could be reputation. In order to assess this constraint we need a behavioral model of the government as an agent that desires to be reelected by its voters (Besley, 2006). A government raising taxes ex post can be seen as an unreliable government by investors but voters with a shorter-term memory and/or perceiving the nuclear tax as a justified tax on excess profits or compensating nuclear risks may consider such a government as defending its interests. The best evidence is that in countries such as Germany and Belgium, such an ex-post nuclear tax has been agreed upon in Parliament. In conclusion, credible commitments not to raise the tax ex post are difficult. The remainder of this section therefore considers multiple cases, with different levels of commitment.

Subsection 4.4.1 first analyzes the effects of the tax $\tilde{\tau}_{sr}$ (from equation (4.12)) on investment incentives, when this tax is made permanent and the government is able to commit to it. Subsection 4.4.2 investigates a situation in which the government cannot make any credible commitment at all. Subsection 4.4.3 investigates the opposite situation, in which the government is capable of making credible commitments to an optimal long-run tax rate, which would be lower than $\tilde{\tau}_{sr}$.

4.4.1 Credible commitment to the short-run tax rate

Suppose the government imposes a tax $\tilde{\tau}_{sr}$ on current nuclear power production, according to Proposition 4.1. When the current nuclear capacity expires and needs to be replaced, nuclear firms will take the tax $\tilde{\tau}_{sr}$ into account when making the reinvestment decision. In this subsection we assume that the government can credibly commit to maintaining – i.e. not increasing – $\tilde{\tau}_{sr}$ after the replacement capacity has been built.

In their long-run reinvestment decision, firms will not only take into account the short-run marginal cost c_n and the tax $\tilde{\tau}_{sr}$ but also the fixed (investment) costs f_n per unit of capacity built. Figure 4.4 illustrates the decision of the nuclear firm. The relevant marginal cost curve is now $MC_{n,lr}$, which intersects with MR_n at E instead of D. The resulting reinvestment decision can be found by replacing c_n with $c_n + \tilde{\tau}_{sr} + f_n$ in equation (4.6):

$$q_{n,reinv} = \frac{1}{2}\left(q_0 - \beta(c_n + \tilde{\tau}_{sr} + f_n) + \frac{c_0 - (c_n + \tilde{\tau}_{sr} + f_n)}{\alpha}\right) \tag{4.15}$$

Using equation (4.12) this can be simplified to:

$$q_{n,reinv} = q_{n,max} - \frac{f_n(1 + \alpha\beta)}{2\alpha} \tag{4.16}$$

which is obviously lower than the technical maximum $q_{n,max}$.[9]

With the existing fringe capacity marginal cost curve, the reduction of nuclear capacity from $q_{n,inst} = q_{n,max}$ to $q_{n,reinv}$ increases electricity prices from p_0^* to p_2^*. One can easily compute that:

$$p_2^* - p_0^* = \frac{f_n}{2} \tag{4.17}$$

i.e. the underinvestment in nuclear capacity due to the introduction of a tax according to Proposition 4.1, may lead to an increase in electricity prices equal to half the per-unit fixed (investment) cost.[10]

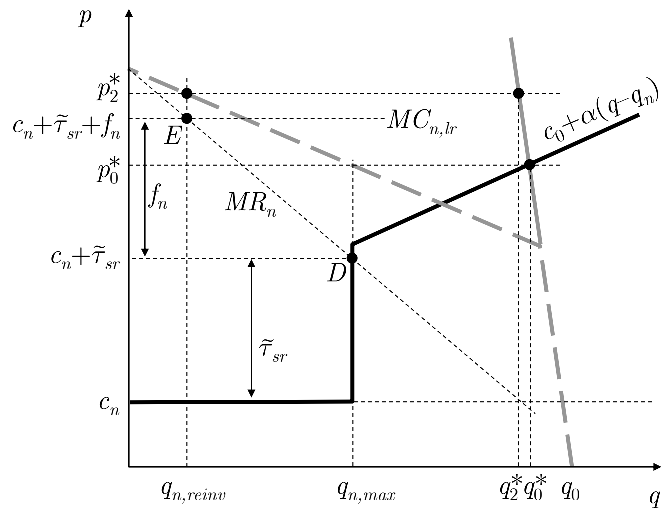
4.4.2 No credible commitment

The caveat with Figure 4.4 is that once the capacity $q_{n,reinv}$ has been installed, the government would be incentivized to increase the tax rate above $\tilde{\tau}_{sr}$. In Figure 4.4, the government could increase the tax to $\tilde{\tau}_{sr} + f_n$ without causing

[9] In the case shown in the figure, we have $f_n \leq \tilde{\tau}_{sr}$, so that in the absence of taxes even a single dominant firm would reinvest up to the the full technical maximum $q_{n,max}$. If $f_n > \tilde{\tau}_{sr}$, this need not be the case but the resulting capacity in the case with taxes is always lower than in the case without taxes.

[10] For the sake of simplicity, we have implicitly assumed that long-run price elasticity of electricity demand is the same as in the short run. In reality, in the long run, households may adapt and industries may decide to relocate to other countries, so long-run demand would be less inelastic.

Figure 4.4: Long-run impact of an unexpected tax on nuclear investments, assuming government can commit to $\tilde{\tau}_{sr}$



any withholding of the capacity $q_{n,reinv}$. In fact, by analogy with Proposition 4.1, it is easy to see that for any level of $q_{n,reinv}$ the government's optimal ex-post short-run tax $\tilde{\tau}'_{sr}$ can be found by following the MR_n curve:

$$\tilde{\tau}'_{sr} = MR_n - c_n \tag{4.18}$$

A rational nuclear firm anticipates the government's ex-post tax increases. Hence, the $MC_{n,lr}$ curve changes to include the ex-post tax increase:

$$MC_{n,lr} = c_n + f_n + \tilde{\tau}'_{sr} = f_n + MR_n \tag{4.19}$$

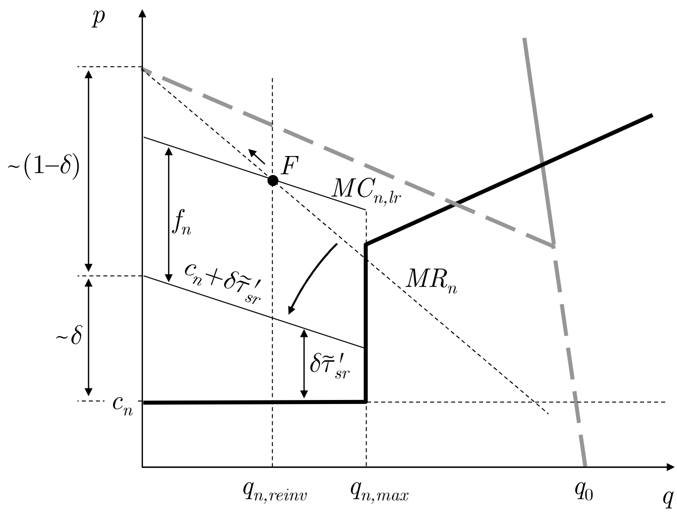
For $f_n > 0$ there is no more intersection between $MC_{n,lr}$ and MR_n, hence no nuclear capacity is built if the government cannot make any credible commitment regarding the future tax rate.

Full loss of commitment power of the government is not required to create the conditions for underinvestment. Indeed, the *possibility* of the government changing tax rates ex-post, may be sufficient to deter investment. Suppose that there is a probability δ that the government makes an ex-post decision to apply the tax $\tilde{\tau}'_{sr}$ according to equation (4.18). Conversely, with probability $(1-\delta)$, the government does not apply any tax. A risk-neutral firm will behave as if government imposed the expected value of the tax, i.e. the weighted average $\delta\tilde{\tau}'_{sr}$. This is illustrated in Figure 4.5. The nuclear firm's reinvestment decision is at point F, the intersection of MR_n and $MC_{n,lr}$, with the latter including the expected value of the tax. As δ increases, the point F moves further up the MR_n curve – as indicated with a little arrow – and reinvestment is further reduced. One can show that $q_{n,reinv}$ becomes 0 when:

$$\delta \geq 1 - \frac{f_n}{\frac{c_0+\alpha q_0}{1+\alpha\beta} - c_n} \tag{4.20}$$

Hence, reinvestment may be completely deterred even if it is not certain that the government cannot commit to a tax rate.

Figure 4.5: Long-run impact of an unexpected tax on nuclear investments, assuming a probability δ that the government cannot credibly commit to a tax rate



4.4.3 Credible commitment to the optimal long-run tax rate

Let us assume that the government can choose a tax rate before the nuclear firm's reinvestment decision, and that the government can make a credible commitment to maintain this tax rate after the capacity has been built. By analogy with Proposition 4.1, essentially by replacing c_n with $c_n + f_n$, we can derive the government's optimal choice of the tax rate. However, since f_n may be large, the analog of condition (4.11) is not necessarily fulfilled, hence two cases need to be considered.

For the sake of analytical simplicity, we will assume $\sigma_n = 0$ in this section.

Proposition 4.2. *Under the conditions of Section 4.2 and assuming $\sigma_n = 0$ and $\beta \ll \alpha^{-1}$, the optimal choice of the nuclear tax rate for a government capable of making credible commitments is:*

$$\tilde{\tau}_{lr} = \max\left\{\tilde{\tau}_{sr} - f_n\,,\ \frac{c_0 - c_n - f_n}{2} + \frac{\alpha\beta}{4}\left(\alpha q_0 - \frac{c_0 - c_n - f_n}{2}\right)\right\} \qquad (4.21)$$

Proof. The proof is analogous to Proposition 4.1. For $\tau \leq \tilde{\tau}_{sr} - f_n$, the constraint $q_{n,reinv} \leq q_{n,max}$ is binding, and lowering τ below $\tilde{\tau}_{sr} - f_n$ cannot increase W.

For $\tau > \tilde{\tau}_{sr} - f_n$ the constraint is not binding anymore. In the absence of the constraint, the government's objective function W has a peak at τ_{peak} as given by equation (4.14) with c_n replaced by $c_n + f_n$.

If $\tau_{peak} \leq \tilde{\tau}_{sr} - f_n$ then W declines when τ is raised above $\tilde{\tau}_{sr} - f_n$, similar to the situation in Figure 4.3. If $\tau_{peak} > \tilde{\tau}_{sr} - f_n$, then the maximum of W is reached for $\tau = \tau_{peak}$. Combining both cases, we find $\tilde{\tau}_{lr} = \max\{\tilde{\tau}_{sr} - f_n\,,\ \tau_{peak}\}$, hence equation (4.21). □

Note that $\tilde{\tau}_{lr}$ according to equation (4.21) is not necessarily positive. In general, it is easy to see that large f_n will make $\tilde{\tau}_{lr}$ small or negative. The government therefore faces a trade-off: on the one hand, short-run welfare maximization would require taxing current nuclear capacity at a rate $\tilde{\tau}_{sr}$, but this would harm government credibility thereby hampering reinvestment. On the other hand, the optimal long-run tax rate $\tilde{\tau}_{lr}$ in this setting is rather low. If, in order to preserve its credibility, the government decides to apply $\tilde{\tau}_{lr}$ also to current capacity, this would leave a very large part of current nuclear rents untouched. Ideally, the government would be able to make a credible commitment to distinguish its tax rate on current capacity from its tax rate on new capacity, but this may be difficult to do.

4.5 Alternative policy instruments

The previous section has highlighted the challenges of imposing an unexpected short-run tax on a dominant nuclear firm. In this section we explore alternative policy measures.

4.5.1 Lifetime extension agreements

Belgium's government first decided to extend the lifetime of the country's nuclear power plants. Only later, it imposed a contested tax on nuclear producers. An alternative is the German approach, in which a tax was agreed between government and the nuclear firms in return for a lifetime extension.[11]

[11] The lifetime extension was later revoked in the aftermath of the Fukushima accident, which is, in fact, an illustration of ineffective government commitment.

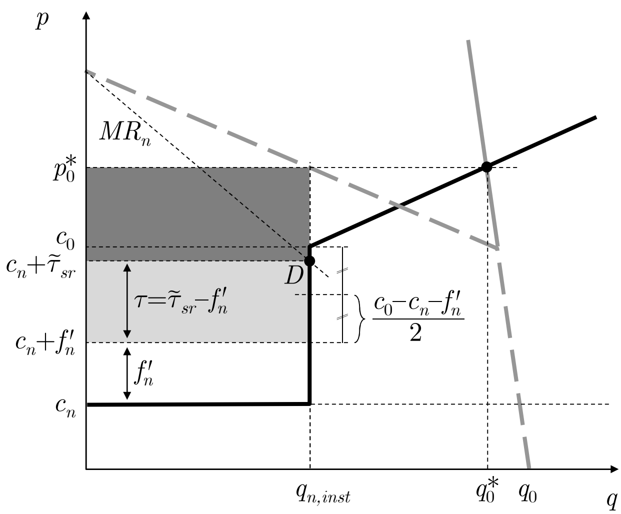
The latter option is arguably preferable from the perspective of government credibility and investment incentives.

We can analyze such a lifetime extension agreement using the reasoning of Section 4.4.3. Indeed, lifetime extension is very similar to reinvestment, but the reinvestment cost f_n is replaced by a significantly lower refurbishment cost f'_n. We simplify the bargaining process between the government and the nuclear industry by assuming a Stackelberg structure in which the government credibly commits to a tax rate τ in the first stage, and a dominant nuclear firm responds in a second stage by deciding on how much of the nuclear capacity will have its lifetime extended. The optimal tax to be set in return for a lifetime extension can then be derived from Proposition 4.2. For the sake of simplicity, we assume $\sigma_n = 0$, so that the optimal tax rate is given by Propostion 4.2. If we assume $\alpha\beta$ to be negligible, then the optimal tax rate becomes:

$$\tau = \max\left\{\tilde{\tau}_{sr} - f'_n\,,\ \frac{c_0 - c_n - f'_n}{2}\right\} \tag{4.22}$$

which can be conveniently analyzed graphically, as in Figure 4.6, where $(c_0 - c_n - f'_n)/2 < \tilde{\tau}_{sr} - f'_n$, hence $\tau = \tilde{\tau}_{sr} - f_n$. We shall maintain this assumption in this and the next section.

Figure 4.6: Taxation in return for lifetime extension



Because $f'_n \ll f_n$, the agreed tax τ allows the government to capture at least a sizeable part of the nuclear lifetime extension rent $(p_0^* - c_n - f'_n)q_{n,inst}$. The part captured by the government is indicated in light shading in Figure 4.6. However, a large part of the rent – shown in dark shading in Figure 4.6

– cannot be captured through a lifetime extension and taxation agreement with a dominant nuclear firm, because a further increase of the tax rate would trigger a reduction of the amount of capacity of which the lifetime is extended, and therefore lead to lower tax revenues.

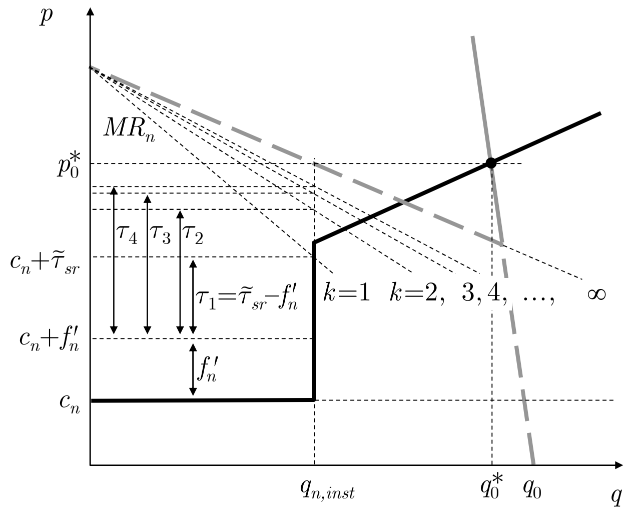
4.5.2 Auctioning

The main underlying reason why a large part of the nuclear rent cannot be captured in the setting of Figure 4.6, is the fact that the nuclear firm is assumed to be a monopolist on the residual demand curve. One way to increase the share of rent captured by taxes – and thus increase welfare – is to introduce competition. More specifically, the government could open up the lifetime extension licensing not only to the incumbent dominant firm but also to other players, in an auction.

Suppose there are k players, including the incumbent. For the sake of simplicity, we model the auctioning process as a Cournot game between nuclear firms. First the government sets a lifetime extension tax τ_k, as in Section 4.5.1. Then, in Cournot competition, the k players each decide on an amount of the capacity to be extended. In the previous sections, the response of the dominant firm as a function of costs plus taxes, was always given by MR_n, which, as is well-known, has twice the slope of the inverse residual demand curve. In a Cournot game the resulting total quantity of the k players, as a function of costs plus taxes, is a similar line, but with a slope of $(k+1)/k$ times the slope of the inverse residual demand curve. This is shown in Figure 4.7. The reasoning from Section 4.5.1 can now be applied almost identically. The optimal tax rate τ_1 in the case of one dominant firm is the same as in Section 4.5.1. The optimal tax rates τ_k for $k > 1$ become larger and larger, thereby allowing the government to capture an ever larger share of the nuclear rent. In the limit case of perfect competition ($k \to \infty$) the tax becomes $\tau_\infty = p_0^* - c_n - f_n'$, and the government captures the entire nuclear rent. There may however be some drawbacks to this approach. In particular, the auctioning process may suffer from asymmetric information. Indeed, the nuclear firm currently operating the power plant may have private information that is unavailable to other potential contenders in the auction. Furthermore, if ownership changes as a result of the lifetime extension auction, the power plant may continue to be operated by the original owner, which could lead to principal-agent problems.

In this section, we have applied the auctioning mechanism to the lifetime extension process. Clearly, the same auctioning mechanism can be applied to the reinvestment process.

Figure 4.7: Auctioning of lifetime extension licenses



4.5.3 Other taxation instruments

In this paper, the optimal nuclear energy tax has been derived for the case of an *excise* tax on nuclear production. This is a fixed amount per MWh of nuclear energy produced. Alternatively, the government could opt for an ad valorem tax, a profit tax, or a lump sum tax. The former two taxes are in fact similar in structure to the excise tax we have considered so far. In theory, in the case of a pure monopoly, an *ad valorem* tax – proportional to the nuclear firm's revenues – yields a larger nuclear output for the same tax revenue than an excise tax.[12] In our case however, the difference will be small since c_n is small. A special *profit* tax – proportional to the nuclear firm's profits – circumvents the problem of withholding, since the nuclear firm's incentives remain unaltered. However, it leads to the same long-term investment problems as the excise tax discussed so far. Moreover, the special nuclear profit tax may be difficult to implement: firms may attempt to evade the nuclear profit tax by inflating costs, shifting costs between non-nuclear and nuclear plants, or overstating the allocation of overhead costs to nuclear subsidiaries. From a legal perspective, it may be challenging to tax the same profit twice, i.e. both through regular corporate taxation and through the special nuclear profit tax. In any case, the profit tax bears some resemblance to the excise tax, because also the optimal excise tax depends on the difference between the price of the fossil fuel alternative (via c_0) and the marginal cost of nuclear power (c_n). A pure *lump sum* tax,

[12] See e.g. Salanié (2003).

set independently of nuclear output, is a superior instrument compared to the excise tax. However, in many countries lump sum taxes would not be constitutional as they are not linked to a defined tax base.

4.5.4 Mandatory investments in renewable energy

Instead of pure taxation, governments are also considering the possibility of requiring nuclear producers to invest a part of the nuclear rent in renewable energy. For instance, the annulled German lifetime extension agreement included a compulsory contribution to a renewable energy fund. In fact, such a scheme is conceptually equivalent to the combination of a tax plus a commitment to invest the tax revenues in renewable energy. Hence, it is only optimal if renewable energy investments are indeed the best use of tax money, among all other possible government investment options. This seems unlikely, unless one assumes that other government expenditures are generally wasted, i.e. λ in equation (4.9) very small.

4.6 Numerical simulations for the case of Belgium

4.6.1 Calibration and baseline

In this section we apply the results from previous sections to a numerical simulation for the case of Belgium. Table 4.2 lists the calibration parameters. More information on how these values are obtained, can be found in 4.A. We assume $\lambda = 1$ as before. Unless specified otherwise, we assume $\sigma_n = 0$, i.e. there are no local shareholders of the nuclear firm, or the government does not include their interests in its objective function.

Table 4.2: Numerical values of parameters for the case of Belgium

Parameter	**Value**	**Unit**
q_0	9422	MW
β	0	MW · MWh/EUR
$q_{n,max}$	5345	MW
c_n	20.1	EUR/MWh
f_n	34.7	EUR/MWh
f'_n	6.9	EUR/MWh
c_0	39.3	EUR/MWh
α	0.0053	EUR/MWh / MW

Since demand is assumed to be completely inelastic ($\beta = 0$), the equilibrium quantity in the absence of any intervention is obviously $q_0^* = 9422$

MW. The equilibrium price in the absence of any intervention is $p_0^* = 60.9$ EUR/MWh, a fairly realistic number.

The total rent obtained by the nuclear firm in this model is given by:

$$(p_0^* - c_n)q_n = 1910 \text{ million EUR} \tag{4.23}$$

which is in line with the estimate of 1.75 to 2.3 billion EUR made by the Belgian electricity regulator CREG (2011). It should be emphasized that the estimation of the nuclear rent is not the objective of our model. Rather, the similarity with CREG (2011) indicates that the calibration of the model is likely to be fairly realistic.

4.6.2 Taxation and the effect on new investments

The parameters of Table 4.2 fulfill condition (4.11). The optimal unanticipated short-run tax according to Proposition 4.1 is:

$$\tilde{\tau}_{sr} = 12.5 \text{ EUR/MWh} \tag{4.24}$$

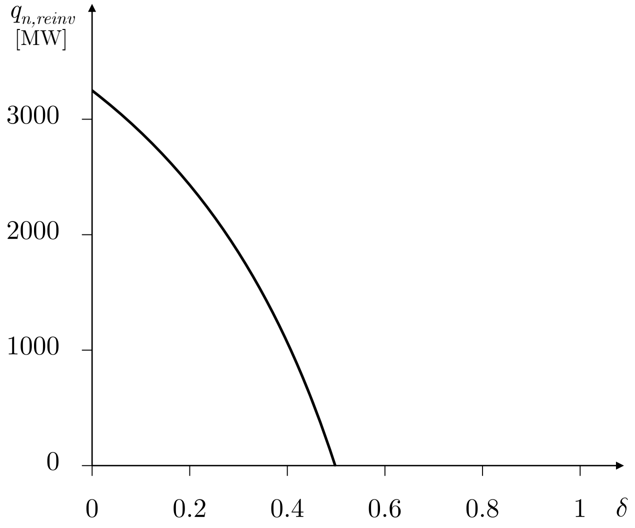
and the corresponding government revenues G are 584 million EUR per year. This tax is quite a bit higher than the 5 EUR/MWh tax introduced by the Belgian government in 2008, but it still captures only 31% of the total rent.

In the longer run, if the 12.5 EUR/MWh tax is credibly maintained for new nuclear investment when current capacity has expired, then equation (4.16) indicates that only 39% of nuclear capacity will be replaced. This is obviously dependent on the assumption that there is only one nuclear firm that can decide to invest in Belgium. The resulting electricity price increase according to equation (4.17) would be 17.4 EUR/MWh. Note that in this dominant firm – competitive fringe model, with only one nuclear firm, the nuclear firm would not replace all nuclear capacity even if no tax is imposed (see footnote 9). Indeed, even in the absence of any taxes, the dominant nuclear firm would only reinvest in 61% of capacity. Still, the 12.5 EUR/MWh tax reduces reinvestment by more than a third.[13]

If no taxes are imposed on new nuclear capacity but there is a probability δ that the government will impose a short-run optimal tax ex-post, new nuclear investment may also be severely reduced, as explained in Section 4.4.2. Figure 4.8 shows how the amount of replacement capacity $q_{n,reinv}$ changes as a function of δ. According to equation (4.20), as soon as $\delta \geq 0.50$, there will be no more nuclear investment. In other words: in this dominant firm – competitive fringe model, a 1-in-2 chance of an ex-post nuclear tax, of which the level will be decided by the government after investment, is sufficient to deter all investment.

[13] An obvious question is then why the current capacity $q_{n,max}$ has been built in the first place, given that a monopolist would build less capacity according to our model. One likely explanation is that the capacity was built during the time of regulated monopoly, in which the incumbent electricity firm would operate in a cost-plus scheme, and not on a free market as in our model.

Figure 4.8: Reinvestment in nuclear capacity as a function of the probability δ that government will impose an ex-post tax



Let us now consider the case in which the government can set a long-term tax rate $\tilde{\tau}_{lr}$ for new investment according to Proposition 4.2 and commit to it. Figure 4.9 shows the optimal tax $\tilde{\tau}_{lr}$ as a function of σ_n, the share of local shareholders in the nuclear firm. The optimal tax is the maximum of two alternatives: $\tilde{\tau}_{lr} - f_n$ and τ_{peak}, as explained in Proposition 4.2. Interestingly, the optimal tax is negative, even when σ_n is very low. Since the dominant firm has an incentive to reinvest less than the technical maximum, it is optimal for the government to *subsidize* the construction of nuclear power, in order to prevent underinvestment and too low production. The subsidy is 7.8 EUR/MWh when $\sigma_n = 0$, and increases with σ_n up to a maximum of 22.2 EUR/MWh for $\sigma_n = 0.51$, after which it remains constant because the maximum possible reinvestment $q_{n,max}$ is reached.

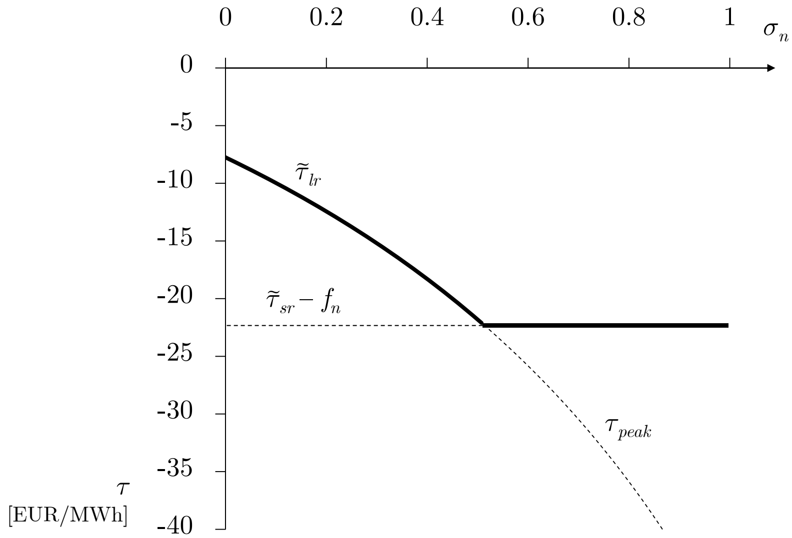
4.6.3 Lifetime extension agreement and auctioning

Let us now analyze the case of a lifetime extension agreement between the government and the nuclear firm. As mentioned in Section 4.5.1 the problem is similar to the question of reinvestment treated in the previous section, but with a lower cost f'_n instead of f_n. Applying equation (4.22) we find:

$$\tilde{\tau}_{sr} - f'_n = 5.6 \text{ EUR/MWh} \tag{4.25}$$

$$\frac{c_0 - c_n - f'_n}{2} = 6.2 \text{ EUR/MWh} \tag{4.26}$$

Figure 4.9: Optimal long-run tax when the government can credibly commit, as a function of σ_n



hence the optimal tax is 6.2 EUR/MWh. The situation is slightly different from what is depicted in Figure 4.6 in that the government will increase the tax rate slightly above the point where the nuclear firm starts to withhold capacity from the lifetime extension. This effect disappears when multiple nuclear firms are invited to an auction of lifetime extension licenses, as in Section 4.5.2. With more than one player, in this numerical setting, the government will set exactly the maximum tax rate that will make sure the lifetime of all capacity gets extended. Figure 4.10 shows the optimal lifetime extension tax as a function of the number of players k in the auction. The total potential rent in this case is 1588 million EUR, which is less than the nuclear rent of 1910 million EUR mentioned earlier because of the lifetime extension cost f'_n. With one player, the above-mentioned tax of 6.2 EUR/MWh captures only 18% of this rent. With more than one player, the taxation potential rapidly increases with k: in an auction with three players, the government would set a tax of 24.5 EUR/MWh, capturing 72% of the maximum potential. In a perfectly competitive setting, i.e. when $k \to \infty$, the government can capture the full rent with a tax of 33.9 EUR/MWh.

As mentioned before, the same principle can be applied to the auctioning of new nuclear power plant licenses, if the government can credibly commit to an optimal long-run tax rate. The total potential rent is much smaller, because $f_n \gg f'_n$. In this numerical simulation, the total rent would be 286 million EUR per year. However, a large number of players would be required in order for the government to capture a significant share of this rent, as shown in Figure 4.11. For $k = 1$ and $k = 2$, we obviously find again the subsidy of 7.8 EUR/MWh shown earlier in Figure 4.9. For higher k the subsidy decreases, and becomes a tax when $k \geq 5$. In the fully competitive situation in which $k \to \infty$, the tax is 6.1 EUR/MWh and the entire rent is captured by the government.

The main numerical results are summarized in Table 4.3.

4.7 Conclusions

In this paper we have studied nuclear taxation using a stylized model of the electricity sector with one dominant nuclear producer and a competitive fringe of fossil-fuel plants. The graphical and analytical results are illustrated using a numerical simulation for the case of Belgium.

We find that an unanticipated tax on nuclear production can generate significant government revenues in the short run without disturbing the equilibrium on the electricity market. In the simulation, the optimal short-run tax is 12.5 EUR/MWh and captures around 31% of the total nuclear rent. However, the tax may harm reinvestment incentives in new nuclear capacity in the long run. Assuming the government commits to maintaining the short-run tax rate in the long run, reinvestment is reduced by a third in the simulation, compared to a situation without the tax. If the probability that the government cannot commit to a tax rate is higher than 50%, reinvestment is completely deterred. If, on the other hand, the government can credibly commit to an optimal long-run tax rate, government revenues would be very low because the socially optimal tax would be very small or negative, due to the market power of the nuclear firm.

An agreement on lifetime extension between the government and a dominant nuclear producer generates less revenues than the unanticipated tax (18% of the lifetime extension rent in the simulation). However, by inviting multiple competing bidders for the lifetime extension licenses, government revenues increase rapidly: in the simulation, a lifetime extension negotiation with three bidders captures 72% of the rent for the government. Likewise, inviting multiple players to bid for reinvestment in new nuclear capacity can increase the potential revenues from reinvestment in new capacity because it can make the socially optimal tax positive.

Government credibility has been proven to be crucial for enabling long-run optimal nuclear taxation policy. One way to achieve such credibility is to transfer some authority to a supranational body, so that appeals are possible when taxation agreements are not honored. The sensitive nature of taxation may make such a transfer difficult to realize in practice, however.

Our stylized representation of the electricity market of a country with nu-

Figure 4.10: Lifetime extension tax as a function of the number of participants in the auction

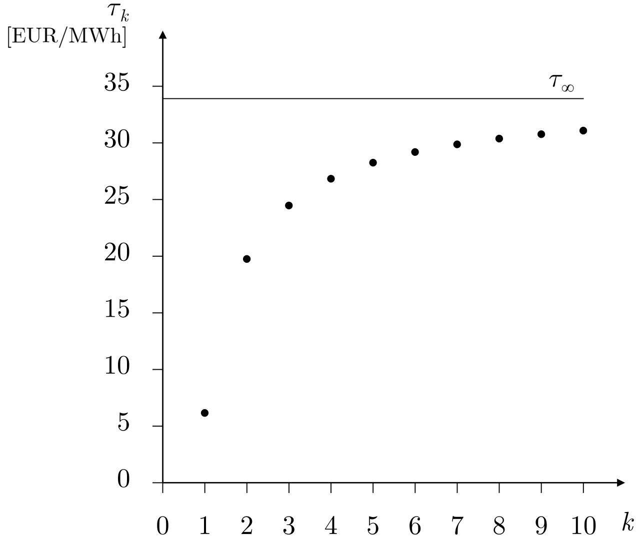


Figure 4.11: Long-term tax on new nuclear investments as a function of the number of participants in the auction

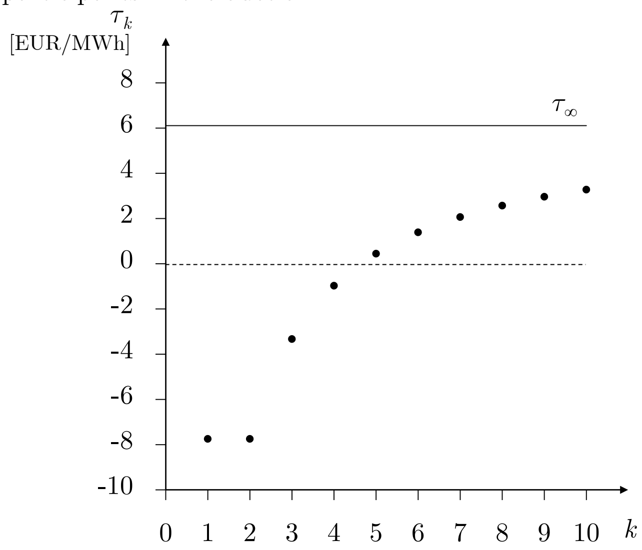


Table 4.3: Summary of numerical results

Situation	**Scenario**	**Tax** EUR/MWh	**Comment**
Current nuclear power plants	Optimal unanticipated short-run tax	12.5	31% of rent captured, but risk of harming reinvestment
Reinvestment in new nuclear power plants	Commitment to same short-run tax	12.5	A third less investment than without tax
	No government commitment (or more than 50% chance of short-run tax ex-post)	/	No new nuclear investments
	Commitment to optimal long-run tax, with one nuclear firm	-7.8	Negative tax, i.e. subsidy, to mitigate monopoly power
	Commitment to optimal long-run tax, with perfect competition for nuclear licenses	6.1	Complete rent captured by government
Lifetime extension of old nuclear power plants	Negotiation with incumbent only	6.2	18% of total rent captured
	Auctioning with three players	24.5	72% of total rent captured
	Perfect competition for licenses	33.9	Complete rent captured by government

clear power could be further refined. In particular, it would be useful to enhance the integration of international imports in our model, as these may be quite important in the event of large underinvestment. Furthermore, since our analysis shows that multi-party negotiation of lifetime extension agreements and auctioning of new nuclear licenses seem to be the most attractive policies, further research on the details of such auctioning processes would be beneficial.

4.8 Acknowledgements

The authors would like to thank the participants at the conference on "The Economics of Energy Markets" at Toulouse School of Economics in January 2010 – in particular the discussant N. Ladoux – the ETE seminar in Leuven in March 2010, and the conference in honor of Yves Smeers at CORE-UCL in June 2010. The usual disclaimers apply.

4.A Annex: Industry marginal cost curve for the Belgian electricity sector

We estimate coefficients c_0 and α based on a cost curve of Belgian electricity supply for the year 2010. The Belgian Transmission System Operator *Elia* provides data on available total power per 15-minute time slice. Average available total power in 2010 was 11927 MW (Elia, 2011). Table 4.4 provides a breakdown of available capacity by fuel.[14] Elia (2011) also provides the available amounts of international import transmission capacity from neighboring countries to Belgium, shown in Table 4.5. The latter capacities are relevant, since the fringe cost curve $c_0 + \alpha q$ should also include potential imports, as mentioned in Section 4.2.1.

Data on efficiencies, emissions and maintenance costs per technology are taken from the European Commission (2008) and summarized in Table 4.6. Coal, gas and oil[15] prices are based on the average price over the period 2006-2010 according to BP (2011). Nuclear fuel price is taken from the European Commission (2008), and includes provisions for waste management. Table 4.7 provides an overview.

In order to construct the industry marginal cost curve, we make a number of additional assumptions:

[14]Note that the available nuclear capacity is less than in Table 4.1, because not all capacity is available at any given time. Overall, average total available capacity is much lower than total installed capacity, which was 17084 MW in 2010 according to Elia (2011).

[15]For simplicity we do not distinguish crude oil from the various refined products used in power generation. Since oil-based generation is anyhow located at the far right-hand side of the cost curve, this does not significantly affect results.

Table 4.4: Average available power generation capacity by fuel

Fuel	**Average available capacity** MW, 2010
Nuclear	5345
Gas	3133
Hydro	1480
Coal	817
Oil (Fuel)	285
Wind	159
Other	753
Total	**11972**

Source: Elia (2011).

Table 4.5: International import transmission capacity from neighboring countries to Belgium

Country	**Net Transfer Capacity – NTC** MW, 2010
France	1700
Netherlands	830
Total	**2530**

Source: Elia (2011).

- The vast majority of hydropower capacity in Belgium is pumped storage, which does not make a net contribution to the power supply. The 1480 MW of hydropower in Table 4.4 is therefore excluded from the analysis.

- The available capacities for Wind and Other in Table 4.4 are assumed to be non-dispatchable capacity, i.e. they do not run as a function of demand but as a result of another constraint (wind, cogeneration of heat and power, etc.). We therefore subtract them from demand (using a 50% load factor) and do not include them in the cost curve.

- Older coal plants typically have lower efficiencies than the Best Available Technology (BAT) efficiencies listed in Table 4.6. For the purpose of the coal part of our cost curve, we therefore use the least favorable characteristics among the two coal technologies listed in Table 4.6.

- Imported generation from the Netherlands is assumed to be gas-based.

- Imported generation from France is assumed to be coal-based. Although France has very large nuclear capacity, our model does predict that this

Table 4.6: Techno-economic characteristics of power plant technologies

Technology	**Typical characteristics**		
	Efficiency	**CO_2 Emissions**	**Maintenance costs**
	Percent	kg/MWh	EUR/MWh
Nuclear – Fission	35%	0	12.1
Coal – PCC	47%	725	8.1
Coal – CFBC	40%	850	9.4
Gas – CCGT	58%	350	3.4
Gas – GT	38%	530	5.4
Oil – CC	53%	505	6.7

Source: European Commission (2008).

Table 4.7: Fuel prices

Fuel	**Reference**	**Price** EUR/MWh(thermal)
Coal	Northwest Europe marker 2006-2010	8.3
Gas	European Union cif 2006-2010	21.9
Oil	Brent (dated) 2006-2010	32.2
Nuclear	Price cited by European Commission (2008)	2.8

Source: BP (2011), European Commission (2008).

capacity is already fully utilized, due to its low marginal costs (see Section 4.2.2). On average, it would therefore not be able to make an incremental contribution to serving Belgian load. A similar argument can be made for France's large hydropower capacity. The next largest technology in France's generation system is coal (Platts, 2010).

- To take into account the additional transmission costs incurred by importing power from neighboring countries, a flat fee of 12.0 EUR per MWh is added to the cost of imported power. This value corresponds to the average revenue of Elia (2011) per MWh of load in Belgium.

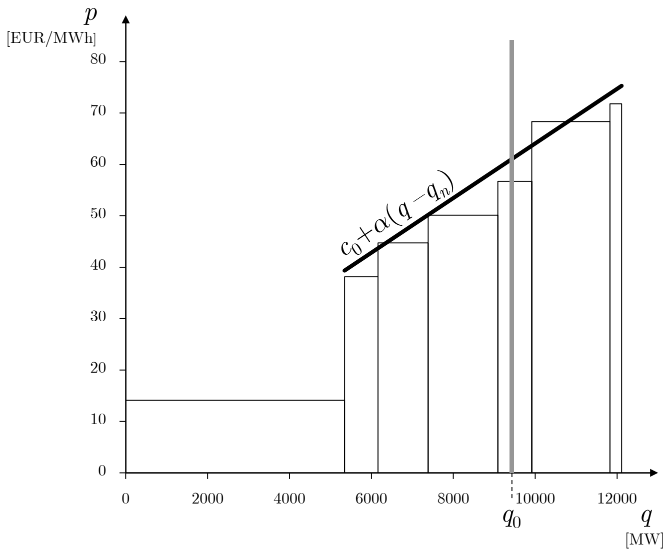
- The maintenance costs cited in Table 4.6 comprise both fixed (FOM) and variable (VOM) costs. For the industry marginal cost curve, only the variable part should be included. We assume this variable part is 50% of the maintenance cost.

- The carbon emissions allowance price is assumed to be 15 EUR per tonne.

The resulting industry marginal cost curve is shown in Figure 4.12. Based on this cost curve, one can estimate c_0 = 39.3 EUR/MWh and α = 0.0053

EUR/MWh/MW. The figure also shows the net demand, which, after subtraction of the Wind and Other capacities, is $q_0 = 9422$ MW. Demand is assumed to be completely inelastic: $\beta = 0$.

Figure 4.12: Simplified industry marginal cost curve for the Belgian electricity sector



The fixed capital charge f_n for nuclear power is estimated using the capex estimate provided by the European Commission (2008), but with a lower discount rate: 7%, which is in line with the pre-tax weighted average cost of capital of a large European utility (see e.g. E.ON, 2009), instead of the 10% discount rate suggested by the European Commission (2008). The result is $f_n = 34.7$ EUR/MWh. For the variable cost c_n of nuclear power, we include not only the marginal cost of 14.1 EUR/MWh, as shown in Figure 4.12 but also the FOM of 6.0 EUR/MWh, bringing the total to $c_n = 20.1$ EUR/MWh. Indeed, although the industry marginal cost curve – as used for pricing – does not include the FOM, the FOM will be relevant for the nuclear firm in deciding whether or not to withhold some of its nuclear capacity over the timeframe of a year. Finally, we assume that the fixed capital charge for lifetime extension is 20% of f_n: $f'_n = 6.9$ EUR/MWh.

Chapter 5

International Transport of Captured CO_2: Who Can Gain and How Much?

5.1 Introduction

Fossil fuels are likely to remain the main source for electricity generation in Europe, at least in the short to medium term, despite the significant ongoing efforts to promote renewable energy technologies and energy efficiency. CO_2 Capture and Storage (CCS) is generally considered as a promising technological option for reducing CO_2 emissions from the power generation sector, as well as from other heavy industries. CCS is a process consisting of the separation of CO_2 from industrial and energy-related sources, transport to a storage location (such as a depleted hydrocarbon field or a saline aquifer) and long-term isolation from the atmosphere (see e.g. IPCC, 2005). CCS may offer a bridge between the fossil fuels dependent economy and the carbon-free future.

Large-scale deployment of CCS in Europe may require the development of an international pipeline network to transport the captured CO_2 from its sources (e.g. power plants) to the appropriate CO_2 storage sites. This paper estimates the magnitude of the benefits associated with international cooperation in the development of the pipeline network, compared to a situation in which countries take individual action. More importantly, this paper uses cooperative game theory to describe how the gains from such cooperation could be distributed among participating countries. Equivalently, the paper answers the question how the costs of an international CO_2 network would be allocated to the participating countries. In particular, we study how the allocation depends on EU legislation for CO_2 pipelines, by considering two possible policy scenarios: national pipeline monopolies on the one hand, and full liberalisation on the other hand.

CCS figures quite prominently in EU energy and climate policies. The EU's Energy Roadmap 2050 (European Commission, 2011) contains 7 scenarios up to 2050, and on average these scenarios project 133 GW of installed CCS power generation capacity by 2050. Such a power plant fleet would correspond to around 1 Gt/y of CO_2 being transported from sources (power plants) to sinks, which would require a network of the extent described in this paper. Despite the prominence of CCS in EU energy system projections, the acceptance of CCS is still low in many countries, hence large-scale cooperation still seems challenging. The policy relevance of this paper is first of all that it points out which monetary transfers may be needed in order to achieve cooperation. Secondly, the paper assesses the effect of EU CO_2 pipeline regulation options on the cooperation game between European countries.

The paper is centered on a case in which there is a need for a large-scale international CO_2 pipeline network. In this context, it needs to be mentioned that there are different views on how the CO_2 transport infrastructure might evolve in Europe. On the one hand, there is often a perception that CCS plants will be built very close to potential storage sites in order to minimise transport costs. On the other hand, proposals for CCS projects that have become public tend to show that their location is dictated by other factors, such as safety and public acceptance concerns that may require that CO_2 is initially stored offshore; or the presence of old power plants that are suitable for retrofitting or refurbishing with CO_2 capture technologies. Furthermore, the large-scale deployment of CO_2 capture facilities in Europe, which would be needed to achieve the decarbonisation of the European energy system by 2050,[1] combined with the fact that CO_2 storage sites and capacities are not uniformly distributed across Europe, will quickly exhaust local storage opportunities and necessitate the construction of an extended transport infrastructure, which will span across national borders when countries do not have adequate domestic CO_2 storage capacity.

The evolution of the CO_2 transport network in Europe will be dictated by the level of CCS deployment and the degree of coordination for its development. The simplest approach for the development of the CO_2 transport infrastructure would be the construction of numerous pipelines linking individual CO_2 sources with sinks, sized to meet the transport needs of individual capture facilities. This implies that pipelines will be constructed in the context of individual CCS projects and their planning and construction will be synchronous to the development of the CO_2 capture facilities. This approach is however likely to impede the large-scale deployment of CCS as it will not allow for the expansion and sharing of the infrastructure with other CO_2 sources, which in

[1] For instance, a scenario such as *Power Choices* (Eurelectric, 2010), which is in line with the EU's 80% to 95% greenhouse gas emissions reductions targets (as repeated in the recent communication by the European Commission, 2010a), projects more than 1 Gt/y of CO2 captured in the EU by 2050.

turn will be required to develop their own pipelines, resulting in deployment delays due to permitting procedures, and additional costs, since pipeline costs do not scale proportionally with transport capacities. Apparently, this situation would be most detrimental for CO_2 sources that are either of small size or located away from suitable storage sites. Alternatively, the development of integrated pipeline networks, planned and constructed initially at regional or national level and oversized to meet the transport needs of multiple CO_2 sources would take advantage of economies of scale and enable the connection of additional CO_2 sources with sinks in the course of the pipeline lifetime. For example the Pre-Front End Engineering Design Study of a CCS network for Yorkshire and Humber (CO2Sense, 2010) showed that initial investment in spare pipeline capacity would be cost-effective even if subsequent developments were not to join the network for up to 11 years. The study also confirmed experience from other sectors i.e. that investing in integrated networks would catalyse the large scale deployment of CCS technologies by consolidating permitting procedures, reducing the cost of connecting CO_2 sources with sinks and ensuring that captured CO_2 can be stored as soon as the capture facility becomes operational. In the longer run, such integrated networks could be expanded and interlinked to reach CO_2 sources across Europe and distant storage sites, leading to the development of a true trans-European network, similar to the existing ones for electricity and gas. A recent communication from the European Commission (2010b) points out that the realisation of such CO_2 transport infrastructure would require a timely start of coordinated infrastructure planning and development at European level.

Recent research has produced a number of models that are capable of determining the optimal (i.e. cost-minimising) CO_2 transport network that can transport CO_2 from sources to sinks, such as Middleton and Bielicki (2009), Broek et al. (2010a,b), Mendelevitch et al. (2010) and Morbee et al. (2010). These studies, however, do not describe how the necessary coordination to achieve such optimal infrastructure would be realised. The case studies investigated by e.g. Middleton and Bielicki (2009) and Broek et al. (2010a) are focused on single countries or states, where coordination may be relatively feasible. However, the trans-European networks described by e.g. Mendelevitch et al. (2010) and Morbee et al. (2010) require coordination and joint pipeline infrastructure investment by a large number of countries. The question we study in this paper is how such international cooperation could be structured in order to achieve the benefits of joint infrastructure optimisation.

In particular, our paper aims to study how the gains from coordination can be allocated between countries in order to ensure participation in the joint coordination. We analyse the allocation by means of the *Shapley value* concept from cooperative game theory. Game theory has already been applied to energy networks by e.g. Hobbs and Kelly (1992), who apply cooperative models to short-run electricity transmission games and a dynamic non-cooperative Stackelberg game to long-run electricity transmission capacity decision games.

Our model, by contrast, applies cooperative game theory to the capacity decision. Kleindorfer et al. (2001) provide an overview of strategic gaming in power markets, but do not focus on transmission infrastructure. Csercsik and Koczy (2011) study transmission networks – including expansion games – using cooperative game theory in a load flow model of the electricity system, and apply it to a stylised 5-node network. Our analysis is less detailed on the technical side (CO_2 transmission is treated here as a simple transport model) but the methodology is applied to an extensive European case study. As mentioned before, the focus of our study is on allocation of the benefits of cooperation, or equivalently, allocation of the network investment cost. Again for the case of electricity, cost allocation has been studied by e.g. Contreras and Wu (2000) and Evans et al. (2003), who use a Kernel approach from cooperative game theory. Gately (1974) provides a game-theoretic analysis of the distribution of the benefits of cooperation in electrical power investments between three regions in the Southern Electricity Region of India. The study considers 5 possible partitions of players, comprising 7 possible coalitions, and compares several game-theoretic methods for distributing the gains from cooperation, with Shapley value and Kernel as some of the possible options. Our model, by contrast, needs to consider a more complex game with 18 players, hence 262143 possible coalitions and 682 billion possible partitions. As a result, our analysis of the strategic game is less extensive, and considers only the Shapley value as a possible allocation.

As is well-known, the Shapley value is an approach for 'fair' allocation of gains from cooperation among participating actors. It has been applied to natural gas by e.g. Hubert and Ikonnikova (2009) and Ikonnikova and Zwart (2010), and to CO_2 emissions by e.g. Albrecht et al. (2002). In the context of CO_2 pipeline networks, the Shapley value determines the bargaining power of individual countries in the international negotiation on CO_2 infrastructure investment, and hence the allocation of the cost burden. The bargaining power of each country depends on how easily the country can be circumvented. As can be intuitively expected, we will find that countries with large storage potential are likely to receive net benefits from the negotiation, while countries with large excess CO_2 quantities (i.e. CO_2 captured which cannot be stored domestically) make a net contribution. Furthermore, the Shapley value shows that countries with a strategic transit location are able to extract some (but limited) rent from their position. In an alternative scenario, in which CO_2 pipeline construction is liberalised and not subject to national monopolies, such transit rents disappear and the bargaining power of countries with excess storage potential increases further.

The paper is structured as follows. First, Section 5.2 describes the potential structure of a trans-European CO_2 transport network, and the potential benefits obtained from international coordination. Section 5.3 describes our game theoretic solution concept. Section 5.4 apply this solution concept to CO_2 infrastructure negotiations, under two scenarios: one scenario with national CO_2 transport monopolies, and one scenario with liberalised pipeline construction. Section 5.5 summarises our conclusions.

5.2 International coordination of CO_2 pipeline networks

The starting point of our investigation is a projection of the optimal CO_2 pipeline network in Europe in 2050. We assume that the European power system evolves according to the *Power Choices* scenario (Eurelectric, 2010).[2] The *Power Choices* scenario, which is based on the PRIMES model, is chosen for this purpose because it is in line with the EU's 80% to 95% greenhouse gas emissions reductions targets by 2050 (implying near-complete decarbonisation of the power sector), and hence provides a view on large-scale pan-European deployment of CCS in the power sector. The scenario implies a reduction of CO_2 emissions from the power sector to 150 Mt/y by 2050, compared to 1423 Mt/y in 2005. This is achieved through more than 40% electricity production from renewable energy sources (RES), close to 30% of nuclear power, and the remaining 30% from fossil fuels. The latter entails the construction of 63 GW of CCS-equipped power stations by 2030 and an additional 128 GW between 2030 and 2050.

Since the *Power Choices* report by Eurelectric (2010) provides the amount of CCS only at aggregate European level, we need to make an assumption on how this breaks down to individual countries. First, we assume that CO_2 capture deployment is limited to the 18 countries in which CCS takes places in the EU's *Baseline 2009* scenario (Capros et al., 2010). Second, we assume that the aggregate European level of CCS (as obtained from the *Power Choices* scenarios) is distributed between countries in proportion to current CO_2 emissions from the power sector, as obtained from E-PRTR (2010). Third, within each country, the amount of CCS is distributed between various industrial 'clusters'. Further details about the clustering approach can be found in Morbee et al. (2012). Size and location of potential CO_2 storage sites (depleted hydrocarbon fields and saline aquifers) is obtained from the EU GeoCapacity project (Vangkilde-Pedersen et al., 2009). Due to technical uncertainty and public acceptance issues, onshore saline aquifers are excluded as potential CO_2 storage sites. Details about the assumptions can be found in Morbee et al. (2012). Table 5.1 provides an overview of the assumed annual amounts of CO_2

[2]Ideally, the amount of CCS-based electricity production per country should be determined simultaneously with the optimisation of the CO_2 network, as in the least cost theory of industrial location (see e.g. Weber, 1909). In such a model the location of CCS-based power plants would depend on the spatial distribution of electricity demand, coal transport costs, and the CO_2 pipeline network construction costs. This, however, would lead to severe computational challenges. As discussed later in Section 5.4 the problem is already very challenging from a computational complexity point of view, even when an exogenous scenario for CCS deployment is taken.

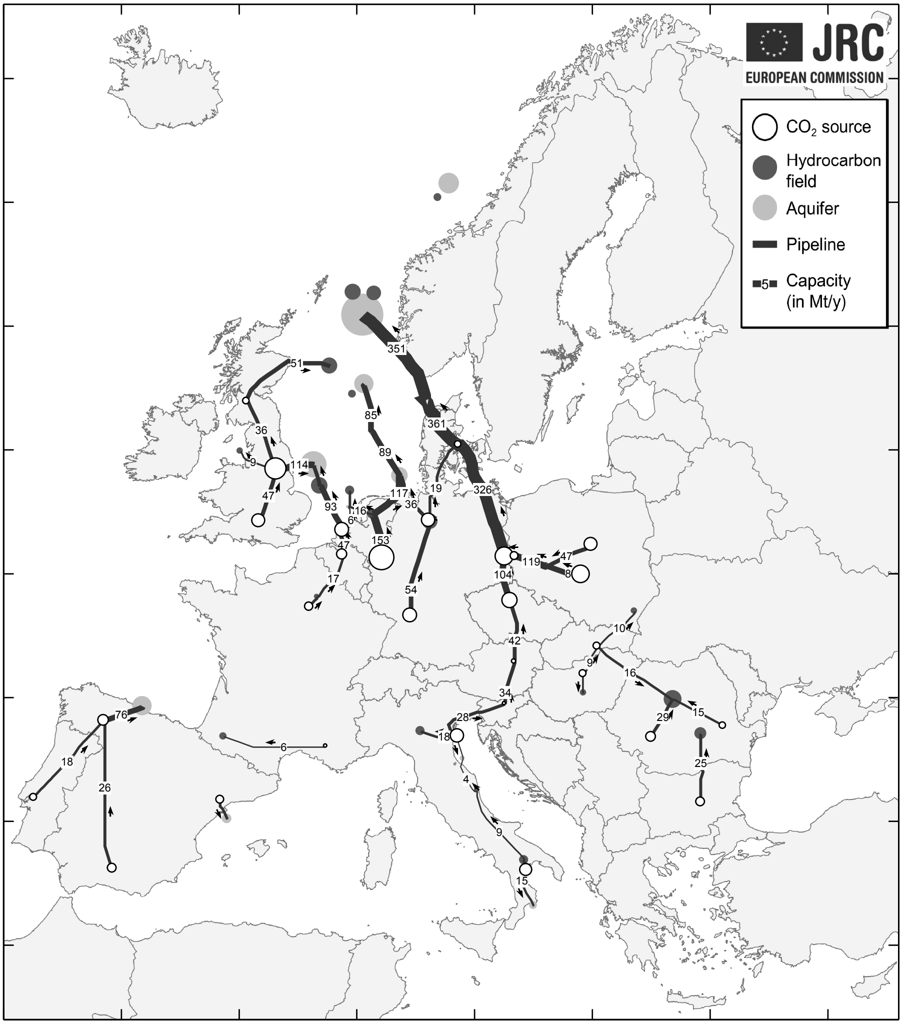
captured in each of the countries, as well as the annual CO_2 storage capacity. CCS activities in Finland have been left out of this picture: since they are geographically far away from the CO_2 network in the rest of Europe, they do not contribute to the negotiation game described in the remainder of the paper. Leaving out Finland from the start reduces computational complexity of the Shapley value approach by an order of magnitude.

Table 5.1: CO_2 capture rates and storage potential assumed in this study

Country	**Annual CO_2 captured** Mt/y (2050)	**Annual CO_2 storage potential** Mt/y (2050)
Austria	8	
Belgium	30	
Bulgaria	25	
Czech Republic	62	
Denmark	15	14
France	27	17
Germany	337	111
Hungary	17	9
Italy	89	61
Netherlands	52	40
Norway		636
Poland	147	17
Portugal	17	
Romania	43	88
Slovakia	16	
Slovenia	6	
Spain	80	108
United Kingdom	173	315
Total	1145	1416

We use the *InfraCCS* model to compute the optimal CO_2 pipeline network in 2050 for the given configuration of sources and sinks. *InfraCCS* is a cost-minimising mixed-integer linear programming model, which takes into account the scale effects of pipelines (see Morbee et al., 2012 for more details). The resulting optimal network is shown in Figure 5.1. The network consists of 11001 km of pipelines, which transport 1145 Mt/y of captured CO_2 from sources to sinks. The total investment required is 28.0 billion euro. The network in Figure 5.1 assumes joint international optimisation. If, by contrast, countries develop networks individually, the resulting pipeline construction would be as

Figure 5.1: CO_2 pipeline network in 2050, assuming joint international optimisation. Total amount of CO_2 captured and stored: 1145 Mt/y.



in Figure 5.2. Since not all countries have sufficient storage potential, not all captured CO_2 projected in the *Power Choices* scenario can be stored. In total, in the network in Figure 5.2, only 565 Mt/y of CO_2 is transported and stored, i.e. less than half of the amount stored under joint international optimisation (Figure 5.1). Non-stored CO_2 can be recognised in the figure as red circles that are not connected with any pipeline. The network in Figure 5.2 is 5097 km in length and costs 6.4 billion euro.

Thus, the benefits of international cooperation are that an additional 580 Mt/y of CO_2 can be captured and stored compared to individual country action, albeit at the cost of a more expensive network. In order to translate the benefits of international cooperation into a single total quantity, we need to make an assumption about the cost of the outside option for the 580 Mt/y of CO_2 that cannot be stored in the non-cooperative case. Clearly, this cost should be lower than the assumed CO_2 emissions allowance price in the EU Emissions Trading System (EU ETS), since the fact that the CO_2 cannot be stored also saves the cost of capturing it in the first place. The description of the PRIMES model in Eurelectric (2010) states that the assumed CO_2 transport and storage cost ranges from 6 to 25 euro per tonne of CO_2. Assuming that (i) the lower bound of the range refers to a situation with only storage costs and no transport costs (i.e. storage very close to the capture site) and (ii) CO_2 storage costs are constant and geographically uniform, we infer that the transport cost in the *Power Choices* scenario ranges from 0 to 19 euro per tonne of CO_2. Hence, if transport costs exceed 19 euros per tonne of CO_2, the PRIMES model will switch technologies and the required emissions reduction will be realised through other means (e.g. wind energy). We therefore assume in our analysis the availability of an 'outside option' that costs 19 euros per tonne of CO_2.[3]

For the sake of simplicity we apply this value uniformly across all countries. Assuming a 7.5% discount rate[4] and a 10-year horizon, the cost of not being able to capture and store 580 Mt/y is 75.7 billion euro. Combined with the investment of 6.4 billion euro, the total cost of the non-cooperative case is therefore 82.1 billion euro, compared with 28.0 billion euro in the cooperative case. In this setting, the benefits of international cooperation are therefore 54.1 billion euro.

The question addressed in this paper is how these benefits can be allocated between participating countries in order to ensure cooperation. Equivalently, the question is how to allocate the cost burden of the 28.0 billion euro invest-

[3]Note that this value is *lower* than the full CO_2 emissions allowance price, which is projected to be 103.2 euro per tonne by 2050 in the scenario under consideration (Eurelectric, 2010). Indeed, if a country decides to pursue the outside option (hence no CCS) it does not need to pay for any other CCS-related costs, such as the CO_2 capture costs, or the residual allowances for non-captured CO_2. These cost savings, together with the 19 euros, pay for the outside option of 103.2 euro per tonne.

[4]This value is midway between a typical social discount rate (5%) and a typical industrial discount rate for this type of investments (10%).

Figure 5.2: CO_2 pipeline network in 2050, assuming individual optimisation per country, without international cooperation. Total amount of CO_2 captured and stored: 565 Mt/y.

ment.

5.3 Bargaining power in multilateral cooperation: the Shapley value

The allocation of benefits between participating countries depends on each country's bargaining power. In the context of this analysis, bargaining power is mainly associated with two types of rents:

Storage rent. Countries with excess CO_2 storage (i.e. more storage capacity than what is required to store the amounts of CO_2 captured within the country) can offer this capacity to other countries who are short of CO_2 storage capacity. Since the availability of additional storage capacity reduces the need for recurrence to the outside option (i.e. switching to an alternative technology at a cost of 19 euro per tonne), it brings about a cost reduction for the coalition partners of a country with excess storage. This increases the bargaining power of countries with excess storage, and allows them to obtain a 'storage rent'.

Transit rent. Some countries have a strategic location, which allows for shortcuts between CO_2 sources and storage sites. For example, in Figure 5.1, the participation of Denmark allows for a near-straight pipeline from Poland to Norway. Non-participation of Denmark would require a detour of the pipeline and hence a higher construction cost. This translates into bargaining power for transit countries, allowing them to obtain a 'transit rent'. This is in fact the reverse of the well-known *Jepma-effect* in international transport networks.The Jepma-effect, first described by Jepma (2001) in the context of liberalisation of the Dutch natural gas transport network, is the observation that gas transport tariff differences between neighbouring countries may incentivise gas shippers to reroute gas flows in order to take advantage of a cheaper neighbouring network, even if the new route is inefficient from a technical perspective. The CO_2 pipeline transit rent described in this paper is essentially the same effect but in the opposite direction: a country with an advantageous transit location may be incentivised to increase CO_2 transport tariffs because it would be even more costly for foreign CO_2 shippers to reroute CO_2 flows in order to circumvent the country.

To assess these rents in an integrated way, we apply the *Shapley value* approach, introduced by Shapley (1953). The Shapley value defines a 'fair' allocation of the benefits of cooperation, taking into account the contributions of each of the players (in this case in particular the storage and transit rents described above). It defines the only allocation that satisfies a set of desirable properties (individual fairness, efficiency, symmetry, additivity and zero-player property). Starting from a set N of n players (in this case: countries), we define the function $v : \mathcal{P}(N) \to \mathbb{R}$, such that, for every subset S of N, $v(S)$ is the pay-off of a cooperation among the countries in S. According to the Shapley value, the amount of benefit received by player $i \in N$ is:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} (v(S \cup \{i\}) - v(S)) \tag{5.1}$$

The sum is computed over all possible coalitions of players. For each coalition, equation (5.1) computes the difference $(v(S \cup \{i\}) - v(S))$ between the pay-off of the coalition with and without player i. The Shapley value ϕ_i is then a weighted average of those values. Intuitively, the formula computes the contribution added by player i to the 'grand coalition' (the coalition of all players) averaged over all possible sequences in which this grand coalition can be formed. As an example, we compute $(v(S \cup \{i\}) - v(S))$ for the subset S that includes all countries except Denmark. i is Denmark. For this case we have $S \cup \{i\} = N$, hence the pay-off $v(S \cup \{i\})$ is the cost of the fully cooperative CO_2 network from Figure 5.1, i.e. 28.0 billion euro. The pay-off $v(S)$ can be determined by running the *InfraCCS* tool without Denmark. This is shown in Figure 5.3. The cost of this network is 32.1 billion euro. Hence, by including Denmark in the coalition, there is a cost saving of 4.1 billion euro, because the participation Denmark permits more efficient routing of pipelines from central Europe to Norway. Furthermore, the inclusion of Denmark in the coalition offers cost savings in Denmark as well because by participating, Denmark does not have to build its own small network, which would cost 0.6 billion euro. In addition, the inclusion of Denmark offers a solution for the 1.6 Mt/y that Denmark would not be able to store domestically (which would cost 0.2 billion euro in order to pay for the NPV of the outside option of 19 euros per tonne). In total, the contribution $(v(S \cup \{i\}) - v(S))$ of Denmark to the coalition S is therefore 4.9 billion euro. This computation is done for all possible subsets S of N and all players i. Equation (5.1) requires a total of 262143 runs of the above-mentioned *InfraCCS* model.

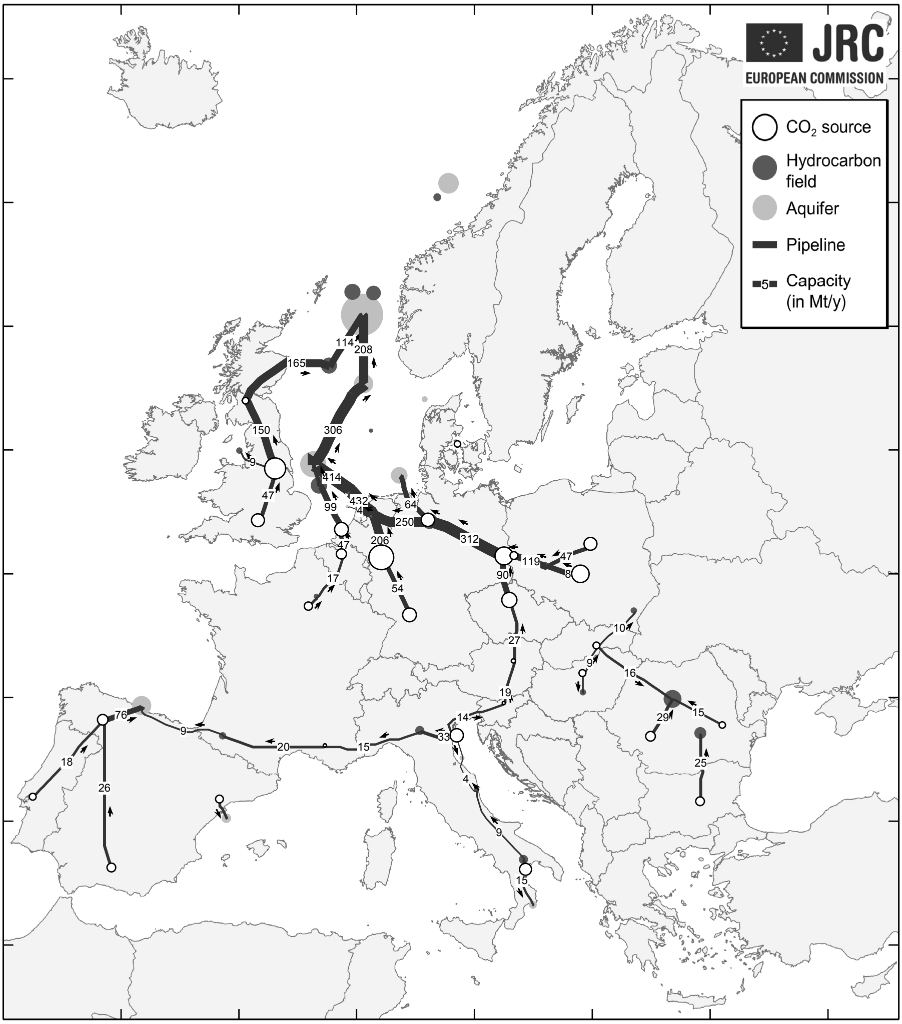
5.4 Simulations

In this section, we apply the Shapley value to the issue of European multilateral CO_2 pipeline infrastructure negotiations. The realisation of the network of Figure 5.1 requires cooperation among $n = 18$ countries: 17 source countries within the EU[5], plus Norway. When applying equation (5.1), we distinguish two cases:

Case 1: National CO_2 pipeline monopolies. In this case, we assume that every country has a monopoly on CO_2 pipeline construction within its

[5] As mentioned before, Finland has been omitted from the set of 18 countries active in CCS according to the simulations by Capros et al. (2010)

Figure 5.3: CO_2 pipeline network in 2050, assuming joint international optimisation without participation of Denmark



territory. As a result, a pipeline through a given country cannot be built by a coalition that does not include this country.

Case 2: Liberalised CO_2 pipeline construction. In this case, any country is free to build pipelines in the entire EU and Norway. This does not mean however, that the land on which these pipelines are constructed is free: a cost to cover the right-of-way is included in the pipeline costing approach embedded in the *InfraCCS* model.

Note that in both cases, CO_2 cannot be stored in a given country by a coalition that does not include that country.

We compute the Shapley value for both cases, which, as mentioned above, requires the computation of the pay-offs of 262143 coalitions in each case. This is a computational challenge, because each pipeline optimisation problem is a mixed-integer problem (MIP), which is NP-hard to solve, i.e. no efficient algorithms for such problems exist today. In Case 1, due to national pipeline monopolies, many coalitions can be broken down into independent contiguous subsets. As a result, Case 1 requires the computation of the pay-offs of only 26922 contiguous coalitions, which can then be combined to obtain the results for all 262143 coalitions. Hence, in Case 1 the number of runs of the *InfraCCS* model can be reduced by a factor 10. In Case 2 however, all coalitions need to be run with the *InfraCCS* model.[6]

Table 5.2 shows the resulting Shapley values in both cases. The Shapley value is shown both in absolute terms and as a percentage of the total. These Shapley values show how the 54.1 billion euro benefit from international cooperation can be allocated between countries. In Case 1, large rents are allocated to Norway and the UK, which are the main net storage providers in this analysis (see Table 5.1). In total, the net storage providers capture 38% of the benefits. A large rent is also allocated to Denmark, which plays a crucial role as transit country in Figure 5.1. The large Shapley value for Germany is related to the fact that it contributes the most CO_2, which allows for avoiding a large 'outside option' cost.

In Case 2, the rents shift more towards the largest storage provider, i.e. Norway. In total, net storage providers capture 45% of the benefits in this case. Due to the liberalisation of pipeline construction, Denmark loses most of its Shapley value, which demonstrates indeed that its bargaining power in Case 1 can be attributed to its transit position. Likewise, one notes that Germany's bargaining power decreases between Case 1 and Case 2, which points to the fact that a portion of its bargaining power in Case 1 is due to its central location in Europe, which allows it to serve as a hub for CO_2 transport. Finally, note that Poland gains significantly in the transition from Case 1 to Case 2: indeed, since Poland is at the end of the pipeline network, it does

[6]For that reason, the MIP computation time is deliberately capped at 3 minutes per coalition in Case 1, and 3 seconds in Case 2. In total, the computation takes about a week on a standard computer.

Table 5.2: Shapley value allocation for the coalitional game described in Section 5.2

Country	**Shapley value**				
	Case 1		**Case 2**		**Difference Case 2 – Case 1**
	Bn EUR	Percent	Bn EUR	Percent	Bn EUR
Austria	0.4	1	0.4	1	0.0
Belgium	1.8	3	1.7	3	-0.1
Bulgaria	1.0	2	1.1	2	0.1
Czech Republic	1.8	3	3.3	6	1.5
Denmark	4.5	8	0.3	1	-4.2
France	1.4	3	0.6	1	-0.8
Germany	14.1	26	11.9	22	-2.2
Hungary	1.0	2	0.5	1	-0.6
Italy	0.9	2	1.0	2	0.1
Netherlands	1.4	3	0.7	1	-0.7
Norway	11.7	22	16.6	31	4.9
Poland	3.5	7	6.5	12	3.0
Portugal	0.6	1	0.7	1	0.1
Romania	2.2	4	2.4	4	0.2
Slovakia	0.6	1	0.8	2	0.2
Slovenia	0.5	1	0.2	0	-0.2
Spain	1.2	2	1.2	2	0.0
United Kingdom	5.4	10	4.0	7	-1.4
Total	54.1	100	54.1	100	

not have an advantageous transit position and therefore stands to gain from pipeline liberalisation.

Table 5.3 translates the Shapley values from Table 5.2 into the allocation of the cost burden of the network. As mentioned in Section 5.2, the total required investment in the CO_2 pipeline network is 28.0 billion euro. Table 5.3 shows how this cost of 28.0 billion euro is shared between individual countries. Note that some countries make a net payment, while others are net recipients

Table 5.3: Investment cost burden sharing for the CO_2 pipeline network shown in Figure 5.1

Country	**Contribution to investment**				
	Case 1		**Case 2**		**Difference Case 2 – Case 1**
	Bn EUR	Percent	Bn EUR	Percent	Bn EUR
Austria	0.6	2	0.6	2	0.0
Belgium	2.1	8	2.2	8	0.1
Bulgaria	2.2	8	2.1	7	-0.1
Czech Republic	6.3	22	4.8	17	-1.5
Denmark	-3.7	-13	0.5	2	4.2
France	0.5	2	1.3	5	0.8
Germany	16.3	58	18.5	66	2.2
Hungary	0.2	1	0.8	3	0.6
Italy	3.3	12	3.2	12	-0.1
Netherlands	0.5	2	1.2	4	0.7
Norway	-11.7	-42	-16.6	-59	-4.9
Poland	13.8	49	10.8	38	-3.0
Portugal	1.6	6	1.5	6	-0.1
Romania	-1.9	-7	-2.1	-8	-0.2
Slovakia	1.5	5	1.3	5	-0.2
Slovenia	0.3	1	0.5	2	0.2
Spain	-0.1	0	-0.1	0	0.0
United Kingdom	-3.8	-14	-2.4	-9	1.4
Total	28.0	100	28.0	100	

from the cooperation. A large share of the cost is borne by countries with large volumes of excess CO_2, such as Germany and Poland. Due to reasons mentioned before, Germany contributes more in Case 2 than in Case 1, while Poland contributes less. Note that Denmark is a net recipient in Case 1, while it is a net contributor in Case 2.

Finally it is possible to translate the Shapley values into prices expressed per tonne of CO_2 stored. For this purpose, we first compute – for each country – the additional investment made when going from the non-cooperative case (Figure 5.2) to the cooperative case (Figure 5.1), i.e. the value from Table 5.3 minus the domestic pipeline investments made to realise the network of Figure 5.2. Secondly, we divide this number by the additional amount of CO_2 captured in this country in the cooperative case compared to the non-cooperative case.[7] Obviously, this computation is meaningful only for net CO_2 exporters. The results of the countries with the highest cost per tonne of CO_2 exported are shown in Table 5.4. One immediately observes that the

Table 5.4: Costs per tonne of CO_2 exported (in EUR per tonne of CO_2)

Case 1		**Case 2**	
Country	**Cost**	**Country**	**Cost**
Poland	15.0	Italy	13.9
Czech Republic	14.7	Portugal	13.1
Italy	14.3	Slovenia	12.8
Portugal	13.7	Bulgaria	12.3
Slovakia	13.5	Poland	11.7
Bulgaria	13.1	Austria	11.6
Austria	11.3	Hungary	11.6
Belgium	10.2	Slovakia	11.6
Germany	9.9	Germany	11.3
Slovenia	7.4	Czech Republic	11.2
Hungary	2.3	France	11.0
Netherlands	1.1	Belgium	10.7
France	-0.3	Netherlands	9.4

spread of costs is much smaller in Case 2 than in Case 1. Indeed, in Case 1 there is much more heterogeneity between countries, depending on their transit position. In Case 2, differentiation between counties is mostly due to their distance from the main storage sites. The range of costs is reduced from over 15 euro per tonne in Case 1, to less than 5 euro per tonne in Case 2. Note that e.g. Slovenia, although located far away from the North Sea, pays a rather low price in Case 1, due to its role as a transit country for Italy. In Case 2 however, this advantage disappears and it ranks as one of the higher-cost countries.

A similar analysis can be done for countries that are net importers of CO_2. As above, we divide the difference in cashflow (when going from the non-

[7] As before, discounting at 7.5% is performed and a 10-year time horizon is assumed.

cooperative case to the cooperative case) by the amount of CO_2 imported, with discounting as above. The results are in Table 5.5. We observe that in

Table 5.5: Revenue per tonne of CO_2 imported and stored (in EUR per tonne of CO_2)

Case 1		**Case 2**	
Country	**Revenue**	**Country**	**Revenue**
Spain	10.0	Spain	10.1
United Kingdom	8.5	Romania	8.6
Romania	7.9	United Kingdom	6.3
Norway	3.9	Norway	5.6
Weighted average	5.1	Weighted average	6.0

Case 1, the revenue per tonne of CO_2 imported and stored ranges from 3.9 to 10.0 euro per tonne of CO_2, with a weighted average of 5.1 euro per tonne. In Case 2, with liberalised pipeline construction, the average revenue increases to 6.0 euro per tonne, although the impact differs per country. While countries with strategically located onshore storage do gain revenues when going from Case 1 to Case 2 (e.g. Spain, Romania), we observe that the revenues of the United Kingdom decrease, due to its remoteness.

As a side-effect of the results of Table 5.5, we can compute an estimate of the resource rent associated with a depleted hydrocarbon field that is to be used for CO_2 storage. As a very approximative rule of thumb, a depleted oil field can store roughly 1 tonne of CO_2 per tonne original recoverable oil reserves. The resource rent of 5-6 euro per tonne of CO_2 stored therefore corresponds to approximately \$1 per barrel of original recoverable oil reserves. This is clearly far below the resource rent that was originally obtained from the oil extraction. For gas fields, the results are more favourable. As a very approximative rule of thumb, a depleted gas field can store roughly 2 tonnes of CO_2 per thousand cubic meters (tcm) of gas in its original recoverable reserves. The resource rent of 5-6 euro per tonne of CO_2 stored therefore corresponds to approximately 1 euro per MWh of gas. This is roughly 5% of the wholesale price of natural gas: e.g. the average German import border price was 20 euro per MWh in 2010 according to BAFA (2011). Overall, therefore, it seems that the rent is relatively small from the perspective of petroleum economics. However, it is relatively large from the perspective of CCS economics. Indeed, typical storage costs are estimated to be 1 to 20 euro per tonne of CO_2 stored depending on such factors as the type of storage site (hydrocarbon field or aquifer), the location (onshore/offshore) and the presence of re-usable legacy wells (see e.g. ZEP, 2011). The rent of 5-6 euros per tonne of CO_2 needs to be added to this number, and represents an increase of 25 to 600% of the costs.

5.5 Conclusions

In this paper, we have analysed bargaining power in the multilateral negotiation process that would be required to develop a cost-minimising trans-European CO_2 transport infrastructure if CO_2 capture and storage (CCS) is deployed on a large scale by 2050. We apply the Shapley value to the coalitional game between 18 European countries, in two different cases: one case with national pipeline monopolies and one case with liberalised pipeline construction. Using the *InfraCCS* pipeline optimisation model, we perform a numerical simulation, which computes each country's contribution to a 28.0 billion euro trans-European CO_2 pipeline network. We find that countries with more storage capacity than capture activity obtain 38% to 45% of the benefits, with the higher number corresponding to the case with liberalised pipeline construction. Countries with a strategic transit location capture significant rent in the case of national pipeline monopolies, e.g. Denmark obtains a net benefit of over 4 billion euro in case of national pipeline monopolies, but loses almost all of this if pipeline construction is liberalised. Finally, the liberalisation of pipeline construction reduces by two-thirds the differences between countries in terms of cost per tonne of CO_2 exported. As a side result of the analysis, we find that the resource rent of a depleted hydrocarbon field (when used for CO_2 storage) is roughly \$1 per barrel of original recoverable oil reserves, or 1 euro per MWh of original recoverable gas reserves. This is small from the perspective of petroleum economics, but corresponds to 5-6 euro per tonne of CO_2 stored, which may increase CO_2 storage costs by 25 to 600%.

The analysis is strongly dependent on the assumptions underlying the Shapley value. Other approaches exist, and the allocation shown in this paper is not necessarily the only possible allocation. Even more importantly, the cooperative game theory framework from which the Shapley value arises, assumes that the grand coalition is eventually formed. This is in stark contract with current developments in Europe: unlike the US, there is no CO_2 pipeline network in Europe yet, and many countries would oppose such developments, which would impede the construction of any trans-European pipeline network.

As a further caveat, it should be mentioned that the rent computed here is only the rent arising from market power in CO_2 transport. In addition there may be a Hotelling (1931) rent for storage sites if storage becomes scarce. Furthermore, supranational regulation and enforcement may be required in order to avoid renegotiation once the network is in place. More generally, there is a question about which market organisation would be suited for the operation of such a jointly optimal network. Finally, an important area for future work is a more thorough understanding of the 'outside option' through better integration with the economic equilibrium models that generate the scenarios of CO_2 capture rates.

5.6 Acknowledgements

This work has been carried out within the multi-annual work programme of the "Assessment of Energy Technologies and Systems" (ASSETS) Action of the European Commission's Joint Research Centre. Any interpretations or opinions contained in this paper are those of the author and do not necessarily represent the view of the European Commission.

Part II

Hedging

Chapter 6

Market Completeness: How Options Affect Hedging and Investments in the Electricity Sector

This chapter has been published in *Energy Economics*.

6.1 Introduction

The specific characteristics of electrical energy create a need for hedging. Electricity cannot be stored economically, and therefore the price for electricity is determined by the supply and demand conditions at each given hour. As demand for electrical energy is very inelastic and of a stochastic nature and as generators face production capacity constraints, spot prices are very volatile. Liberalized electricity markets are therefore typically organized around regional spot markets for energy, which determine hourly spot prices, complemented with markets for long-term contracts, which help coordinate the actions of the players and allow for hedging of volume and price risks. The extent to which a firm can hedge its exposure, depends on the availability of markets, their liquidity (determined by such parameters as trading volume and bid-ask spread), and the presence of speculators who can absorb part of the risk. These factors change as markets evolve from pure OTC to sophisticated spot and futures markets, and to more complete markets in which there is a liquid trade of a broad set of derivatives.[1]

[1]Note that vertical integration of electricity production and retail is an alternative way of creating a 'complete' set of hedging instruments between production and retail.

Recognizing that electricity markets are typically very incomplete, the objective of this paper is to analyze the effect of increasing *market completeness* on welfare and on investment incentives in the electricity sector. In our paper, market completeness is measured as the number of electricity options available to producers and retailers, in addition to a forward contract. Indeed, as more options with different strike prices become available, firms have more instruments to trade risks and markets become more complete.[2]

This paper develops an equilibrium model of the electricity market, which includes the production process, spot market trades and trade of derivatives. For illustrative purposes, the model is calibrated on the German electricity market, although an exact analysis of the German market is not the objective of this paper. First, the results show that adding option markets is welfare-enhancing, but that most of the benefits are obtained with one to three options. In particular, if firms have strong aversion of negative shocks (shocks that would cause firm bankruptcy), then no equilibrium can be found unless option contracts are available in order to protect retailers against bankruptcy under all conditions. Second, we analyze how investment decisions by small firms are affected when an increasing number of derivatives are traded. We show that market incompleteness typically leads to underinvestment. The effects are, however, different for base load plants and peak load plants: the presence of forward contracts only (i.e., no options) is sufficient for investment in base load plants to reach the same level as in case of market completeness, but there will be underinvestment in peak load plants until there is a sufficient number of option contracts (which allow the investor to hedge market risk associated with the investment). Increasing the number of derivatives may, however, also lead to 'crowding-out' of certain investments in power plants: if the investor can trade a financial contract that is highly correlated with the profit of a potential investment and the financial contract provides a more attractive risk-return ratio, then the investor will only invest in the financial market, as its risk-taking capabilities are limited. The amount of information contained in the equilibrium market prices, increases with the number of financial products being traded: it is shown that the quality of power plant investment decisions that are based on risk-free probabilities inferred from market prices, improves with the number of contracts being traded. If markets are not sufficiently complete, players basing their investment decisions on risk-free probabilities may significantly overinvest.

The model proposed in this paper is complementary to the traditional financial models for derivatives pricing, which are based on the no-arbitrage approach. In fact, it has been observed that it is difficult to apply the traditional no-arbitrage approach to the case of electricity derivatives, because the

[2]The paper assumes that demand shocks are the only source of risk. In such a setting, the market is complete if options at every strike price can be traded. However, if there are also firm-specific shocks, then additional derivatives should be added for the market to be complete.

non-storability of electricity means that the well-known cost-of-carry relationship and delta-hedging strategy cannot be implemented, and hence pricing of electricity forwards and options cannot be done in the usual manner.[3] For that reason, Bessembinder and Lemmon (2002) adopt an equilibrium approach and explicitly model the economic determinants of market clearing forward prices. Bessembinder and Lemmon's (2002) model was only focused on forward contracts, and in our paper we extend their model to include an increasing number of options in addition to a forward contract. We then use the model to study the effects of increasing market completeness on welfare and on investment incentives.

The paper is organized as follows. First, Section 6.2 provides an overview of relevant research on incomplete markets, including the applicability to electricity markets. Next, Section 6.3 describes the electricity market model that is used to obtain the results of this paper, while Section 6.4 describes the model data. Section 6.5 verifies the welfare effects of an increasing number of markets. Sections 6.6 and 6.7 analyze the effect on investment incentives, based on welfare considerations (Section 6.6) and on risk-free probabilities, i.e. the 'finance approach' (Section 6.7). Finally, Section 6.8 summarizes our conclusions.

6.2 Literature review

The topic of this paper is closely related to the literature on incomplete markets and financial innovation, as well as to the literature on hedging in electricity markets. In this section we first introduce the concept of incomplete markets. Next, we discuss the main results of the literature. Finally, we highlight the relevance for electricity markets and discuss related work on hedging in electricity markets. We base our discussion on market completeness mainly on Staum (2008) and Duffie and Rahi (1995).

6.2.1 Incomplete markets

Markets are incomplete when *perfect risk transfer* between the agents is *impossible*. There might be several reasons why this would be the case. First, the marketed set of assets may be insufficient to hedge the class of risk one wishes to hedge. This type of incompleteness deals with the spanning role of securities (see also Allen and Gale, 1994). Second, markets might be imperfect due to the existence of transaction costs and/or trading constraints. For instance, firms might not be able to take a short position in a traded security. These costs and/or constraints make it effectively impossible to transfer risk

[3] Eydeland and Geman (1998) present a pricing model for power options that relies on assumptions regarding the evolution of forward power prices. They show that the approach is adequate to manage monthly and yearly power options, but that it does not offer a safe solution for daily options.

perfectly. In our paper we focus on the first type of market incompleteness: the missing markets problem.

In practice, markets are never complete, as not all risk factors are traded on a market. Hence, when might market incompleteness be relevant for hedging or pricing decisions? We mention two situations in which this might be the case. The first situation is when some of the variables one would like to hedge are derived from non-market prices, as is the case for weather derivatives. The second typical situation of market incompleteness occurs when the price of an asset does not follow a standard random walk process – where prices changes are 'infinitesimally small' – but contains 'large' price jumps. The problem with price jumps is that a hedging strategy which dynamically adjusts a portfolio containing the underlying asset and a risk-free bond, is no longer possible, as the payout is non-linear in the size of the shock. In order to complete the market one would need to add a forward market and a set of option markets with different strike prices.

6.2.2 Research results on incompleteness

The first main result of the literature on welfare effects and pricing of additional assets is that *welfare* in an incomplete market is lower than in a complete market because not all risk is perfectly allocated in the market.[4] This is a rather intuitive result: as in an incomplete market not all potential gains from trade are exhausted, total welfare can be improved by a sufficient number of additional markets until the market is complete. This simple intuition does, however, not carry over to situations where only one additional market is added to the economy, without completing the market. Hart (1975) shows that adding a financial product might make every one in the economy worse off. Extending this result, Elul (1995) and Cass and Citanna (1998) show that in an economy with many consumption goods one can always find an asset that makes everyone worse off, or an asset that makes everyone better off, or an asset that makes any combination of individuals better or worse off.[5] Note that introducing all financial assets (completing the market) does not necessarily make everyone better off. Complete markets are Pareto efficient, but not necessarily Pareto dominant with all possible incomplete market allocations. Willen (2005) studies the impact of market innovation in more detail and shows that, when agents have exponential utility and risk is normally distributed, the

[4]In this paper we assume that a Walrasian equilibrium exists, even when markets are incomplete. In a general equilibrium setting with multiple goods, (where securities can contain different bundles of goods), this is not guaranteed. However, when we restrict ourselves to economies where financial claims only have a pay-off in terms of a single numeraire good, existence is guaranteed. On existence of equilibria in a general equilibrium setting, see Duffie and Shafer (1985) and Duffie and Shafer (1986).

[5]Similar results were obtained earlier by Milne and Shefrin (1987) in a specific model set-up. Note that the results are not applicable here because our model assumes only one relevant good: money.

effect of a financial innovation can be split up in a portfolio effect and a price effect. Elul (1999) studies the welfare effects of a financial innovation in a single-good market.

Boyle and Wang (2001) study the *pricing* of a new derivative in an incomplete market. They show that one should not use the standard arbitrage assumptions typically used in the financial (engineering) literature, as the prices of existing assets may change once a new asset is added to the economy.[6] Instead, they recommend to make explicit assumptions on the preferences of the agents in the economy and to use an equilibrium model to derive the prices of the different assets. Staum (2008) and Carr et al. (2001) argue however that results of equilibrium models depend very much on the choice of the utility function, the initial endowment of the firms, and the parameters of the probability measure, and are therefore not useful for trading decisions.

6.2.3 Incompleteness and hedging in electricity markets

The electricity market is an interesting example of a very incomplete market. Since electricity cannot be stored economically and electricity prices are very volatile, it is difficult to hedge even the most basic forward contracts and options, when they are not traded directly in the market.[7] Our paper focuses on the electricity market and builds further upon existing studies on contracting and hedging in this market.[8]

First of all, Bessembinder and Lemmon (2002) develop a partial equilibrium model of the spot market and *one* forward market. They derive analytical solutions for forward and spot prices in a setting in which firms are risk averse, production cost are convex, retail prices are fixed and demand is stochastic. Their theoretical predictions on risk premia are verified empirically: the model correctly predicts when markets should be in backwardation or in contango. Siddiqui (2003) completes the Bessembinder and Lemmon (2002) model by introducing a forward market for ancillary services (reserve capacity) and deriving analytical results that link the forward prices of electricity and ancillary services with the statistical properties of the spot price. Our paper also extends the framework of Bessembinder and Lemmon (2002) and allows for multiple financial products to be traded – not just one forward contract.[9] Furthermore,

[6] They also show that the condition of arbitrage-free pricing does not determine a unique price for the newly created asset.

[7] If it were possible to store electricity economically, then the forward contract could be hedged through a combination of stored electricity and a loan. If, in addition, electricity prices did not have spikes, then electricity options could be hedged through a dynamically adjusted portfolio of stored electricity and a loan (i.e., delta-hedging).

[8] In the review we limit ourselves – with the exception of Oum et al. (2007) – to studies that rely on equilibrium models of the electricity market. The alternative to equilibrium models is the study of *one* firm's contracting and production decisions for an exogenously given stochastic spot price process and forward price.

[9] With a forward contract this paper refers to a contract for future delivery of a fixed quantity of a good at a fixed price. We will not explicitly specify whether these contracts

our paper analyzes the effects of speculators trading in a number of derivatives markets and studies an alternative, more realistic formulation of risk aversion. In addition, the effect on investments is analyzed.

The usefulness of financial instruments other than forwards to hedge risks in electricity markets is discussed by Oum et al. (2007). They show that a regulated retail firm can use a combination of forwards, call options and put options to hedge its volumetric risk, and draw attention to the regulated firm's difficulty to hedge when regulators forbid trade in derivatives that look speculative, such as weather derivatives, and rebuff contracting positions that require the firm to pay a sum ex-ante. The optimal hedging strategy is found by optimizing the firm's utility, subjective to the financing constraint. The results are derived for the CARA and the mean-variance utility functions, with an endogenously given price and quantity distribution function. In our paper we develop an equilibrium model of the market and show that option contracts are important instrument *to transfer volumetric risks* from generators to retailers, even more so when firms might face liquidity constraints.[10] We also show the importance of options for investment decisions.

Baldursson and von der Fehr (2007) study vertical integration, forward contracting and hedging in an equilibrium electricity market model. They show that vertical integration might increase the equilibrium risk premia in the market and lower overall welfare, compared with forward contracting. The reason why this happens in their model is that they assume that a vertically integrated firm has a 'smaller capacity' to take up risk than two separate entities combined. Even though our model does not represent vertical integration explicitly, the case of vertical integration in our model corresponds to the case in which perfect risk-transfer between producer and retailer is possible, i.e., the case of market completeness. In such a setting, the implicit assumption is that the vertically integrated firm has the 'same capacity' to take up risk as the two separate firms combined. In our opinion, this is a more realistic assumption. We see the difference between vertical integration and contracting by means of a forward contract, as follows: within the vertically integrated firm, risk sharing between generation and retail is perfect, while risk sharing by trading forward contracts is imperfect, leaving part of the risk untraded.[11]

are traded over the counter (OTC), or whether they are traded as 'futures' on a centralized power exchange.

[10]The notion *liquidity constraints* in this context refers to the constraints faced by an individual firm in the financing of its activities. We model these through a CRRA utility function, as described further in the paper. Separately, there is the entirely different issue of limited market liquidity, which refers to the fact that in general, markets for energy derivatives are relatively thin. As mentioned in the introduction, market liquidity is one of the elements of market completeness. In our paper, limited market liquidity is therefore modeled by assuming that not all possible contracts are traded. An alternative way of modeling limited market liquidity is by assuming a large bid-ask spread. Our paper, however, does not consider transaction costs, hence we do not study market liquidity in this way.

[11]An additional difference between vertical integration and trading derivatives is that in a derivatives market, financial investors can reduce the risk premia in the market.

Also Aid et al. (2006) study vertical integration, forward contracting, hedging, and retail competition. They develop an equilibrium model in which firms have a mean-variance utility function and show that both vertical integration and forward contracting allows for a better risk sharing between retailers and generators and leads to lower retail prices, increased market share for small generators, and a reduction of the profits of retailers. Compared with long-term contracts, vertical integration leads to perfect risk sharing between generators and retailers. Additionally, forward markets might not develop under some parameters of the game in which case no risk is shared between upstream and downstream firms. The results of Aid et al. (2006) on the comparison of vertical integration and forward contracting are driven by the change of the utility function (and the implied capacity of firms to take up risks) and the quality of risk transfer between upstream and downstream firms (market completeness). In our paper we single out the effect of market completeness. We do not, however, study retail competition. Our paper assumes a perfectly competitive market and neglects strategic issues associated with long-term contracting that have been reported in the literature.

Allaz and Vila (1993) study the role of forward contracts, not as a tool to hedge risks, but as an instrument used by oligopolists to strategically affect market outcomes. It is shown that in a Cournot setting, generation firms sell forward contracts in order to commit to compete more aggressively in the spot market. Hence forward contracts make markets more competitive. Willems (2006) shows that a similar mechanism is at work with financial call options: the market equilibrium is even more competitive than with future contracts.

Green (2003) studies the combined hedging and strategic roles of forward contracts while at the same time examining different types of competition in the retail market. He shows that retail competition may lower the amount of forward contracts firms will sign. The current paper does not allow for retail competition, – consumers cannot switch retail supplier – and assumes, as in Bessembinder and Lemmon (2002), that retail prices are fixed.

Green (2007) models investment decisions and the technology choice in a long-term oligopolistic equilibrium model with risk averse firms in which firms can sign forward contracts. The setting of our paper allows us to analyze the relation between market completeness and investment decisions of generators.

6.3 Model description

We extend the competitive market equilibrium model of the forward and spot markets developed by Bessembinder and Lemmon (2002). The main difference with their model is that we allow for multiple financial products to be traded on the market. We start with a description of the spot market and continue with a description of the derivatives markets.

We consider an electricity sector with N_g identical generation firms and N_r

identical retailers. Each of the N_g generation firms is assumed to have a total production cost with a fixed and a variable component:

$$\widetilde{F} + \widetilde{a}\frac{\widetilde{Q}^c}{c} \tag{6.1}$$

where $\widetilde{F}, \widetilde{a}$ and c are parameters that determine the shape of the cost function, and $\widetilde{Q}$ is the production level of an individual firm. The total production cost of the industry is given by:

$$C(Q) = F + a\frac{Q^c}{c} \tag{6.2}$$

with $F = N_g\widetilde{F}$ and $a = \widetilde{a}/N_g^{c-1}$ and Q the total production of the industry.

Demand for electricity D is inelastic and stochastic. The spot market is perfectly competitive, and the wholesale price for electricity P is determined by market clearing:

$$P = C'(D) = a\,D^{c-1} \tag{6.3}$$

Each generation firm produces D/N_g. As demand is a random variable, so is the spot price.

The combined profit of the generators is equal to spot market revenue minus production costs:

$$\pi_g = P \cdot D - C(D) \tag{6.4}$$

Retailers buy energy on the spot market and sell it at a fixed retail rate R to consumers.[12] Each retailer supplies a volume D/N_r. The combined profit of the retailers is equal to:

$$\pi_r = (R - P)D \tag{6.5}$$

Both retailers' and generators' profits are affected by the stochastic nature of demand.

In the derivatives market, a derivative $i \in \{1, .., I\}$ is traded at a price F_i. The derivative promises a payment $T_i(P)$, which is conditional on the spot price P. This paper assumes that the only derivatives which are traded are call options. Hence:

$$T_i(P) = \max(P - S_i, 0) \tag{6.6}$$

with S_i the strike price of option i. A derivative with strike price zero corresponds to the standard forward contract.

The combined profit Π_j $(j = r, g)$ that is made by retailers and generators, respectively, when the retailers/generators buy a total of k_i^j derivatives in the derivatives market, is equal to:

$$\Pi_j = \pi_j(P) + \sum_{i=1}^{I} k_i^j \cdot (T_i(P) - F_i) \tag{6.7}$$

[12] The fixed rate R is either a regulated rate, or a fixed price contract offered to customers in a deregulated market. The case of real-time pricing, which would allow retailers to transfer upward price risk to the consumers, is discussed in Section 6.5.

The firms' profit is the sum of the profit they make in the spot market, and the profit they make on the derivatives they have bought. Both terms are stochastic as they depend on the realization of the demand level.

We assume that retailers and generators are risk averse, and that the utility of individual retailers and generators can be described by a mean-variance utility function with risk aversion parameter $N_j A$ ($j = r, g$). The risk aversion parameter contains N_j to account for the fact that a larger number of firms would lead to a smaller average size per firm, and therefore a proportionally smaller risk-bearing capacity, i.e. a higher absolute risk aversion. When $N_g = N_r$, the risk aversion of all firms (both generators and retailers) is the same, a reasonable assumption. If U_r and U_g represent the utility of retailers and generators, respectively, then each identical individual firm will maximize its utility U_j/N_j:

$$\frac{U_j}{N_j} = \mathrm{E}\left(\frac{\Pi_j}{N_j}\right) - \frac{N_j A}{2}\mathrm{Var}\left(\frac{\Pi_j}{N_j}\right) \qquad j = r, g \tag{6.8}$$

Maximizing (6.8) is equivalent to maximizing the following:

$$U_j = \mathrm{E}(\Pi_j) - \frac{A}{2}\mathrm{Var}(\Pi_j) \qquad j = r, g \tag{6.9}$$

which has the intuitively appealing benefit of not containing N_j anymore. One could say that the risk aversion parameter A measures the risk aversion of either the generation sector or the retail sector as a whole. We can proceed with the analysis as if there was only one generator and one retailer. Aggregate market welfare W is equal to the sum of the utility of retailers and generators: $W = U_r + U_g$.

In the contracting stage, firm j maximizes its utility U_j, by choosing the amount of derivatives $k_1^j, ..., k_i^j, ..., k_I^j$ it buys or sells. The equilibrium contract positions are given by:

$$\vec{k}^j = \Sigma^{-1}\frac{\mathrm{E}(\vec{T}) - \vec{F}}{A} - \Sigma^{-1}\mathrm{Cov}\{\pi_j, \vec{T}\} \tag{6.10}$$

with $\vec{k}^j = (k_1^j, ..., k_I^j)$, the vector of equilibrium quantities bought by player j, $\Sigma = \mathrm{Cov}\{\vec{T}, \vec{T}\}$ the I by I covariance matrix of the contracts $\vec{T} = (T_1, ..., T_I)$, $\vec{F} = (F_1, ..., F_I)$ the derivative price vector, and $\mathrm{Cov}\{\pi_j, \vec{T}\}$ the 1 by I covariance matrix of contracts and firm j's profit.

Equation (6.10) shows that the amount of contracts firm j buys is the sum of two terms. The first term is the pure speculative amount of contracts a firm would like to buy. If a financial derivative has an expected positive return, then the firm will buy some of it, as long as it does not increase the variance of its portfolio too much. The second term is the pure hedging demand by the firm. A firm j will buy derivatives in order to hedge its profit risk. It will buy more of a certain derivative, if it is more correlated with the profit it wants to hedge, and if the impact on the variance of the portfolio is smaller.

In equilibrium the demand and supply of derivative products should be equal. Hence, if there are no speculators active in market $i$we find:

$$k_i^r + k_i^g = 0 \tag{6.11}$$

and using equation (6.10) the equilibrium price of derivative i is given by:

$$F_i = \mathrm{E}(T_i) - \frac{A}{2}\mathrm{Cov}\{\pi_g + \pi_r, T_i\} \tag{6.12}$$

Hence, the price of a derivative is equal to the expected pay-off of the derivative minus a term which reflects the fact that the derivative is used to hedge the risk of the individual firms. The last term depends on the risk aversion of all the firms and the covariance of industry profit with financial instrument i. It is worth noting that the price of the derivative does not depend on the number of products traded in the market.[13]

If risk neutral speculators are active in derivatives market i, then the risk premium becomes zero, and the price of the derivative should be equal to its expected value:

$$F_i = \mathrm{E}(T_i) \tag{6.13}$$

6.4 Model data

The model is calibrated on the German electricity market, using technical and market data recorded in the first two months of 2006. Note that the purpose of the calibration is to allow us to perform simulations that produce intuitively relevant results. The numbers thereby serve as an illustration – this paper does not claim to make exact statements about the impact of option trade on the German electricity sector.

The marginal production cost curve $C'(Q)$ is calibrated on the actual German marginal production cost curve, as explained in Appendix 6.A. Demand is assumed to be normally distributed with mean 60 GW (which is the average of the observed sample) and standard deviation 17 GW. The standard deviation is chosen in such a way that the standard deviation of the resulting power price (according to equation (6.3)) corresponds to the standard deviation of the sample of observed prices. Given the assumptions about supply and demand, we can derive the wholesale price distribution. The distribution has a mean of 48 EUR/MWh and a standard deviation of 35 EUR/MWh. Bessembinder and Lemmon (2002) show that as the industry marginal cost function is convex, the price distribution is skewed.

[13]In standard mean-variance settings, risk pricing is not affected. Specifically, in quadratic or CARA-normal economies, the price of any risky security relative to the bond is unaffected by changes in the span. See Oh (1996).

Retailers and generators have the same risk aversion parameter $A = 0.0025$, which has the unit (h/1000 EUR). Furthermore, we assume that the fixed cost parameter $F = 1200$ (expressed in 1000 EUR/h), and that retailers sell their energy at a fixed price of 58 EUR/MWh. Note that prices and quantities are expressed in (EUR/MWh) and (GW), respectively, and hence profits and total costs are expressed in (1000 EUR/h).

6.5 Welfare effects

In this section, we use the model to calculate the optimal hedging strategy of generators and retailers, and analyze the welfare effects of adding additional derivatives to the market. In the first part of the simulations we assume that *no speculators* are active on the market, and hence supply and demand of financial contracts is only from retailers and generators. We consider four scenarios with a different number of derivative markets present. In Scenario 1, only a forward market exists. In Scenarios 2 through 4, the forward market is supplemented with one, three, and eleven additional option markets, respectively.[14]

Table 6.1 shows the simulation results for all scenarios. It shows for each of the twelve derivative contracts the net amount traded by generators and retailers. Positive numbers represent long positions, negative numbers represent short positions. The option contracts have strike prices ranging from 0 to 143 EUR / MWh, with the zero strike price (contract 1) corresponding to the forward contract. The range of option strike prices covers the 95% confidence interval of price levels.

The results show that if there are only forward contracts, firms overhedge their positions. Generators sell 68 GW forward, while in expected terms they will only produce 60 GW. The intuitive explanation for this is that generators and retailers want to hedge *volumetric* risk (or *quantity* risk), in addition to price risk. If there were only price risk (i.e., the quantity of electricity demanded would be deterministic), then forward contracts – which are specifically suited for hedging price risk – would be sufficient. The number of forward contracts would exactly correspond to the deterministic demand quantity. However, in the setting of this paper (and in reality), generators and retailers are exposed to both volumetric risk and price risk, because both quantities and prices are stochastic. If no options are traded, then volumetric risk can be hedged using additional forwards, because price and quantity are positively related. Another way of explaining this effect is that, because price and quantity are positively related, overall risk exposure is convex in the underlying state variable (demand) and hence the number of forward contracts exceeds expected demand. The price of the forward contract is 45.3 EUR/MWh, which is below the expected spot price of 48 EUR/MWh.

[14]The numerical model is written as a Mixed Complementarity Problem (MCP) in GAMS. See Appendix 6.B.

Table 6.1: Market equilibrium without speculation

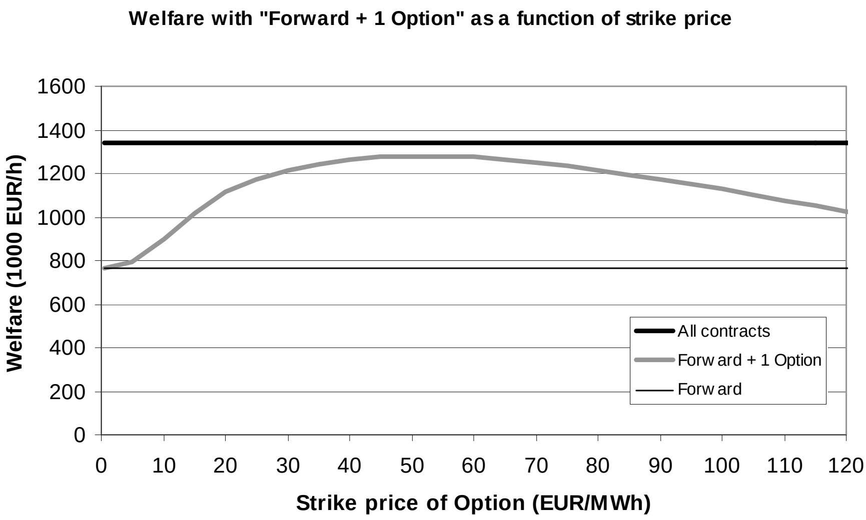
			Net Contract Position							
			Scenario 1 Forward		Scenario 2 Forward + 1 Option		Scenario 3 Forward + 3 Options		Scenario 4 All Contracts	
(i)	S_i	F_i	k_i^g	k_i^r	k_i^g	k_i^r	k_i^g	k_i^r	k_i^g	k_i^r
1	0	45.3	-68.0	68.0	-52.0	52.0	-32.1	32.1	5.3	-5.3
2	13	33.6							-37.9	37.9
3	26	25.2							-16.2	16.2
4	39	19.1					-37.7	37.7	-11.1	11.1
5	52	14.5							-8.3	8.3
6	65	10.9							-6.9	6.9
7	78	7.9			-45.4	45.4	-15.9	15.9	-5.9	5.9
8	91	5.6							-5.2	5.2
9	104	3.8							-4.3	4.3
10	117	2.4					-14.5	14.5	-6.6	6.6
11	130	1.3							3.7	-3.7
12	143	0.7							-13.1	13.1
	Welfare	W	**768**		**1224**		**1322**		**1337**	

In Scenarios 2 to 4, extra financial instruments are added to the market. Table 6.1 shows that once more instruments become available, generators reduce the amount of standard forward contracts they sell and substitute these contracts with option contracts. The generators and the retailers reduce their supply and demand of forward contracts. Although both demand and supply functions shift, the price of the forward contract remains 45.3 EUR/MWh as shown in derivation (6.12). As we pointed out in footnote 13, this effect is due to the use of the mean-variance utility function.

The last row in Table 6.1 is the aggregate welfare, measured in certainty equivalents (1000 EUR/h). Increasing the number of contracts traded clearly increases market efficiency. The introduction of one option contract, when none existed before, increases welfare by approximately 50 %. Adding extra markets for option contracts increases welfare further, but to a lesser extent. For instance, increasing the number of option markets from 3 to 11, increases welfare by 1.2 %. Hence risk sharing between generation and retail is close to optimal once one option contract (or a few option contracts) are traded. In fact, the welfare effects in Scenario 2 (Forward + 1 Option) can be improved even further, by modifying the strike price of the one available option contract. The strike prices of the option contracts in the simulations of Table 6.1 are chosen so that they span the 95% confidence interval of price levels. Scenario 2 assumes that one option is available, with a strike price roughly in the middle of this price range. Figure 6.1 shows how welfare would change when a different strike price is used.

If the strike price of the option is very low, then welfare is equal to welfare obtained with a forward contract only, because the option does not add

Figure 6.1: Welfare obtained with “Forward + 1 Option”, for different strike prices (compared with welfare obtained with only a “Forward” and with “All contracts”)



any new hedging possibilities. Welfare reaches a plateau optimum of 1280 (expressed in 1000 EUR/h) when the strike price is in the 50-55 EUR/MWh range. However, for all strike prices in the 30-80 EUR/MWh range, adding the one option contract to the forward contract lifts welfare above 1200 (in 1000 EUR/h), thereby capturing more than 75% of the potential benefits of market completeness.

Until now, we have assumed that retailers sell at a fixed price R, which can be either a regulated rate, or a fixed price contract offered to customers in a deregulated market. Given the continuous development of more sophisticated metering systems, it is interesting to consider what would happen if real-time pricing were possible. If all consumer contracts were based on real-time prices, then this would eliminate all risk for the retailers. However, generators would still have a desire to hedge. Smart retailers could therefore develop structured consumer contracts that take away risks from the generators and transfer them to the consumers who are willing to take on the risks. Such consumers would be rewarded with a lower expected power price. In its simplest form, such a structured contract could be similar to a fixed price contract. If options are available on the wholesale market (in addition to forwards) then more sophisticated structures would be possible, thereby hedging the generators' risk better and better, and improving welfare. In practice, in order to preserve the demand incentives created by real-time pricing, such structured contracts are still likely to price a consumer's individual demand based on real-time prices. However, at the end of each period, consumers could expect a check that settles the structured part of the consumer contract, i.e. the hedge, with the amount of the check depending on the overall demand and price developments in the spot market in the course of the period.[15]

For the second part of the simulations, we assume that *speculators* can actively participate in the market, by taking positions in the electricity derivative markets and financially closing their position in the spot market. We assume they trade away the risk premia in the market: the price of the derivatives becomes equal to the expected value of the derivative. As speculators provide extra liquidity to the market, the supply of derivatives by generators does no longer need to exactly balance the demand by retailers. The difference of generators' supply and retailers' demand is the position speculators take in the market. For the same four scenarios as before, Table 6.2 gives the net position of generators and retailers. In Scenario 1, only forward contracts exist, and generators sell 69.1 GW forward, retailers buy 67 GW, and speculators buy 2.1 GW. The results indicate that the more derivatives markets are introduced, the larger the gap between supply and demand for forward contracts, and the larger the role played by speculators. In Scenario 4, in which there is one

[15] Borenstein (2007) discusses retail markets and hedging in more detail. He shows how retail contracts can be developed that base the marginal price of electricity consumption on the real time price, but at the same time include a hedge which reduces monthly bill volatility.

forward market and eleven option markets, generators sell 34.2 GW, retailers *sell* 44.8 GW and speculators buy 79 GW.

The introduction of speculators increases welfare, as the players can share their risk with players outside the market, the speculators. Hence, the addition of speculators does not change our previous conclusions. Speculators play an active role in the electricity market by taking up market risk and by decreasing the risk premia in the market. As the number of markets increases, the amount of risk that speculators take away from market participants increases, but the positive welfare effect of additional markets levels off after a few products.

Table 6.2: Market equilibrium with speculation

			Net Contract Position							
			Scenario 1 Forward		Scenario 2 Forward + 1 Option		Scenario 3 Forward + 3 Options		Scenario 4 All Contracts	
(i)	S_i	F_i	k_i^g	k_i^r	k_i^g	k_i^r	k_i^g	k_i^r	k_i^g	k_i^r
1	0	48.4	-69.1	67.0	-58.7	45.3	-48.0	16.2	-34.2	-44.8
2	13	36.0							-13.2	62.6
3	26	25.9							-8.8	23.6
4	39	18.1					-20.5	55.0	-6.6	15.7
5	52	12.4							-5.5	11.2
6	65	8.3							-4.7	9.1
7	78	5.5			-29.3	61.5	-12.2	19.7	-4.2	7.6
8	91	3.5							-3.8	6.6
9	104	2.1							-3.2	5.4
10	117	1.2					-10.5	18.4	-4.9	8.2
11	130	0.7							2.7	-4.7
12	143	0.3							-9.9	16.4
	Welfare	W	**772**		**1285**		**1403**		**1423**	

Finally, for the third part of the simulations, we repeat the previous simulation but we now use a different assumption for firms' utility functions: instead of the utility functions from equation (6.7), we use the well-known *CRRA utility function* (i.e., the utility function with *constant relative risk aversion*). As a result of the CRRA property, firms become very averse of potential shocks that would lead to very low or negative profits. In other words, the CRRA utility function models a world in which firms want to avoid the risk of liquidity problems or bankruptcy. Practically, we choose the coefficient of relative risk aversion to be 4, which is in the middle of the typical 2-6 range (see e.g. Palsson, 1996). The simulation results with speculation and CRRA utility functions for producers and retailers are shown in Table 6.3. Generally speaking, the results are very similar to the results of Table 6.2, although the retailer seems to have a slightly increased preference for options over forwards (as compared to Table 6.2).

The most interesting observation is that with the CRRA utility function, it is not possible to find a sufficiently hedged solution in the case when only forwards are present. In other words, if no options are introduced, welfare remains 'infinitely low' for CRRA utility functions. The reason is that forwards alone do not allow the retailer and the generator to limit their exposure in all 'states-of-the-world'. The intuition for this effect is the following: since a negative result in one potential state-of-the-world is strongly penalized by the CRRA function, the retailer and the generator would like to avoid – at all costs – any outcomes in which their profit is below a certain threshold, in order to avoid bankruptcy. The retailer faces a negative shock when demand is high (it faces a high wholesale price, and has to buy a large volume of power), and when demand is low (sales volume is too low to cover fixed costs). The generator faces a negative shock when demand is very low (low price and low volume). As the retailer wants to avoid bankruptcy at all cost, its demand for forward contracts is undetermined for any price of forward contracts. With only the forward contract, the retailer is unable to hedge against both the risk of having high demand and the risk of having low demand. Based on these results for a CRRA utility function, it is clear that the introduction of options is especially welfare-enhancing if there is a strong risk aversion for negative shocks that could lead to bankruptcy.

A practical implication of this phenomenon would be that a retailer alone would have difficulty to survive if no liquid option market is available. Anecdotal evidence of this effect is the case of Centrica in the UK. After the demerger of British Gas (Centrica, 12/2/1997), Centrica was essentially a gas retailer in the UK. At the end of 1997, Centrica entered the electricity market as a pure retailer (without any generation assets) and acquired its first electricity customers (Centrica, 1/12/1997). Rather than staying a stand-alone gas and electricity retailer, Centrica started to invest in gas-fired power generation in 2001 (Centrica, 29/5/2001 and 24/8/2001). Centrica stated the following reason for the investments in power generation: "*As part of its risk management strategy the company has said it plans to source 20-25 per cent of its future peak electricity requirements from its own generating capacity. This strategy offers increased long term stability and protection against electricity price fluctuations and spikes*" (Centrica, 24/8/2001). In other words: the investment in gas-fired power generation is meant primarily to protect the retailer against electricity price volatility in case of peak demand. From a financial perspective, an investment in gas-fired power generation (which has relatively low investment cost and relatively high marginal cost) can be considered as the purchase of a call option on electricity, with a relatively high strike price[16]. The absence of a market for such options forces retailers to invest in the physical equivalent, because staying unhedged is not a viable alternative.

[16] Strictly speaking, a gas-fired power plant is an option on a (clean) spark spread, i.e., the difference between the electricity price and the input prices (gas, carbon emission rights).

Table 6.3: Market equilibrium with speculation, with CRRA utility functions

			Net Contract Position							
			Scenario 1 Forward		Scenario 2 Forward + 1 Option		Scenario 3 Forward + 3 Options		Scenario 4 All Contracts	
(i)	S_i	F_i	k_i^g	k_i^r	k_i^g	k_i^r	k_i^g	k_i^r	k_i^g	k_i^r
1	0	48.4	No solution		-58.8	38.7	-48.1	11.3	-33.5	-45.8
2	13	36.0							-13.4	63.6
3	26	25.9							-10.8	23.7
4	39	18.1					-20.5	59.6	0.0	15.7
5	52	12.4							-18.7	11.2
6	65	8.3							6.1	9.1
7	78	5.5			-29.3	63.8	-12.1	20.2	-3.7	7.9
8	91	3.5							-3.8	5.6
9	104	2.1							-9.9	9.4
10	117	1.2					-10.5	18.2	-2.5	-9.6
11	130	0.7							-7.3	49.6
12	143	0.3							10.4	-47.0
	Welfare	W	**"- ∞"**		**-1216**		**-349**		**-314**	

6.6 Investment decisions by small firms

Above we have shown that the welfare effect of adding contracts levels off after a relatively small number of contracts. However, implicitly we have assumed that all production firms have a diversified portfolio of generation plants. Indeed: we assume that there are N_g identical production firms, which implies that the production cost curve of each firm is just a horizontally scaled version of the aggregate production cost curve of the generation industry. In other words, the portfolio of each firm contains power plants with relatively low marginal costs (e.g., nuclear power plants), which will be run in nearly all demand scenarios, and power plants with relatively high marginal costs (e.g., gas-fired power plants), which will be run only if demand is high. As a result, all generation firms are reasonably diversified, and the demand/price risk is adequately distributed across generation firms. However, if some firms have only base load power plants and other firms have only peak load power plants, then some firms' financial results are much more sensitive to certain demand/price scenarios. Intuitively, this could make the potential social value of a comprehensive set of financial contracts (which would allow risk-transfer under accurately defined demand/price scenarios) significantly higher.

In order to test the impact of market completeness when firms have different types of portfolios, we analyze how risk trading modifies investment behavior. Specifically, we determine whether a small, non-diversified firm would invest in a single power plant with marginal cost c and fixed investment cost F.[17] The

[17] Since such a power plant is in fact (almost) equivalent to an option contract, this essentially means that we analyze the implications of providing one player with one additional

small firm is assumed to be risk averse, with mean-variance utility function (as in the first part of our simulations). We assume no speculators in the market.

The firm invests in this production plant if the investment increases its expected utility. The expected utility without investments is equal to

$$\begin{aligned} U^{NI} &= \max_{k_1,\dots,k_I} \left[\mathrm{E}\{\pi\} - \tfrac{A}{2}\mathrm{Var}\{\pi\}\right] \\ \text{with} \qquad \pi &= \textstyle\sum_{i=1}^{I} k_i \cdot (T_i(P) - F_i) \end{aligned} \tag{6.14}$$

while the expected utility with investments is equal to

$$\begin{aligned} U^{INV}(c,F) &= \max_{k_1,\dots,k_I} \left[\mathrm{E}\{\pi\} - \tfrac{A}{2}\mathrm{Var}\{\pi\}\right] \\ \text{with} \qquad \pi &= \textstyle\sum_{i=1}^{I} k_i \cdot (T_i(P) - F_i) + (\max\{p-c,0\} - F) \end{aligned} \tag{6.15}$$

The firm invests as long as

$$U^{NI} > U^{INV}(c,F) \tag{6.16}$$

Equation (6.16) defines implicitly the maximum fixed cost for which the firm is willing to invest in new generation capacity with marginal cost c. Hence investment occurs as long as

$$F < F^{cr}(c) \tag{6.17}$$

Therefore, the function $F^{cr}(c)$ represents the investment behavior of the firm.[18] $F^{cr}(c)$ obviously depends on the number and types of financial contracts traded in the market, and on the risk aversion of the firm. Indeed, generally speaking, as more contracts are traded, the firm is able to better hedge the output of the production plant, thereby reducing its risks. This makes it more interesting for the firm to build a power plant. In certain cases, there is, however, a non-monotonic relation between market completeness and investments decisions, as will see below. Let $\overline{F}^{cr}(c)$ denote the case of market completeness, i.e., the firm's investment behavior if a full set of option contracts is available.

Figure 6.2 compares the decision behavior of risk averse firms in case of market incompleteness with the decision behavior in case of market completeness. More specifically, the figure shows the "adjustment factor" κ:

$$\kappa = \frac{F^{cr}(c)}{\overline{F}^{cr}(c)} \tag{6.18}$$

When calculated for different levels of market completeness, the factor κ describes how the firm's investment decisions change as markets become more

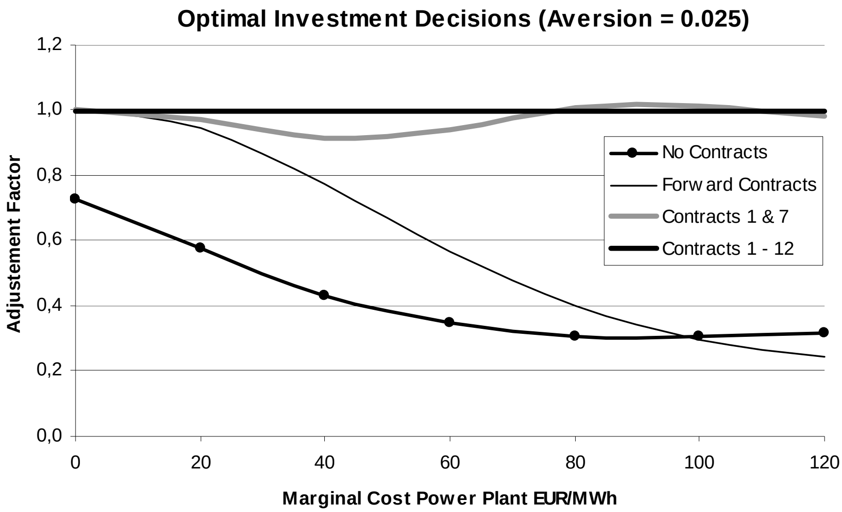
contract.

[18] If we consider the power plant as an option contract, then $F^{cr}(c)$ is the maximum price (option premium) the firm would be willing to pay for this option contract.

complete. Those investments decisions are the profit-maximizing decisions for the firm (conditional on the available contracts). Note that the firm's risk aversion is chosen at a higher level than before, because we analyze a small firm and the mean-variance utility function does not scale.

If $\kappa < 1$, the firm invests less in a particular type of generation than if markets were complete. If $\kappa > 1$, the firm invests more in a particular type of generation than if markets were complete.

Figure 6.2: Optimal investment decisions



The effect of increasing market completeness is very different for base load plants on the one hand, and peak load plants on the other hand. Once the forward contract is introduced, a firm with a *base load plant* ($c \approx 0$) would be able to hedge its position completely. Adding additional derivatives to the market does not change the investment decisions for a base load power plant, as the firm already has perfect information in evaluating the value of the power plant using the forward contract. Speculation and investment decisions are decoupled: a firm wanting to invest in a base load power plant can do so without taking any market risk (i.e., it can focus on the operational aspects), while a speculator can decide to assume some base load market risk without actually having to build a power plant.

For *peak load plants* (c large), the results are quite different. Once the forward contract is introduced (and no options), investment in certain peak load power plants with very high marginal cost (higher than 100 EUR/MWh in

Figure 6.2) may actually be *less* than if no contracts are traded. The reason for this is that it may be more profitable for the firm to speculate on the forward market (without building a power plant), than to build a power plant and use financial contracts to hedge its portfolio. Hence, financial investments 'crowd out' the investments in physical assets: investment and speculation decisions are coupled. As more and more contracts are introduced, we see that investment in peak generation increases dramatically, because there are better instruments to hedge the risk of the production output of the firm.[19] As a result, the investment decision and the speculation decision become decoupled again. The figure also shows that for the technology with marginal production costs around 78 EUR/MWh, adding additional contract markets on top of contract number 7 does not change the results. Contract 7 has a strike price of approximately 78 EUR/MWh, hence the investment valuation of the firm is perfect, regardless of any additional contracts being added.

Note that in certain cases the adjustment factor κ might be larger than one, which implies that a firm might invest more when markets are incomplete than when markets are complete. This may happen when the investment increases the risk of the existing firms in the sector. In that case it would be cheaper for the firm to buy a financial option with an equivalent strike price, than to invest in physical capacity. Such a financial option would be available at a 'depressed' price, i.e. a price below its expected value, because it reduces risk of the existing firms. Put otherwise, the availability of an extra derivative market creates additional investments opportunities for the firm. If those opportunities are very profitable, then the firm uses its capital to speculate on the derivatives market, instead of investing it in new power plants. In other words, the opportunity cost of risk-bearing capital has increased with the availability of new investment opportunities. Similar to what we have observed for peak load plants when a forward contract is introduced, we see that financial investments 'crowd-out' physical investments.[20]

[19]It is important to note the difference with the Centrica case, in which – as mentioned in Section 6.5 – peak investment was due to a *lack* of tradable call options. Before the investment, Centrica was a pure *retailer*, hence it had a natural *short* position in peak electricity. If tradable call options had been available, such a retailer would have closed its position by taking a long position (i.e. buying) in call options. Since these call options were not available, Centrica invested in physical peak generation in order to close its position. In contrast, Section 6.6 studies the incentives of a small *generator*, who is considering making an investment in power generation, which would give it a natural *long* position in peak electricity. If tradable call options are available, such a firm invests in physical peak generation and closes its position by taking a short position (i.e. selling) in call options. If tradable call options are not available, then the only way to keep a closed position is to *not* invest, which leads to the underinvestment in peak generation that one can observe in Figure 6.2 for the case where no contracts or only forward contracts are traded. In summary, the availability of tradable call options increases investment incentives for generators, who then sell call options to retailers. A lack of tradable call options leads to underinvestment by generators and leaves retailers with no other alternative than to become an integrated generator-retailer themselves.

[20]Crowding out of physical investments can only happen in incomplete markets. Once a

The previous discussion assumes that the investor actually has access to all contracts traded in the market. It is interesting to consider the investment incentives of a completely new entrant who trades neither physically nor financially before investing. Suppose for instance that forward contracts (but no options) are traded in the market. For power plants with marginal cost smaller than 100 EUR/MWh, an investor who has no market access would underinvest, as can be seen from Figure 6.2. Indeed, given that the investor does not utilize the hedging possibilities offered by the market (i.e., the forward contract), his risk aversion reduces his investment, especially for plants that can be hedged very well using the available contracts (in particular: base load plants). The investor therefore foregoes potential profits by not having market access. The opposite happens for power plants with marginal cost larger than 100 EUR/MWh: investing in the power plant is in fact a less profitable use of the investor's 'risk budget' (i.e., his capital) than investing in forward contracts. An investor without market access would overinvest in power plants, and forego the potential profits of speculation with forwards.

6.7 Information content of derivatives prices: risk-free probabilities

In the previous section, we studied the effect of financial contracts on the investment decisions of a firm as a function of the technology parameter c. In this section we look at the information content contained in the prices of financial products, and derive optimal investment decisions based upon a typical financial approach using risk-free probabilities. This approach uses the price data from the financial market to estimate the risk-free probabilities and then computes the market value of an asset as its expected value under the risk-free probability measure. The investment decision is made by comparing the market value of the asset with the investment costs of the asset. In a sense, this approach measures the 'information content' of the derivatives prices. The approach assumes that the market is sufficiently complete to create a portfolio of contracts which replicates the pay-off of the physical asset. In this section we will test at which point markets are sufficiently complete to use the risk-free probabilities approach. We will compare the investment decisions based on the risk-free probabilities approach with the optimal decisions we found in the previous section.

The market equilibria in Table 6.1 can be represented by means of a risk-free probability distribution θ, different from the true distribution. Under the risk-free probability distribution, the contracts' prices are equal to their

power plant is fully hedged, crowding-out no longer occurs. The investment in the power plant becomes risk-free, and will therefore no longer put a burden on the risk-bearing capacities of the firm.

expected values:

$$\mathrm{E}_\theta(\vec{T}) = \vec{F} \tag{6.19}$$

and the generator and the retailer act as risk neutral agents who optimize expected profit:

$$j = r, g \qquad U_j = \mathrm{E}_\theta(\Pi_j) \tag{6.20}$$

Figure 6.3 shows the risk-free probabilities for different assumptions regarding the number of products being traded.[21] When all financial contracts are traded, the risk-free probability distribution assumes that extreme events, especially low prices, are more likely to occur than they do in reality. When only forward contracts are traded, however, the risk-free probabilities calculated on the basis of forward prices, give extreme events a too small probability. Adding just one extra financial market brings the distribution relatively close to the situation in which all 12 financial products are traded, and greatly improves the information that firms obtain.

In the risk-free probabilities approach, firms invest in a power plant when the expected net present value, calculated using the risk-free probabilities, is larger than the fixed investment cost. Hence a firm invests if

$$F < NPV_\theta(c) = \mathrm{E}_\theta\{\max(p - c, 0)\} \tag{6.21}$$

As in Section 6.6 we compare these critical values with $\overline{NPV}_\theta(c) = \overline{F}^{cr}(c)$, the net present value of a power plant with marginal cost c calculated using the risk-free probabilities inferred when all contracts are traded. Figure 6.4 shows – for different types of generation plants – the ratio of both numbers. We use the index RF (risk-free probabilities) to distinguish the result from the previous section:

$$\kappa^{RF} = \frac{NPV_\theta(c)}{\overline{NPV}_\theta(c)} \tag{6.22}$$

The results are similar to those obtained in Section 6.6. In order to see whether the information content of the prices is sufficient, we compare the decisions of the risk-free probabilities approach with the optimal decisions described in Section 6.6. Figure 6.5 shows the difference between the two

[21] As the set of forward and option markets is incomplete, the risk-free probability distribution is not uniquely determined by the observable market prices. We estimate the risk-free probabilities using a slightly modified version of the approach suggested by Stutzer (1996). The approach of Stutzer (1996) is essentially a Bayesian framework in which the risk-free probabilities are estimated using the historical real probabilities as a prior (thereby maximizing the observed entropy). In our model setting, we use the available real probabilities instead of 'historical' real probabilities. Furthermore, we replaced Stutzer's maximum-entropy objective function with the minimum-distance objective function suggested by Rubinstein (1994), in order to reduce computational complexity. Jackwerth and Rubinstein (1996) had already pointed out the computational challenges of the maximum-entropy objective function and had observed that results with a minimum-distance objective function are very similar.

Figure 6.3: Risk-free probabilities without speculators

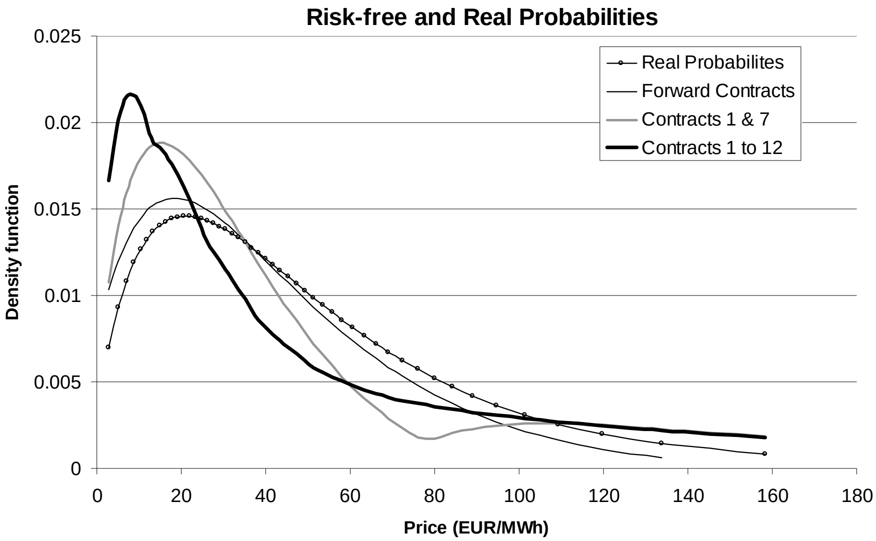
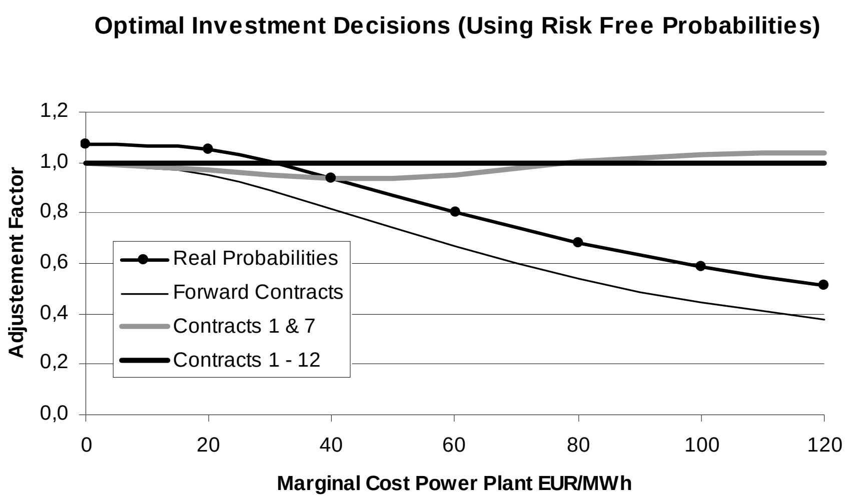


Figure 6.4: Investment decisions using risk-free probabilities



decision rules:

$$\Delta\kappa = \frac{NPV_\theta(c) - F^{cr}(c)}{\overline{NPV}_\theta(c)} = \kappa^{\mathrm{RF}} - \kappa \tag{6.23}$$

Figure 6.5: Difference of the risk-free probabilities approach and the optimal investments

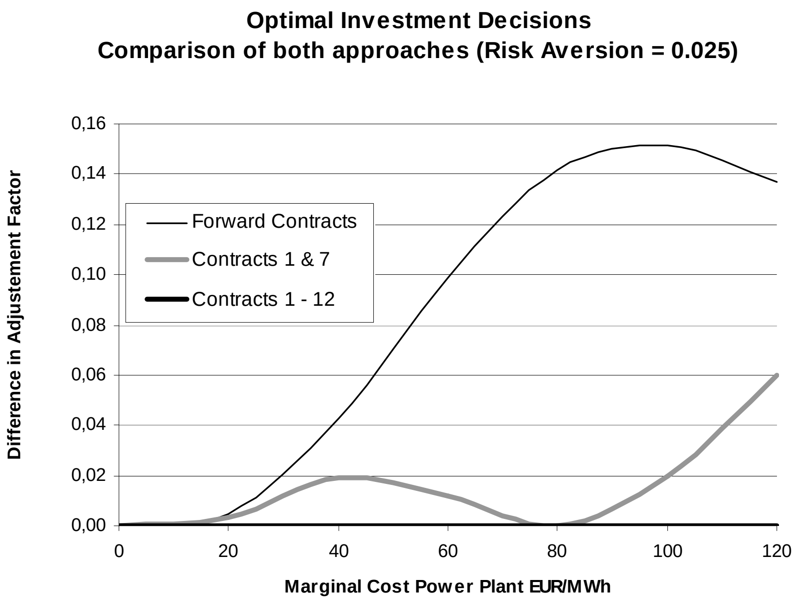


Figure 6.5 shows that if the marginal cost of the power plant is equal to the strike price of one of the contracts traded in the market, the two approaches produce identical results. In general, the error in using risk-free probabilities is smaller when the contracts traded correspond better to the risk profile of the power plant being built. Hence, the decisions to build base load power plants are always efficient when there is a forward contract. Similarly, if the market trades an option with a strike price very close to the marginal cost of a certain peak power plant, the investment decision for that power plant based on risk-free probabilities is optimal.

On the other hand, the risk-free probabilities approach leads to overinvestment in power plants for which no close financial substitutes are traded in the market. For example, if only forward contracts are available, the investment decision for a peak power plant can be seriously distorted. In our model setting firms may make an error of more than 15% when evaluating this decision. The addition of one option contract eliminates this error for almost all types of power plants. In general, as more contracts are being traded, the risk-free probabilities approach leads to decisions that are closer to the optimal decisions. When markets are complete, the two approaches yield identical results.

6.8 Conclusions

The contribution of this paper is that it studies the effects of market completeness on welfare and on investment incentives, in an equilibrium model of the spot and derivative markets in the electricity sector.

With respect to *welfare*, the numerical results of the model (calibrated with German market data) show that welfare is enhanced when options are offered in the market in addition to forward contracts. However, it turns out that most of the welfare benefits are achieved with one to three options. The need for options is especially relevant if firms have a strong aversion of liquidity problems (bankruptcy risk): with a CRRA utility function, aggregate welfare is infinitely low when no options are present. Allowing speculators to actively trade in the market, eliminates the risk premium, and increases aggregate welfare. The beneficial effect of speculators increases as more contracts are traded.

With respect to *investment incentives*, financial contracts are important for (i) hedging the risk of the entrant, and (ii) signaling to entrants how they could reduce the overall sector risk. When no financial contracts are traded, risk averse firms will tend to invest less than if a complete set of financial contracts is present. When forward contracts are traded, investment in base load power plants increases (investment and speculation are decoupled), but investment in certain peak load plants declines because it is more attractive to speculate with forward contracts instead (investment and speculation are coupled, and financial investments may 'crowd-out' physical investments). When options are added to the market, investment in peak power plants increases again dramatically (investment and speculation become more and more decoupled).

We test at which point markets are sufficiently complete so that a firm can base its investment decisions on a financial approach using *risk-free probabilities,* in other words, at which point the information content contained in the prices is sufficient for investment purposes. We show that as long as a perfectly matching contract is traded, the risk-free probabilities approach leads to optimal investment decisions. However, for power plants for which no close financial proxy is available, the risk-free probabilities approach can lead to significant overinvestment, especially for peak load power plants. As more and more options are added to the markets, the investment error decreases. This shows that the quality of the information contained in market prices improves as markets become more complete.

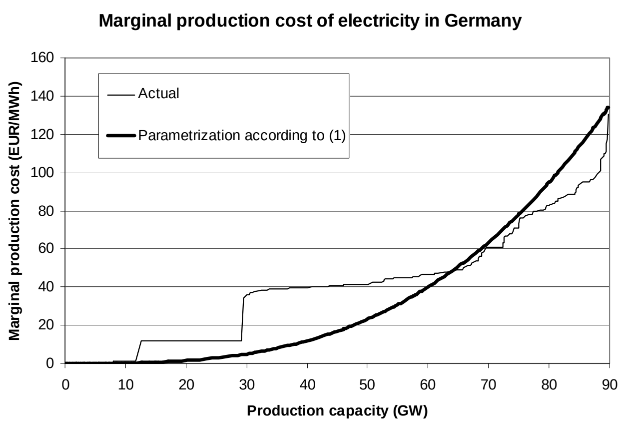
6.9 Acknowledgements

The authors would like to thank the participants of the workshop on "Policymaking Benefits and Limitations from Using Financial Methods and Modelling in Electricity Markets" in Oxford (July 2008) and seminar audiences in Tilburg, Kiev, Cologne and Ljubljana. Special thanks to the discussant Thomas Tangeras, as well as to three anonymous reviewers. Furthermore, the authors would like to thank the three anonymous referees of Energy Economics, who provided helpful and constructive feedback.

6.A Annex: Construction of the German marginal production cost curve

The marginal production cost curve $C'(Q)$ (see equations (6.2) and (6.3)) is calibrated on the actual German marginal production cost curve, as shown in Figure 6.6. The parameters chosen in order to obtain a reasonable fit, are $c = 4$, $a = 1.852\ 10^{-4}$.

Figure 6.6: Industry marginal cost



Generation capacities and ownership are obtained from VGE (2006). More than 300 power plants are considered, totaling 100 GW of generation. Wind, biomass and solar capacities are not considered within the firm's generation portfolios. Plant capacities are decreased by seasonal availability factors following Hoster (1996). Using a type-specific algorithm based on Schröter (2004) with construction year as proxy, we calculate a plant-specific efficiency to derive marginal costs. Fuel prices are taken from BAFA (2006) and resemble average monthly cross-border prices for gas, oil and coal. We include the price of CO_2-emission allowances in the cost estimate based on fuel type and plant efficiency. Allowance prices are taken from the EEX.

6.B Annex: Numerical model

The equilibrium model is solved as a Mixed Complementarity Problem (MCP) in GAMS. By writing the problem as a MCP, we can simultaneously determine the equilibrium prices and quantities of derivatives.

6.B.1 Spot market

The following equations are considered for the spot market:

$$\begin{aligned} C_s &= F + \tfrac{a}{c} D_s^c & P_s &= a\, Q_{gs}^{c-1} \\ \pi_{gs} &= P_s \cdot D_s - C_s & \pi_{rs} &= (R - P_s) D_s \end{aligned} \tag{6.24}$$

with s the state of the world and the indices $g$and r indicating generator and retailers.

6.B.2 Call option

The call option i pays T_{is} in state s, with $T_{is} = \max(P_s - S_i, 0)$.

6.B.3 Equilibrium in the derivatives markets

$$\begin{aligned} \Pi_{gs} &= \pi_{gs} + \textstyle\sum_{i=1}^{I} k_{gi} \cdot (T_{is} - F_i) \\ \Pi_{rs} &= \pi_{rs} + \textstyle\sum_{i=1}^{I} k_{ri} \cdot (T_{is} - F_i) \\ \mathrm{E}\Pi_g &= \textstyle\sum_{s=1}^{S} \rho_s \, \Pi_{gs} \\ \mathrm{E}\Pi_r &= \textstyle\sum_{s=1}^{S} \rho_s \, \Pi_{rs} \\ \textstyle\sum_{s=1}^{S} (1 - A\,(\Pi_{gs} - \mathrm{E}\Pi_g))\; \rho_s \; T_{is} &= F_i \\ \textstyle\sum_{s=1}^{S} (1 - A\,(\Pi_{rs} - \mathrm{E}\Pi_r))\, \rho_s \; T_{is} &= F_i \\ k_{ri} + k_{gi} &= 0 \end{aligned} \tag{6.25}$$

6.B.4 Risk-free probabilities

The risk-free probabilities are chosen such that the price of each financial instrument equals its expected value under the risk-free probability measure, and that the difference between the risk-free probabilities and the true probabilities is minimized:

$$\begin{array}{rcccl} \sum_{s=1}^{S} \theta_s \; T_{is} & = & F_i & \perp & \mu_i \\ \sum_{s=1}^{S} \theta_s & = & 1 & \perp & \omega \\ 2(\theta_s - \rho_s) - \omega - \sum_{i=1}^{I} \mu_i T_{is} & = & 0 & \perp & \theta_s \geq 0 \end{array} \tag{6.26}$$

Chapter 7

Risk Spillovers and Hedging: Why Do Firms Invest Too Much in Systemic Risk?

7.1 Introduction

This paper studies whether investments by a small competitive firm are socially efficient, when market outcomes are uncertain and financial markets are incomplete. We show that decisions about real investments, i.e. investments in productive assets, may be suboptimal, because the presence of those assets changes the distribution of overall industry risk, which, if financial markets are incomplete, creates a risk externality for the firms already active in the market. In particular, private investment decisions will lead to a market in which the industry as a whole takes too much risk by investing too much in production activities with highly correlated risk profiles. Firms that could reduce the overall risk of the industry by offsetting the aggregate risk, do no enter often enough, and firms that increase the overall industry risk, enter too often. It is important to note that if the entrant only invests in financial products and not in real physical assets, then its investment decisions will be socially optimal, even if the market is incomplete.

To illustrate the point, let us consider an industry in which the production cost is strongly dependent on an input factor of which the price is uncertain. As an example, we consider the production of tomatoes in heated greenhouses. The production costs are strongly dependent on the price ω of the gas that is used to heat the greenhouses. The gas price ω is determined on international markets and is considered an exogenous random variable in this analysis. We assume that the gas price is the only source of uncertainty: the state-of-the-world is fully defined by ω. Producer and consumer surplus are a function of ω. For simplicity, let us assume that the tomato market is perfectly competitive and that demand is completely inelastic. In this market, a higher gas price would result in higher costs for the marginal tomato producer, hence higher tomato prices. Therefore, *consumer surplus* declines rapidly as a function of the gas price. On the other hand, *producer surplus* increases slightly as a function of the gas price. Indeed, with higher gas price, the surplus of the marginal producer would not change, since the gas price increase is passed on to consumers through higher tomato prices. However, inframarginal capacity consisting of more efficient greenhouses that consume less gas would obtain higher profits when gas prices are higher, because the tomato price increases more than the gas cost per tomato. Consumer surplus $\mathbf{CS}$ and producer surplus $\mathbf{\Pi}$ are shown schematically as a function of ω in Figure 7.1. Since they depend on ω, both $\mathbf{CS}$ and $\mathbf{\Pi}$ are random variables.

Now, let us consider an entrant who invests in setting up a transport chain to import tomatoes from a warmer country. Due to transport costs, such imported tomatoes are only competitive when tomato prices are high enough. This is the case when the gas price ω exceeds a given price ω_e. Therefore, the advent of the entrant has an impact on consumer surplus and producer surplus of existing players in all states-of-the-world for which $\omega > \omega_e$. The effect is shown in Figure 7.1: due to the entrant, the producer surplus $\mathbf{\Pi}$ of existing producers is reduced to $\mathbf{\Pi}'$ while consumer surplus increases from $\mathbf{CS}$ to $\mathbf{CS}'$. For a very small entrant, the impact $d\mathbf{\Pi}$ of the entrant on $\mathbf{\Pi}$ is the exact opposite of the impact $d\mathbf{CS}$ of the entrant on $\mathbf{CS}$. Hence, the entrant does not cause a net change in total surplus of existing players (i.e. the sum of existing producer and consumer surplus). However, the entrant does cause a change in the *variation* in surplus. In this particular example, the variation in both producer surplus of existing players and consumer surplus is lower when the entrant is present, as can be observed in Figure 7.1. Hence, the entrant takes risk out of the market. If the existing producers and consumers are risk-averse, the entrant therefore has a positive externality on the industry. The entrant does not get rewarded for this, however. If the investment in the entrant's envisaged transport chain is marginally too expensive compared to the expected profits, the entrant would not invest, while it would be socially optimal to invest given the positive externality. Entry in the tomato transport option may be inefficiently low.

Conversely, consider an entrant who invests in a very inefficient old greenhouse technology with very high gas consumption but low other operating costs. Such a greenhouse would only produce if gas prices are low enough. There would only be an impact on $\mathbf{\Pi}$ and $\mathbf{CS}$ for gas prices lower than a given price level, say ω_e'. Above that gas price level, the entrant's greenhouse would not be competitive due to high gas costs. Again, the impact of the entrant

Figure 7.1: Illustration of the effect of an infinitesimal entrant on producer surplus of existing producers and consumer surplus, as a function of the gas price ω. The entrant *reduces* variation in outcomes for existing players.

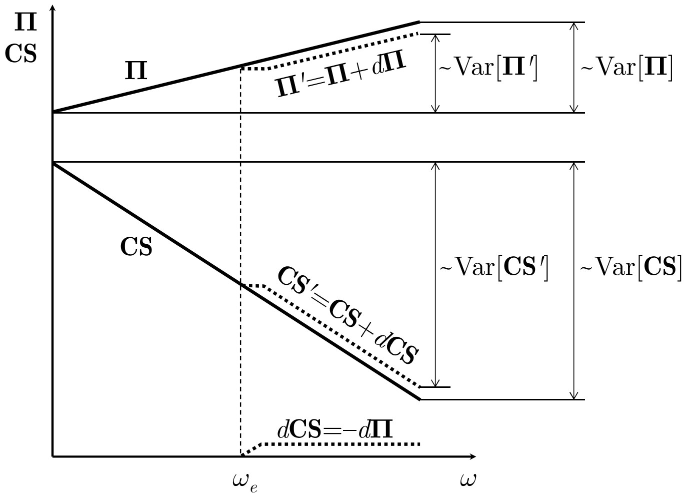
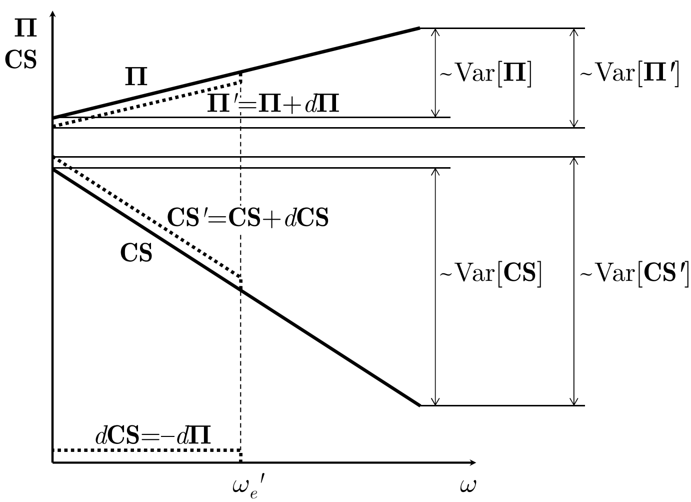


Figure 7.2: Illustration of the effect of another infinitesimal entrant on producer surplus of existing producers and consumer surplus, as a function of the gas price ω. The entrant *increases* variation in outcomes for existing players.



on $\mathbf{\Pi}$ and $\mathbf{CS}$ is symmetric and there is no change in total surplus of existing players if the entrant is small enough. However, in this case the entrant causes an *increase* in variation of producer and consumer surplus of existing players, as shown in Figure 7.2. Hence, the entrant has a negative externality on the existing producers and consumers, if they are risk-averse. The entrant does not 'see' this external cost when making the investment decision. If the investment in the entrant's envisaged greenhouse is marginally lower than the expected profits, the entrant would invest, while it would be socially optimal not to invest given the negative externality. Entry in the old greenhouse technology may be inefficiently high.

The problem is therefore one of incomplete property rights to the risk of existing producers and consumers. If, as in the traditional Coase (1960) approach, property rights could be assigned and enforced, the inefficiency would disappear. In practice this is hard to do. There is however another way to eliminate the efficiency, namely by introducing financial markets for risk-sharing. If there are financial instruments – e.g. a combination of options on the gas price – that have a pay-off similar to the impact $d\mathbf{CS} = -d\mathbf{\Pi}$ of the entrant, then producers and consumers can optimally share the risk between them. Any impact of the entrant could be hedged, and there would be no externality. If markets are complete, then such a combination of financial instruments exists for all types of entrants. If markets are incomplete, only a limited set of risk-sharing instruments is available, and the impact of the entrant cannot be perfectly replicated with the available instruments. With incomplete markets, part of the impact of the entrant may be hedged, but there is no guarantee that the inefficiency decreases or disappears.

The results of our discussion are relevant for specific sectors such as electricity or oil, in which investment costs are significant, financial markets do not cover all potential contingencies[1] and firms can choose between different technologies or locations with different risk profiles. In such situations, the industry as a whole becomes too risky. In those cases sector-specific regulation might be necessary. Sector-specific regulation could take the form of entry-regulation – if the regulator is capable of adequately measuring the risk – or incentives for the creation of markets of additional financial instruments.[2]

The fact that entry decisions might be socially inefficient in an *oligopolistic* market structure is well known. The most obvious case is the case of entry deterrence by oligopolistic incumbents, a topic that has been studied extensively

[1] Markets might be incomplete because productive assets have a long lifetime, or because some sources of risks are non-tradable. For instance, there might not be financial instruments to hedge regulatory uncertainty.

[2] One may wonder what prevents industry actors from creating those markets by themselves. In many cases, however, it turns out that markets are not sufficiently liquid to constitute a realistic hedging solution. Even the market for oil futures or forwards becomes relatively illiquid for delivery dates that are more than a few years out. Industry players often need to hedge through vertical integration, as in the Centrica example described by Willems and Morbee (2010).

in the industrial organization literature, following the seminal work by Bain (1949), Sylos Labini (1969) and Modigliani (1958). With entry deterrence, there is *too little* entry from a welfare standpoint. For example, Spence (1977) – later extended by Dixit (1980) and Schmalensee (1981) among others – shows that entry deterrence through prior capacity commitments by the incumbent may result in larger costs than are necessary for a given output level, and higher prices. Our paper does not consider the preemptive strategic actions of incumbents and focuses on the potential entrant's investment decision. In such a context, one finds not only cases with too little entry, but also cases with *excessive* entry: von Weizsäcker (1980) shows that there are plausible parameter configurations under which welfare would be improved by limiting entry. Similar to Mankiw and Whinston (1986), we use a two-stage model with capacity investment decisions by entrant(s) in stage one and actual production in stage two. As Mankiw and Whinston (1986) point out, suboptimal entry is due to the fact that the entrant's evaluation of the desirability of his entry is different than the 'social planner's' evaluation – a phenomenon one could call *investment externalities*. In the analysis by Mankiw and Whinston (1986), the externality is due to 'business-stealing' from other players: the entrant will gain some profit by reducing the profit of the existing players. This leads to a redistribution of industry profits, but not necessarily to an increase in the total surplus of the industry. Note that the business stealing effect disappears if the stage-two game is perfectly competitive and the post-entry market price reflects the marginal cost of firms.[3] Our model is quite different from the industrial organization literature because it has a perfectly competitive post-entry market. Furthermore, the models by Spence (1977), Dixit (1980) and Schmalensee (1981) either assume a minimum entry capacity or a fixed set-up cost – independent of entry capacity. In contrast, our model allows for infinitesimal capacity investment by the entrant. Finally, our model incorporates uncertainty, a feature which has also been added to the above-mentioned models, by e.g. Perrakis and Warskett (1983) and Maskin (1999). Most importantly however, we assume *imperfect financial markets*. The investment externality in our model turns out to be a 'risk externality': the real investment changes the risk profile of future shocks.

Investment under uncertainty has been thoroughly studied in the real option framework (Dixit and Pindyck, 1994): firms should take into account the option value of an investment opportunity. By delaying the investment the firm learns more about the likely profitability of the project and might be able to avoid investments that are likely to be loss-making. Recently, Miao and Wang (2007) and Hugonnier and Morellec (2007) have extended the real option framework to the case of incomplete markets, using a utility-based approach. They study how market incompleteness affects the investment decisions of firms. Miao and Wang (2007) for example, find that – unlike in

[3]Reaching a sufficient number of entrants in order to satisfy this condition typically requires the absence of fixed set-up costs that would create barriers to entry.

standard real options analysis – an increase in project volatility can accelerate investment if the agent has a sufficiently strong precautionary savings motive. Although we use a similar utility-based model framework, our point of view is complementary in that we do not focus on how the entrant should make investment decisions, but rather study the social welfare implications of those decisions.

This paper is an extension and generalization of Willems and Morbee (2010), in which the effect of increasing market completeness on an entrant's investment decisions was examined numerically for the case of the electricity market. In the current paper we develop a general analytical model. In the next section, we first demonstrate the possibility of suboptimal entry when risk markets are incomplete. We will start from the traditional deterministic model, where entry is optimal, and subsequently include uncertainty and risk aversion, which may lead to suboptimal entry. Then, in section 7.3 we study the effect of increasing market completeness, i.e. increasing availability of instruments to trade risk between market participants. Section 7.4 summarizes our conclusions and briefly provides policy recommendations and areas for future research.

7.2 Suboptimal entry with incomplete markets

Building on the industrial organization literature, we first describe a deterministic version of our model, in which entry is always socially optimal, as there are no risk spillovers. In a second step, we demonstrate the possibility of suboptimal entry when uncertainty and risk aversion are introduced.

7.2.1 Traditional deterministic model

Following Mankiw and Whinston (1986), we model entry as a two-stage game. In the first stage, the *investment stage*, the entrant decides whether to enter the industry by investing in capacity. In stage two, the *production stage*, firms produce and sell a homogeneous product in a perfectly competitive market. $P(Q)$ denotes the inverse demand function, where Q is aggregate output, and assume $P'(Q) \leq 0, \forall Q$. Before any entry takes place, the industry marginal cost curve is given by $C'(Q)$, with $C''(Q) \geq 0, \forall Q$. In the absence of entry, the competitive market equilibrium is (p^*, Q^*) with $p^* = P(Q^*) = C'(Q^*)$. We consider an entrant who, in the first stage, has the possibility to invest in an infinitesimal amount of production capacity dq, at an investment cost of $k\,dq$. If the entrant decides to invest, she will have access to a production capacity dq with marginal production cost c in the second stage.[4]

[4]We assume an infinitesimal small entrant as we want to model the behavior of a competitive, price-taking entrant. An infinitesimal investor will affect market outcomes only marginally, which justifies the price-taking assumption. Note that in contrast with many entry models we do not assume that investment decisions are lumpy. Spence (1977), Dixit

Figure 7.3: Effect of entry on stage-two Marshallian aggregate surplus

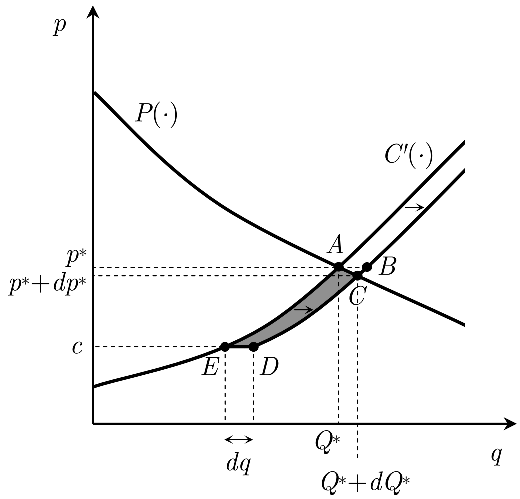


Figure 7.3 shows how the entry decision in stage one affects the outcome of the production stage, for the case in which $c \leq p^*$. Entry reduces the equilibrium price to $p^* + dp^*$, and increases the equilibrium quantity to $Q^* + dQ^*$. Entry increases stage-two Marshallian aggregate surplus by an amount corresponding to the shaded area $\widehat{ACDE}$. Since the area $\widehat{ABC}$ is only a second-order effect (it is approximately given by $\frac{1}{2}dp^* \cdot dq$), the surface area of $\widehat{ACDE}$ can be approximated by $\widehat{ABDE}$, which corresponds to $(p^* - c)dq$. Taking into account the investment cost $k\,dq$ incurred by the entrant in stage one, the net effect of entry on social welfare W is therefore:

$$dW = (p^* - c)dq - k\,dq \tag{7.1}$$

This amount dW corresponds exactly to the entrant's profit $d\pi$, hence the entrant's incentives are perfectly aligned with social interest: the entrant invests if and only if it is socially optimal to do so. This is the well-known textbook result about the social efficiency of free entry in a perfectly competitive market.

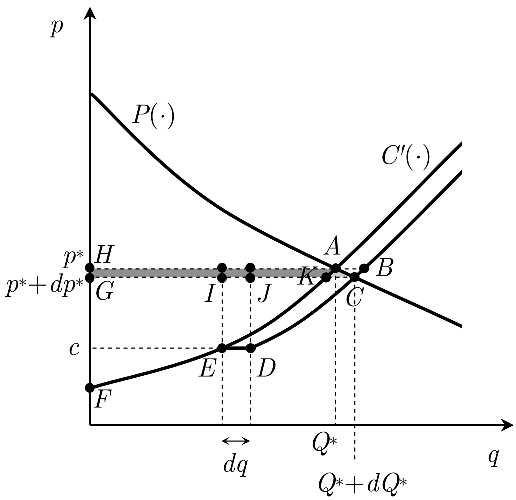
At this point it is useful to take a closer look at equation (7.1). In general, the change in social welfare caused by entry is given by:

$$dW = d\Pi + dCS + d\pi \tag{7.2}$$

where Π represents aggregate industry profits (producer surplus) of existing producers, CS represents aggregate consumer surplus and $d\pi$ the profit of the

(1980) and Schmalensee (1981) either assume a minimum entry capacity or a fixed set-up cost – independent of entry capacity. By assuming away lumpiness we eliminate a possible source of inefficiency in the market, and the results of our model become stronger.

Figure 7.4: Effect of entry on producer surplus of existing firms and on consumer surplus



infinitesimal entrant. Since we concluded above that $dW = d\pi$, we must have:

$$d\Pi = -dCS \tag{7.3}$$

Equation (7.3) is illustrated in figure 7.4. The investment in capacity dq causes a (negative) price change $dp^* = \frac{dp^*}{dq}dq$. As a result, CS increases by the area $\widehat{HACG}$, which in first order corresponds to $-Q^*dp^*$. The effect on Π is similar, but slightly more complicated to compute. In the absence of entry, the producer surplus Π of existing firms is given by $\widehat{HAF}$. In the event of entry, the producer surplus of existing firms changes to $\widehat{GIEF} + \widehat{JCD} = \widehat{GKF}$. Hence, $d\Pi = -\widehat{HAKG}$, which in first order corresponds to $Q^*dp^* = -dCS$.[5] Again, this result stems from $\widehat{ACK} = \widehat{ABC} = \frac{1}{2}dp^* \cdot dq$ being a second-order effect, which should be ignored in infinitesimal analysis. The fact that there is no net effect of entry on $\Pi + CS$ is due to the perfectly competitive nature of the production stage. In a non-competitive setting, entry would have an additional negative externality, due to 'business-stealing' from existing inframarginal capacity. In a perfectly competitive setting, existing firms produce up to the point where price equals marginal cost, hence the only existing capacity that is being displaced by business-stealing is the capacity at the margin, which does not have any net social value.

[5]Note that $d\Pi = Q^*dp^*$ is in fact nothing but Hotelling's lemma for the case of a one-good economy, while $dCS = -Q^*dp^*$ is Roy's identity for the case of a quasilinear utility function (as is implicitly assumed in our partial equilibrium setting).

The above reasoning is for the case in which $c \leq p^*$. The alternative case $c > p^*$ is trivial: since neither Π nor CS is affected by the entrant's investment, we obviously have $dW = d\pi$ and $d\Pi = -dCS = 0$ in this case.

7.2.2 Model including uncertainty and risk aversion

We will extend the deterministic model from section 7.2.1 to include the effects of uncertainty and risk aversion. Uncertainty is included by making the second stage stochastic: stage two takes place in a random *state-of-the-world* denoted ω, chosen stochastically among a range of possible states Ω. As a result, the variables Π, CS and $d\pi$, as well as all equilibrium prices and quantities, become random variables, which will be denoted using boldface.[6] The randomness may be caused by uncertainty in demand (as in Willems and Morbee, 2010), but may also be due to other factors, such as e.g. uncertainty in the prices of input factors, possible unforeseen outages of some of the production capacity, or regulatory uncertainty. Our reasoning is not limited to any of these sources of uncertainty.

The random nature of stage two requires additional assumptions about the social welfare function. As before we use a utilitarian social welfare function, i.e. the sum of the individual utilities of existing firms, consumers and entrant. As for the individual utility functions, we incorporate *risk aversion*, i.e. a preference for more certain outcomes over more uncertain outcomes for a given expected value of the outcome. For the sake of analytical convenience, we assume that the aggregate utility U_p of the existing producers is given by the well-known *mean-variance utility function:*

$$U_p = \mathrm{E}[\mathbf{\Pi}] - \frac{A_p}{2}\mathrm{Var}[\mathbf{\Pi}] \tag{7.4}$$

and, likewise, that the aggregate utility U_c of consumers is given by:

$$U_c = \mathrm{E}[\mathbf{CS}] - \frac{A_c}{2}\mathrm{Var}[\mathbf{CS}] \tag{7.5}$$

with the risk aversion parameters for producers and consumers $A_p, A_c \geq 0$. The expected-value and variance operators $\mathrm{E}[\cdot]$ and $\mathrm{Var}[\cdot]$ in equations (7.4) and (7.5) are computed on the sample space Ω. Social welfare is assumed to be given by:

$$W = U_p + U_c + U_e \tag{7.6}$$

in which U_e represents the utility of the entrant, for which we do not make functional-form assumptions.

Proposition 7.1. *When social welfare is given by equations (7.4), (7.5) and (7.6), the effect of an infinitesimal entrant with capacity dq, on social welfare, is given by:*

$$dW = dU_e + \mathrm{Cov}[A_p\mathbf{\Pi} - A_c\mathbf{CS}, \mathbf{x}]dq \tag{7.7}$$

[6] Later we will interpret random variables as vectors in $\#\Omega$-dimensional space.

where:

$$\mathbf{x} = -\mathbf{Q}^* \frac{d\mathbf{p}^*}{dq} \tag{7.8}$$

Proof. If the entrant decides to invest in capacity dq in stage one, this will have an effect on $\mathbf{\Pi}$ and $\mathbf{CS}$ in stage two. The effect may be different in each state-of-the-world ω. However, equation (7.3) will hold for each ω. Therefore, we can write the effect of entry on $\mathbf{\Pi}$ and $\mathbf{CS}$ as:

$$d\mathbf{\Pi} = -d\mathbf{CS} = -\mathbf{x}dq \tag{7.9}$$

with $\mathbf{x}$ as in equation (7.8). The effect on U_p is obtained by differentiation of equation (7.4):

$$\begin{aligned} dU_p &= d(\mathrm{E}[\mathbf{\Pi}] - \frac{A_p}{2}\mathrm{Var}[\mathbf{\Pi}]) \\ &= \mathrm{E}[d\mathbf{\Pi}] - \frac{A_p}{2}(\mathrm{Var}[\mathbf{\Pi} + d\mathbf{\Pi}] - \mathrm{Var}[\mathbf{\Pi}]) \\ &= \mathrm{E}[-\mathbf{x}dq] - \frac{A_p}{2}(\mathrm{Var}[\mathbf{\Pi} - \mathbf{x}dq] - \mathrm{Var}[\mathbf{\Pi}]) \\ &= -\mathrm{E}[\mathbf{x}]dq + A_p\mathrm{Cov}[\mathbf{\Pi}, \mathbf{x}]dq \end{aligned} \tag{7.10}$$

where we have used the fact that $\mathrm{Var}[\mathbf{x}dq] = \mathrm{Var}[\mathbf{x}](dq)^2$ can be ignored as a second-order term. Using an analogous reasoning, we obtain

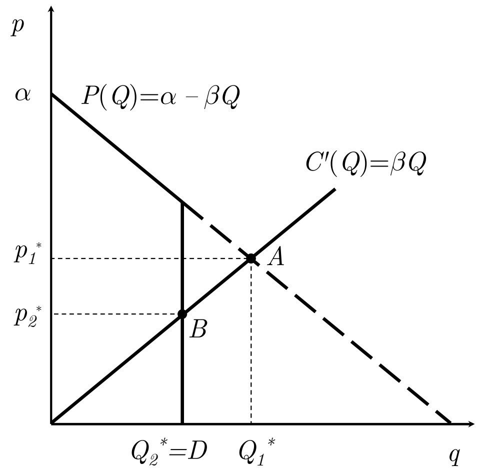
$$dU_c = \mathrm{E}[\mathbf{x}]dq - A_c\mathrm{Cov}[\mathbf{CS}, \mathbf{x}]dq \tag{7.11}$$

Putting equations (7.6), (7.10) and (7.11) together, we find equation (7.7). □

The first term dU_e in equation (7.7) is the effect of entry on the utility of the entrant himself. The second term in equation (7.7) is an *externality*: unlike dU_e, this effect of entry on social welfare is not fully internalized by the entrant, and hence is not included in his investment decision. Of particular interest are the cases in which the entrant would like to enter, $dU_e > 0$ but it would not be socially optimal, $dW < 0$, or vice versa. Such cases exist when dU_e and $\mathrm{Cov}[A_p\mathbf{\Pi} - A_c\mathbf{CS}, \mathbf{x}]$ have different signs and A_p and/or A_c are large enough. In such circumstances, a decision to enter (or refrain from entering) may be privately optimal, but socially detrimental. The externality is due to the fact that the investment in new capacity leads to a shift in market outcomes, which affects risk-sharing between existing producers and consumers. Hence, despite the perfectly competitive nature of stage two, the combination of uncertainty and risk-averse agents may lead to suboptimal entry. We will illustrate the potential inefficiency of free entry in the following example.

Example 7.2. Let us consider a sector with linear inverse demand $P(Q) = \alpha - \beta Q$ with $\alpha, \beta > 0$. The industry marginal cost curve is given by $C'(Q) = \bar{c}Q$, with $\bar{c}$ a constant. For analytical convenience we assume $\bar{c} = \beta$. In the absence

Figure 7.5: Graphical illustration of Example 7.2



of any intervention, the competitive equilibrium will be $(p_1^*, Q_1^*) = (\frac{\alpha}{2}, \frac{\alpha}{2\beta})$, shown as point A in Figure 7.5. We introduce uncertainty into this market by assuming that with probability ψ, government will intervene and forbid the lowest-value applications of the product.[7] The result of this intervention would be that demand becomes flat as soon as it reaches a certain level D. The part of the inverse demand curve to the right of $Q = D$ is clipped and becomes a vertical line at $Q = D$. Hence, with probability ψ the equilibrium is $(p_2^*, Q_2^*) = (\alpha - \beta D, D)$, which is shown as point B in Figure 7.5. Conversely, with probability $1 - \psi$ there is no government intervention and the equilibrium is at point A. No other source of uncertainty is assumed. Furthermore, we assume that $Q_2^* < Q_1^*$, i.e. $D < \frac{\alpha}{2\beta}$. Social welfare is assumed to be given by equations (7.4), (7.5) and (7.6), with $A_p = 0$ and $A_c > 0$. Hence, producers are risk-neutral while consumers are risk-averse.

Let us now consider an entrant who has access to two technologies: a 'peak' technology with marginal cost c_P such that $p_2^* < c_P < p_1^*$, and a 'base' technology with marginal cost c_B such that $0 < c_B < p_2^*$. The unit investment cost of the two technologies is k_P and k_B, respectively. Hence, the 'peak' technology will be activated only if there is no government intervention (probability $1 - \psi$), while the 'base' technology will be activated both in the case of government intervention and in the case of no government intervention. Obviously, a necessary condition for the 'peak' technology to be attractive, is that $k_P < k_B$. More than that, we will assume that both technologies

[7] Environmental concerns would be a typical reason for this kind of interventions. One could think, for example, of a ban on using tap water for filling up swimming pools, or – as it exists in some European countries – a fine for installing electrical heating in new houses.

yield equal, zero NPV for the investor. Assuming the entrant is risk-neutral like the other producers (in fact, the entrant could be one of the existing producers), we would then have $dU_e = 0$, which would make the entrant indifferent between investing and not investing in either technology. To make matters more interesting, we shall assume that k_B is infinitesimally smaller and k_P is infinitesimally larger, so that the entrant would have a marginal preference for investing in the 'base' technology and not investing in the 'peak' technology.

Since $dU_e = 0$, the social welfare impact of the investment is only the 'investment externality': $dW = dU_p + dU_c$. Using Proposition 7.1, or by directly computing U_p and U_c, one can demonstrate that the social welfare impact of an infinitesimal investment dq_B in the 'base' technology, for the case $\psi = \frac{1}{2}$, is given by:

$$\frac{dW}{dq_B} = -\frac{3A\beta^2}{8}\left(\frac{\alpha}{2\beta} - D\right)\left(D - \frac{\alpha}{6\beta}\right)\left(D - \frac{\alpha}{4\beta}\right) \tag{7.12}$$

while the welfare impact of an infinitesimal investment dq_P in the 'peak' technology, for the case $\psi = \frac{1}{2}$, is given by:

$$\frac{dW}{dq_P} = \frac{3A\beta^2}{8}\left(\frac{\alpha}{2\beta} - D\right)\left(D - \frac{\alpha}{6\beta}\right)\frac{\alpha}{4\beta} \tag{7.13}$$

Assuming that $D > \frac{\alpha}{4\beta}$, we find that $dW/dq_B < 0$ while $dW/dq_P > 0$. Hence, from a social welfare point of view, the entrant would overinvest in the 'base' technology, and underinvest in the 'peak' technology. The underlying cause is that the 'peak' technology takes costly risk out of the market, but the entrant is not rewarded for this. In this example, we have assumed that the uncertainty is due to unhedgeable political factors. In the next section, we will examine the case in which risk-sharing instruments are available.

7.3 Effects of increasing market completeness

As mentioned before, the demonstration of suboptimal entry in section 7.2.2 is related to imperfect risk-sharing between market participants. Indeed, the setting described above does not offer any instruments that would allow market participants to trade risk between them: markets are *incomplete*. In this section we will examine the case of increasingly complete markets.

7.3.1 Increasing market completeness without entry

Let us consider a case with n tradable financial instruments, such as forwards and options. Such instruments are fully represented by prices $F_i, i = 1, \ldots, n$ and their pay-offs $\mathbf{T}_i, i = 1, \ldots, n$, the latter being random variables because they depend on the state-of-the-world $\omega \in \Omega$ in stage two. Buying (selling) an instrument i means paying (receiving) a fixed price F_i in stage one, and receiving (paying) an uncertain pay-off $\mathbf{T}_i$ in stage two. Without loss of generality, we can assume that $\mathrm{E}[\mathbf{T}_i] = 0, \forall i$. To study the impact of the availability of these financial instruments on the behavior of producers and consumers, it is convenient to consider random variables as 'vectors'. Indeed, the space of zero-mean random variables (i.e. all functions $\mathbf{X} : \Omega \to \mathbb{R}$ with $\mathrm{E}[\boldsymbol{X}] = 0$) can be augmented with an inner product $\langle \mathbf{X}, \mathbf{Y} \rangle = \mathrm{E}[\mathbf{XY}] = \mathrm{Cov}[\mathbf{X}, \mathbf{Y}]$, to form a Hilbert space. The instrument pay-offs $\mathbf{T}_i, i = 1, \ldots, n$ span a subspace of this Hilbert space. Through orthogonal projection of the two zero-mean random variables $\mathbf{\Pi} - \mathrm{E}[\mathbf{\Pi}]$ and $\mathbf{CS} - \mathrm{E}[\mathbf{CS}]$ onto this subspace, we can uniquely rewrite $\mathbf{\Pi}$ and $\mathbf{CS}$ as:[8]

$$\begin{aligned} \mathbf{\Pi} &= \mathrm{E}[\mathbf{\Pi}] + \vec{\lambda}_p^T \tilde{\mathbf{T}} + \boldsymbol{\varepsilon}_p && (7.14) \\ \mathbf{CS} &= \mathrm{E}[\mathbf{CS}] + \vec{\lambda}_c^T \tilde{\mathbf{T}} + \boldsymbol{\varepsilon}_c && (7.15) \end{aligned}$$

with $\mathrm{E}[\boldsymbol{\varepsilon}_p] = \mathrm{E}[\boldsymbol{\varepsilon}_c] = 0$ and $\mathrm{Cov}[\mathbf{T}_i, \boldsymbol{\varepsilon}_p] = \mathrm{Cov}[\mathbf{T}_i, \boldsymbol{\varepsilon}_c] = 0, \forall i$. The arrow $\vec{\cdot}$ denotes an n-dimensional column matrix, and $\cdot^T$ denotes matrix transposition. Furthermore, we write $\vec{\mathbf{T}} = [\mathbf{T}_1 \ldots \mathbf{T}_n]^T$ and $\vec{F} = [F_1 \ldots F_n]^T$. Finally, note that $\boldsymbol{\varepsilon}_p$ and $\boldsymbol{\varepsilon}_p$ are stochastic, while $\vec{\lambda}_p$ and $\vec{\lambda}_c$ are deterministic.

The trade of financial instruments modifies producers' profits and consumer surplus. The resulting quantities are:

$$\begin{aligned} \tilde{\mathbf{\Pi}} &= \mathbf{\Pi} + \vec{k}_p^T (\vec{\mathbf{T}} - \vec{F}) && (7.16) \\ \tilde{\mathbf{CS}} &= \mathbf{CS} + \vec{k}_c^T (\vec{\mathbf{T}} - \vec{F}) && (7.17) \end{aligned}$$

with the column matrices $\vec{k}_p$ and $\vec{k}_c$ denoting the amount of each of the n instruments bought by producers and consumers, respectively. Negative amounts represent 'selling'. The resulting utility levels $\tilde{U}_p$ and $\tilde{U}_c$ are related to $\tilde{\mathbf{\Pi}}$ and $\tilde{\mathbf{CS}}$ in the same way as in equations (7.4) and (7.5).

Lemma 7.3. *In the absence of other players on the financial markets, the equilibrium quantities and prices of financial instruments bought and sold by producers and consumers are given by:*

$$\vec{k}_p = -\vec{k}_c = \frac{A_c \vec{\lambda}_c - A_p \vec{\lambda}_p}{A_c + A_p} \qquad (7.18)$$

and

$$\vec{F} = -\frac{A_c A_p}{A_c + A_p} \Sigma (\vec{\lambda}_c + \vec{\lambda}_p) \qquad (7.19)$$

with Σ the $n \times n$-dimensional covariance matrix of $\vec{\mathbf{T}}$.

[8]Uniqueness requires that the $\mathbf{T}_i, i = 1, \ldots, n$ not be linearly dependent. We assume here that this condition is fulfilled.

Proof. From equations (7.4), (7.14) and (7.16), we find that:

$$\tilde{U}_p = \mathrm{E}[\mathbf{\Pi}] - \vec{k}_p^T \vec{F} - \frac{A_p}{2}((\vec{\lambda}_p + \vec{k}_p)^T \Sigma (\vec{\lambda}_p + \vec{k}_p) + \mathrm{Var}[\boldsymbol{\varepsilon}_p]) \tag{7.20}$$

using the fact that $\mathrm{Cov}[\mathbf{T}_i, \boldsymbol{\varepsilon}_p] = 0, \forall i$. The gradient in $\vec{k}_p$, assuming price-taking behavior on the financial market, is easily derived as:

$$\vec{\nabla}_{k_p} \tilde{U}_p = -\vec{F} - A_p \Sigma (\vec{\lambda}_p + \vec{k}_p) \tag{7.21}$$

from which the first-order equilibrium condition for $\vec{k}_p$ can be determined:

$$\vec{k}_p = -\left(\frac{1}{A_p}\Sigma^{-1}\vec{F} + \vec{\lambda}_p\right) \tag{7.22}$$

A completely analogous condition can be derived for $\vec{k}_c$. In the absence of other players on the financial markets, we must have $\vec{k}_p + \vec{k}_c = \vec{0}$, from which we can derive equation (7.19). Substituting (7.19) into (7.22), we obtain (7.18). □

Equation (7.18) represents the optimal risk-sharing between producers and consumers, for the given set of available financial instruments.

Example 7.4. Assume the market is complete ($\boldsymbol{\varepsilon}_p = \boldsymbol{\varepsilon}_c = 0$) and $A_p = A_c = A$. Then $\vec{k}_p = \frac{\vec{\lambda}_p - \vec{\lambda}_c}{2}$, hence $\tilde{\mathbf{\Pi}} = \mathrm{E}[\mathbf{\Pi}] - \vec{k}_p^T \vec{F} + \left(\frac{\vec{\lambda}_p + \vec{\lambda}_c}{2}\right) \vec{\mathbf{T}}$, so that $\tilde{\mathbf{\Pi}}$ becomes identical to $\frac{\mathbf{\Pi} + \mathbf{CS}}{2}$, except for a non-stochastic component. The same holds for $\mathbf{CS}$, hence risk is perfectly distributed between producers and consumers: the only remaining risk is the sector risk $\mathbf{\Pi} + \mathbf{CS}$, which is shared equally between producers and consumers. The transition from an incomplete market (as in Section 7.2.2) to a complete market as in this example, increases social welfare from $\mathrm{E}[\mathbf{\Pi} + \mathbf{CS}] - \frac{A}{2}(\mathrm{Var}[\mathbf{\Pi}] + \mathrm{Var}[\mathbf{CS}])$ to $\mathrm{E}[\mathbf{\Pi} + \mathbf{CS}] - \frac{A}{4}(\mathrm{Var}[\mathbf{\Pi} + \mathbf{CS}])$.[9]

7.3.2 Production entry in an increasingly complete market

Let us now study the effect of an infinitesimal entrant in the case of an increasingly complete financial market. Analogous to equations (7.14) and (7.15), we can write $\mathbf{x}$ (defined as in Section 7.2.2) as:

$$\mathbf{x} = \mathrm{E}[\mathbf{x}] + \vec{\lambda}_x^T \vec{\mathbf{T}} + \boldsymbol{\varepsilon}_x \tag{7.23}$$

Lemma 7.5. *In first order, an entrant who only invests in physical capacity and does not enter the financial markets, does not change the prices of tradable*

[9] Note that market completion does not increase social welfare when the stochastic variations in $\mathbf{\Pi}$ and $\mathbf{CS}$ are identical. In that case, social welfare remains the same before and after the introduction of a complete market, because there are no gains to be made from trading risk.

financial instruments ($d\vec{F} = 0$), while the quantities of financial instruments traded change by an amount corresponding to the hedgeable part of the impact of the entrant ($d\vec{k}_p = -d\vec{k}_c = \vec{\lambda}_x dq$).

Proof. From equations (7.9), (7.14), (7.15), (7.23) and the uniqueness of orthogonal projection, one can infer that the effect on $\vec{\lambda}_p$ and $\vec{\lambda}_c$, of an infinitesimal entrant with capacity dq, is $d\vec{\lambda}_p = -d\vec{\lambda}_c = -\vec{\lambda}_x dq$. Lemma 7.5 then follows directly from Lemma 7.3. □

Proposition 7.6. *Assume the same conditions as in Proposition 7.1. When tradable financial instruments $\vec{\mathbf{T}}$ are available, the effect of an infinitesimal entrant with capacity dq, on social welfare, is given by:*

$$d\tilde{W} = dU_e + \mathrm{Cov}[A_p\boldsymbol{\varepsilon}_p - A_c\boldsymbol{\varepsilon}_c, \boldsymbol{\varepsilon}_x]dq \tag{7.24}$$

with $\boldsymbol{\varepsilon}_p$, $\boldsymbol{\varepsilon}_c$ and $\boldsymbol{\varepsilon}_x$ defined as in equations (7.14), (7.15) and (7.23).

Proof. Using reasoning analogous to the proof of Proposition 7.1, we find

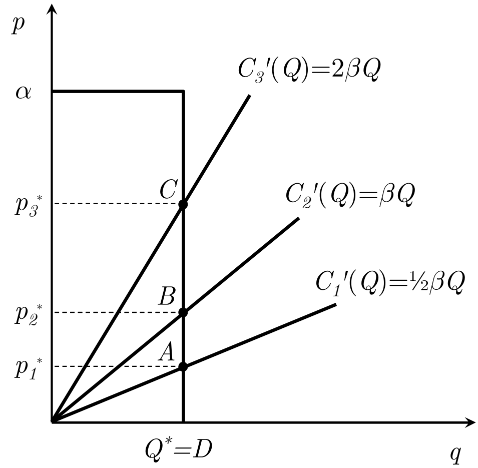
$$\begin{aligned}
d\tilde{U}_p &= d\left(\mathrm{E}[\mathbf{\Pi}] - \vec{k}_p^T\vec{F} - \frac{A_p}{2}((\vec{\lambda}_p + \vec{k}_p)^T\Sigma(\vec{\lambda}_p + \vec{k}_p) + \mathrm{Var}[\boldsymbol{\varepsilon}_p])\right) \\
&= \mathrm{E}[d\mathbf{\Pi}] - \vec{\lambda}_x^T\vec{F}dq - A_p(\vec{\lambda}_p + \vec{k}_p)^T\Sigma d(\vec{\lambda}_p + \vec{k}_p) \\
&\quad -\frac{A_p}{2}(\mathrm{Var}[\boldsymbol{\varepsilon}_p - \boldsymbol{\varepsilon}_x dq] - \mathrm{Var}[\boldsymbol{\varepsilon}_p]) \\
&= -\mathrm{E}[\mathbf{x}]dq - \vec{\lambda}_x^T\vec{F}dq - A_p(\vec{\lambda}_p + \vec{k}_p)^T\Sigma(-\vec{\lambda}_x dq + \vec{\lambda}_x dq) \\
&\quad -\frac{A_p}{2}(-2\mathrm{Cov}[\boldsymbol{\varepsilon}_p, \boldsymbol{\varepsilon}_x dq]) \\
&= -\mathrm{E}[\mathbf{x}]dq - \vec{\lambda}_x^T\vec{F}dq + A_p\mathrm{Cov}[\boldsymbol{\varepsilon}_p, \boldsymbol{\varepsilon}_x]dq
\end{aligned}$$

and analogous for $d\tilde{U}_c$. Putting both expressions together, we find equation (7.24). □

As before, the second term in equation (7.24) is an externality that may lead to over- or underinvestment. Proposition 7.6 clearly demonstrates the impact of increasing market completeness. As markets become more complete, the subspace spanned by the instruments $\vec{\mathbf{T}}$ approaches the complete space of random variables. As a result, $\mathrm{Var}[\boldsymbol{\varepsilon}_p] = \|\boldsymbol{\varepsilon}_p\|^2 \to 0$, and likewise for $\boldsymbol{\varepsilon}_c$, so that, in a complete market, the externality disappears and the entry decision is socially optimal.

Example 7.7. Let us consider a sector in which consumers have fixed inelastic demand D, with reservation price α. We assume that the industry marginal cost curve takes one of the following three forms: $C_1'(Q) = \frac{1}{2}\beta Q$, $C_2'(Q) = \beta Q$, or $C_3'(Q) = 2\beta Q$. In words: costs can take a reference value, or double the reference value, or half the reference value. As an example, the uncertainty in industry costs may be due to uncertainty in prices of input products, such as

Figure 7.6: Graphical illustration of Example 7.7



oil. To focus our thoughts, let us assume indeed that the industry is strongly dependent on oil and that all the above-described uncertainty in production costs is due to uncertainty about the oil price. The choice between the three cost curves is stochastic. We assume that the three states-of-the-world are equally likely. The sector assumptions are illustrated in Figure 7.6. Note that we assume $\alpha > 2\beta D$. The competitive equilibrium price in the absence of entry is $p_1^* = \frac{1}{2}\beta D$ (point A), $p_2^* = \beta D$ (point B), or $p_3^* = 2\beta D$ (point C), each with probability $\frac{1}{3}$. The top part of Table 7.1 summarizes the pay-offs $\mathbf{\Pi}$ and **CS** in each of the states-of-the-world. Social welfare is assumed to be given by equations (7.4), (7.5) and (7.6), with $A_p = A_c \equiv A$. Hence, consumers and producers are equally risk-averse.

Let us now consider an entrant who has access to three technologies: a 'peak' technology with marginal cost c_P such that $p_2^* < c_P < p_3^*$, a 'base' technology with marginal cost c_B such that $0 < c_B < p_1^*$, and a 'medium' technology with marginal cost c_M such that $p_1^* < c_M < p_2^*$. We assume that the costs c_P, c_B and c_B do not exhibit any uncertainty. In the story of our example: they are independent of the oil price. The unit investment cost of the three technologies is k_P, k_B and k_M, respectively. The 'peak' technology will be activated only in state ω_3, the 'medium' technology will be activated only in states ω_2 and ω_3, while the 'base' technology will be activated in all three states. As in Example 7.2, we assume that all three technologies yield equal, zero NPV for the investor. Again assuming the entrant is risk-neutral, we would then have $dU_e = 0$ for all technologies, which would make the entrant indifferent between investing and not investing in any of the technologies. We will assume that in this case the entrant does not invest. The bottom part of

Table 7.1: Example 7.7: Profits of existing producers and consumer surplus in the three states-of-the-world (top part of the table), and impact of the entrant (bottom part of the table)

State-of-the-world $\omega =$	ω_1	ω_2	ω_3
$C'(Q) =$	$\frac{1}{2}\beta Q$	βQ	$2\beta Q$
$\mathbf{\Pi} =$	$\frac{1}{4}\beta D^2$	$\frac{1}{2}\beta D^2$	βD^2
$\mathbf{CS} =$	$(\alpha - \frac{1}{2}\beta D)D$	$(\alpha - \beta D)D$	$(\alpha - 2\beta D)D$
$\frac{d\mathbf{CS}}{dq_B} = -\frac{d\mathbf{\Pi}}{dq_B} = \mathbf{x}_B =$	$\frac{1}{2}\beta D$	βD	$2\beta D$
$\frac{d\mathbf{CS}}{dq_M} = -\frac{d\mathbf{\Pi}}{dq_M} = \mathbf{x}_M =$	0	βD	$2\beta D$
$\frac{d\mathbf{CS}}{dq_P} = -\frac{d\mathbf{\Pi}}{dq_P} = \mathbf{x}_P =$	0	0	$2\beta D$

Table 7.2: Example 7.7: Pay-offs of the tradable financial instruments

State-of-the-world $\omega =$	ω_1	ω_2	ω_3
$\mathbf{T}_1 =$	-1	0	1
$\mathbf{T}_2 =$	-1	-1	2
$\mathbf{T}_3 =$	4	-5	1

Table 7.1 shows the impact (on profits of existing producers and on consumer surplus) of an entrant investing in an infinitesimal amount of 'base' capacity (dq_B), 'medium' capacity (dq_M), or 'peak' capacity (dq_P), respectively. Since $dU_e = 0$, the social welfare impact of the investment is only the 'investment externality': $d\tilde{W} = d\tilde{U}_p + d\tilde{U}_c$. Looking at Table 7.1 and considering Proposition 7.1, it is easy to see that $d\tilde{W} > 0$ for all three technologies, because they reduce the variability of profits of existing producers and consumer surplus. However, since the entrant does not invest, we have a case of underinvestment (insufficient entry) compared to the social optimum.

Now let us introduce tradable financial instruments. Since there are three states-of-the-world, the Hilbert space of zero-mean random variables is two-dimensional. Hence, two linearly independent instruments $\mathbf{T}_1$ and $\mathbf{T}_2$ are sufficient to make the market complete. Let us define the pay-offs of $\mathbf{T}_1$ and $\mathbf{T}_2$ as in Table 7.2. $\mathbf{T}_1$ could be considered as a 'future' contract on the oil price, while $\mathbf{T}_2$ could be considered as a 'call option' contract. The table also mentions $\mathbf{T}_3$, which will be considered later on. The availability of tradable

Table 7.3: Example 7.7: Social welfare impact of the entrant (for each of the three technologies) as a function of the available tradable instruments

	No instruments	Only $\mathbf{T}_1$	Only $\mathbf{T}_2$	Both $\mathbf{T}_1$ and $\mathbf{T}_2$
$\frac{d\tilde{W}}{dq_B} =$	$\frac{7}{12}$	$\frac{1}{48}$	$\frac{1}{16}$	0
$\frac{d\tilde{W}}{dq_M} =$	$\frac{3}{4}$	0	$\frac{1}{8}$	0
$\frac{d\tilde{W}}{dq_P} =$	$\frac{5}{6}$	$\frac{1}{12}$	0	0

Note: all values in this table need to be multiplied by $A\beta^2 D^3$.

financial instruments alters the risk-sharing between existing producers and consumers. As a result, the risk-reducing external benefits of an investment by the entrant may be less important. Using Proposition 7.6, we can compute the impact of an infinitesimal investment on social welfare, when an increasing number of tradable instruments are available. Table 7.3 provides an overview of the results, for each of the three technologies. The presence of either $\mathbf{T}_1$ or $\mathbf{T}_2$ reduces the positive externalities of entry. When both instruments are present, the market is complete, hence risk-sharing between producers and consumers is perfect and entry (or, in this example, lack thereof) is socially optimal. Finally, it is interesting to note that the externalities for some types of entry may become 0 even when the market is not yet fully complete. This is the case when an instrument is available with exactly the same risk profile as the impact of the entrant. For example, the presence of only $\mathbf{T}_1$ already makes the externality of the 'medium' technology disappear. The same holds for $\mathbf{T}_2$ and the 'peak' technology.

The observation in Table 7.3 that the investment externality goes down for each tradable instrument added, is however not general:

Corollary 7.8. *Adding a tradable instrument does not necessarily decrease the investment externality computed in Propositions 7.1 or 7.6.*

To see this, let us consider a market in which the instruments $\mathbf{T}_i, i = 1, \ldots, n$ are available. The investment externality per unit of investment dq according to Proposition 7.6 is given by:

$$\mathrm{Cov}[A_p\boldsymbol{\varepsilon}_p - A_c\boldsymbol{\varepsilon}_c, \boldsymbol{\varepsilon}_x] = \langle \boldsymbol{\varepsilon}_{pc}, \boldsymbol{\varepsilon}_x \rangle \tag{7.25}$$

with $\boldsymbol{\varepsilon}_{pc} = A_p\boldsymbol{\varepsilon}_p - A_c\boldsymbol{\varepsilon}_c$. Consider the addition of a new instrument $\mathbf{T}_{n+1}$. Let $\mathbf{T}'_{n+1}$ denote the component of $\mathbf{T}_{n+1}$ that is orthogonal to $\mathbf{T}_i, i = 1, \ldots, n$. We can now write $\boldsymbol{\varepsilon}_{pc} = \boldsymbol{\varepsilon}'_{pc} + a_{pc}\mathbf{T}'_{n+1}$ and $\boldsymbol{\varepsilon}_x = \boldsymbol{\varepsilon}'_x + a_x\mathbf{T}'_{n+1}$, with $\left\langle \boldsymbol{\varepsilon}'_{pc}, \mathbf{T}'_{n+1} \right\rangle = \left\langle \boldsymbol{\varepsilon}'_x, \mathbf{T}'_{n+1} \right\rangle = 0$. The new value of the investment externality is now given by

$\left\langle \boldsymbol{\varepsilon}'_{pc}, \boldsymbol{\varepsilon}'_x \right\rangle$. We find:

$$\begin{aligned}
\langle \boldsymbol{\varepsilon}_{pc}, \boldsymbol{\varepsilon}_x \rangle &= \left\langle \boldsymbol{\varepsilon}'_{pc} + a_{pc}\mathbf{T}'_{n+1}, \boldsymbol{\varepsilon}'_x + a_x \mathbf{T}'_{n+1} \right\rangle && (7.26)\\
&= \left\langle \boldsymbol{\varepsilon}'_{pc}, \boldsymbol{\varepsilon}'_x \right\rangle + a_x \left\langle \boldsymbol{\varepsilon}'_{pc}, \mathbf{T}'_{n+1} \right\rangle + a_{pc} \left\langle \mathbf{T}'_{n+1}, \boldsymbol{\varepsilon}'_x \right\rangle + && (7.27)\\
&\quad + a_{pc} a_x \left\langle \mathbf{T}'_{n+1}, \mathbf{T}'_{n+1} \right\rangle && (7.28)
\end{aligned}$$

hence:

$$\left\langle \boldsymbol{\varepsilon}'_{pc}, \boldsymbol{\varepsilon}'_x \right\rangle = \langle \boldsymbol{\varepsilon}_{pc}, \boldsymbol{\varepsilon}_x \rangle - a_{pc} a_x \left\| \mathbf{T}'_{n+1} \right\|^2 \tag{7.29}$$

Clearly, when $a_{pc} a_x < 0$ (i.e. when $\operatorname{sgn} \left\langle \boldsymbol{\varepsilon}_{pc}, \mathbf{T}'_{n+1} \right\rangle \neq \operatorname{sgn} \left\langle \boldsymbol{\varepsilon}_x, \mathbf{T}'_{n+1} \right\rangle$), the investment externality increases. If in addition, $\langle \boldsymbol{\varepsilon}_{pc}, \boldsymbol{\varepsilon}_x \rangle > 0$, then the investment externality increases also in absolute terms. By analogy, the same holds when no instruments are available yet and the instrument added is the first (i.e. $n = 0$). As mentioned before, however, when sufficiently many instruments are added so that the market becomes complete, the externality always tends to 0.

Example 7.9. *(Continuation of Example 7.7)* Consider the same set-up as in Example 7.7. Suppose that we do not introduce $\mathbf{T}_1$ and $\mathbf{T}_2$, but instead we introduce $\mathbf{T}_3$ (and only $\mathbf{T}_3$), an instrument with pay-offs shown in Table 7.2. In the story of the example, $\mathbf{T}_3$ can be considered as an asymmetric long straddle option on the oil price. The investment externality after introduction of $\mathbf{T}_3$ is shown in Table 7.4. The introduction of $\mathbf{T}_3$ increases the investment externality of entry in 'peak' technology. Hence, if only the 'peak' technology is available, the introduction of $\mathbf{T}_3$ increases the inefficiency in entry. To illustrate this point, suppose that instead of $dU_e = 0$, we have $dU_e = -\frac{71}{84} A\beta^2 D^3 dq_P$. Clearly, the entrant would not invest. When no tradable instruments are available, this would also be the socially optimal behavior, since $dW = (-\frac{71}{84} + \frac{5}{6}) A\beta^2 D^3 dq_P < 0$. Now suppose that the instrument $\mathbf{T}_3$ is available. The entrant obviously still would not invest. But in this case, this would not be socially optimal, since $d\tilde{W} = (-\frac{71}{84} + \frac{6}{7}) A\beta^2 D^3 dq_P > 0$.

7.3.3 Entry in financial markets

Until now, we have assumed that the entrant invests only in physical capacity and does not trade on the financial markets. Let us now consider an entrant on the financial markets. As before, the available financial instruments are $\mathbf{T}_i, i = 1, \ldots, n$. The pre-entry equilibrium on the financial markets is described by Lemma 7.3. Entry here means that the entrant invests in an infinitesimal amount of financial instruments $d\vec{k}_e$ in stage one, thereby causing a change $d\vec{F}$ in the prices $\vec{F}$ of financial instruments, and a change $d\vec{k}_p$ and $d\vec{k}_c$, respectively, in the quantities $\vec{k}_p$ and $\vec{k}_c$ of financial instruments bought by existing producers and consumers, respectively. In the absence of other players on the financial markets, we must have $d\vec{k}_e + d\vec{k}_p + d\vec{k}_c = 0$.

Table 7.4: Example 7.7 – continued: Social welfare impact of the entrant (for each of the three technologies) as a function of the available tradable instruments

	No instruments			Only $\mathbf{T}_3$		
$\frac{d\tilde{W}}{dq_B} =$	$\frac{7}{12}$	$=$	0.583	$\frac{4}{7}$	$=$	0.571
$\frac{d\tilde{W}}{dq_M} =$	$\frac{3}{4}$	$=$	0.750	$\frac{5}{7}$	$=$	0.714
$\frac{d\tilde{W}}{dq_P} =$	$\frac{5}{6}$	$=$	0.833	$\frac{6}{7}$	$=$	0.857

Note: all values in this table need to be multiplied by $A\beta^2 D^3$.

Lemma 7.10. *In response to a change $d\vec{F}$ in the price of financial instruments – caused by infinitesimal entry on the financial markets – the existing producers and consumers change their quantities of financial instruments bought, by:*

$$d\vec{k}_j = -\frac{1}{A_j}\Sigma^{-1} d\vec{F} \qquad j = p, c \tag{7.30}$$

Proof. The proof follows directly from differentiation of equation (7.22). □

Proposition 7.11. *Entry on the financial markets without production entry, does not have an externality on the existing producers and consumers:*

$$d(\tilde{U}_p + \tilde{U}_c) = 0 \tag{7.31}$$

Proof. Differentiation of equation (7.20) yields:

$$d\tilde{U}_p = -d\vec{k}_p{}^T \vec{F} - \vec{k}_p^T d\vec{F} - A_p(\vec{\lambda}_p + \vec{k}_p)^T \Sigma d\vec{k}_p \tag{7.32}$$

Using Lemma 7.10, we obtain:

$$\begin{aligned} d\tilde{U}_p &= \left(\frac{1}{A_p}\Sigma^{-1} d\vec{F}\right)^T \vec{F} - \vec{k}_p^T d\vec{F} \\ &\quad - A_p(\vec{\lambda}_p + \vec{k}_p)^T \Sigma \left(-\frac{1}{A_p}\right) \Sigma^{-1} d\vec{F} \qquad (7.33) \\ &= \frac{1}{A_p} d\vec{F}^T \Sigma^{-1} \vec{F} + \vec{\lambda}_p^T d\vec{F} \qquad (7.34) \end{aligned}$$

and a completely analogous expression for $d\tilde{U}_c$. Putting both together, we find:

$$d(\tilde{U}_p + \tilde{U}_c) = \left(\frac{1}{A_p} + \frac{1}{A_c}\right) d\vec{F}^T \Sigma^{-1} \vec{F} + d\vec{F}^T (\vec{\lambda}_p + \vec{\lambda}_c) \tag{7.35}$$

Substituting $\vec{F}$ from Lemma 7.3 into the last factor of the first term, we find that the first term and the second term cancel out, hence equation (7.31). □

Proposition 7.11 is equivalent to saying $d\tilde{W} = d\tilde{U}_e$. The entrant on the financial markets therefore 'sees' the full societal impact of its entry. Entry decisions in the financial market are therefore always optimal from a societal perspective.

7.4 Conclusions

In this paper we have developed a model of investment in a perfectly competitive industry. We have shown that a combination of risk aversion of existing players and incomplete financial markets, leads to a situation in which entrants' investment decisions in productive assets may be inefficient. In particular, we have demonstrated that there are situations in which new entrants overinvest in one technology and underinvest in another technology, compared to the socially optimal investment decisions. The underlying cause is that presence of the new productive assets changes the distribution of overall industry risk, which, if financial markets are incomplete, creates a risk externality for the firms already active in the market. The availability of an additional tradable financial instrument (without making the market complete) does not necessarily reduce the externality. When financial markets become complete however, the externality disappears. If the entrant invests in the financial market instead of in productive assets, there are no externalities, hence entry decisions in the financial market are always optimal from a societal perspective.

The result of the above is that the industry as a whole takes too much risk by investing too much in production activities with highly correlated risk profiles. Firms that could reduce the overall industry risk, do no enter often enough, while firms that increase overall industry risk, enter too often. Governments could attempt to reduce these inefficiencies by stimulating the creation of financial markets. More than that, in the absence of financial markets, the results could provide a ground for sector-specific regulation of investment decisions. Indeed, one could imagine a regulatory setting in which all project proposals need to be screened in advance by the regulator in order to assess the impact of the proposed investment on systemic risk. Approval would be given when project benefits weigh up against a possible negative risk spillover. A major obstacle to this approach, however, is that it may be very difficult for the regulator to adequately measure the risk. Finally, from the perspective of competition policy, the analysis of this paper shows that, in the absence of complete financial markets, an efficiency defense based on optimal risk-sharing may be a valid argument in vertical mergers.

Our model takes the number and types of tradable financial instruments as an exogenous input. Future work could endogenize the degree of market completeness, in order to study e.g. whether incumbent firms might have strategies to create market incompleteness as an entry barrier. Furthermore, the model assumes mean-variance utility, which allows for simple closed-form expressions of welfare impacts of entrants. Using numerical methods, one could study the effect of assuming a different structure for the utility functions. Finally, our model makes no assumptions about the risk behavior of the entrant. By making such assumptions, one could make an integrated study of the effect of market completeness on both the risk externality and the entrant's decision-making under (hedgeable) uncertainty.

Possibilities for Further Research

Chapter 2. The results of this paper are obtained using a partial equilibrium model of the market for long-term gas import contracts, with differentiated competition between one potentially unreliable 'dominant firm' (Russia) and a reliable 'competitive fringe' of other non-European import suppliers. Future research could examine the impact of the other suppliers becoming unreliable as well. Another possible extension is to turn our model into a repeated game. In such a game, δ could become endogenous as part of a mixed Russian strategy. Finally, the topic of this paper could be placed in a broader comparison of policy measures (import taxes, rationing, interruptible consumer contracts, etc.) that can be used to address gas import challenges.

Chapter 3. The model of this paper can be further improved by modelling the behavior of citizens in a more detailed way, thereby endogenizing the election probability or non-revolution probability. Furthermore, an extension to more than two periods or to continuous time would enable a more refined modelling of the complex trade-off between current and future earnings. Finally, the use of longer time series of resource taxation data would allow for more conclusive panel data estimates.

Chapter 4. Our stylized representation of the electricity market of a country with nuclear power could be further refined. In particular, it would be useful to enhance the integration of international imports in our model, as these may be quite important in the event of large underinvestment. Furthermore, since our analysis shows that multi-party negotiation of lifetime extension agreements and auctioning of new nuclear licenses seem to be the most attractive policies, further research on the details of such auctioning processes would be beneficial.

Chapter 5. The analysis is strongly dependent on the assumptions underlying the Shapley value. The allocation shown in this paper is therefore not necessarily the only possible allocation. Furthermore, supranational regulation and enforcement may be required in order to avoid renegotiation once the network is in place. An important area for future work is a more thorough understanding of the 'outside option' through better integration with the economic equilibrium models that generate the scenarios of CO_2 capture rates.

Chapter 6. The results of this paper are obtained in a numerical simulation for the electricity sector. The phenomenon can be generalized beyond the electricity sector and studied in an analytical model. This extension is pursued in Chapter 7.

Chapter 7. Our model takes the number and types of tradable financial instruments as an exogenous input. Future work could endogenize the degree of market completeness, in order to study e.g. whether incumbent firms might have strategies to create market incompleteness as an entry barrier. Furthermore, the model assumes mean-variance utility, which allows for simple closed-form expressions of welfare impacts of entrants. Using numerical methods, one could study the effect of assuming a different structure for the utility functions. Finally, our model makes no assumptions about the risk behavior of the entrant. By making such assumptions, one could make an integrated study of the effect of market completeness on both the risk externality and the entrant's decision-making under (hedgeable) uncertainty.

Bibliography

Acemoglu, D., 2005. Constitutions, Politics, and Economics: A Review Essay on Persson and Tabellini's *the Economic Effects of Constitutions*. Journal of Economic Literature 43(4): 1025-1048.

Acemoglu, D., S. Johnson and J. Robinson, 2001. The colonial origins of comparative development: An empirical investigation. American Economic Review 91(5): 1360-1401.

Acemoglu, D. and Robinson, J., 2006. Economic origins of dictatorship and democracy. Cambridge University Press, New York.

Aguilera, R., Eggert, R., Lagos, G. and Tilton, J., 2009. Depletion and the Future Availability of Petroleum Resources. The Energy Journal 30(1):141-174.

Aid R., Porchet, A., and N. Touzi, 2006. *Vertical Integration and Risk Management in Competitive Markets of Non-Storable Goods*. Centre de Mathmatiques Appliques – Ecole Polytechnique (29 September 2006).

Albrecht, J, Francois D, Schoors K, 2002. A Shapley decomposition of carbon emissions without residuals. Energ Policy 30(9): 727-736

Allaz, B. and J.L. Vila, 1993. "Cournot competition, forward markets and efficiency". Journal of Economic Theory 59: 1-16.

Allen, F. and D. Gale, 1994. Financial Innovation and Risk Sharing. Cambridge: MIT press.

Angevine, G., Brown, M. and Cervantes, M., 2009. *Fraser Institute Global Petroleum Survey 2010.* The Fraser Institute, Vancouver, BC.

Angevine, G. and Cameron, B., 2007. *Fraser Institute Global Petroleum Survey 2007.* The Fraser Institute, Vancouver, BC.

Angevine, G. and Cervantes, M., 2010. *Fraser Institute Global Petroleum Survey 2010.* The Fraser Institute, Vancouver, BC.

Angevine, G. and Cervantes, M., 2011. *Fraser Institute Global Petroleum Survey 2011.* The Fraser Institute, Vancouver, BC.

Angevine, G. and Thomson, G., 2008. *Fraser Institute Global Petroleum Survey 2008.* The Fraser Institute, Vancouver, BC.

Auty, R. M., 1993. *Sustaining Development in Mineral Economies: The Resource Curse Thesis*, London/New York:

Auty, R. M., 2001. *Resource Abundance and Economic Development.* Oxford: Oxford University Press.

Auty, R. M. and A. H. Gelb, 2001. Political Economy of Resource-Abundant States. In: Auty (2001). Routledge.

BAFA, 2006. `http://www.bafa.de/1/de/aufgaben/energie/`. Retrieved 21.08.06.

BAFA, 2009. *Monatliche Erdgasbilanz und Entwicklung der Grenzübergangspreise, ausgewählte Statistiken zur Entwicklung des deutschen Gasmarktes*, Bundesministerium für Wirtschaft und Technologie – Bundesamt für Wirtschaft und Ausfuhrkontrolle, `http://www.bmwi.de/BMWi/Navigation/Energie/Energiestatistiken/gasstatistiken.html`, last consulted in June 2009.

BAFA, 2011. Monatliche Erdgasbilanz und Entwicklung der Grenzübergangspreise, ausgewählte Statistiken zur Entwicklung des deutschen Gasmarktes. Bundesministerium für Wirtschaft und Technologie Bundesamt für Wirtschaft und Ausfuhrkontrolle, http://www.bmwi.de/BMWi/Navigation/ Energie/Energiestatistiken/gasstatistiken.html, last consulted in February 2011

Bain, J., 1949. A Note on Pricingin Monopoly and Oligopoly. American Economic Review 39(2):448-464.

Baldursson, F. and N.-H. von der Fehr, 2007. Vertical Integration and Long-Term Contracts in Risky Markets. Conference on "The Economics of Energy Markets", Toulouse (January 15-16, 2007).

Besley, T., 2006. *Principled Agents? The Political Economy of Good Government*, Oxford: Oxford University Press.

Bessembinder, H. and M.L. Lemmon, 2002. Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets. The Journal of Finance 57(3): 1347-1382.

Boots, M., F. Rijkers and B. Hobbs, 2004. Trading in the Downstream European Gas Market: A Successive Oligopoly Approach. The Energy Journal 25(3): 73-102.

Borenstein S., 2007. Customer risk from real-time retail electricity pricing: Bill volatility and hedgability. The Energy Journal 28(2): 111-129.

Borenstein S., Bushnell J., Knittel C.R., 1999. Market power in electricity markets: beyond concentration measures. The Energy Journal 20(4): 65-88.

Boucher, J., T. Hefting and Y. Smeers, 1987. Economic Analysis of Natural Gas Contracts, in Golombek et al. (1987).

Boyle, P. and T. Wang, 2001. Pricing of new securities in an incomplete market: The catch 22 of no-arbitrage pricing. Mathematical Finance 11: 267-284.

BP, 2006/2007/2008. *Statistical Review of World Energy June 2006/2007/2008*, London, BP p.l.c.

BP, 2011. *Statistical Review of World Energy June 2011*, London, BP p.l.c.

Broek, M vd, Ramirez A, Groenenberg H, Neele F, Viebahn P, Turkenburg W, Faaij A, 2010a. Feasibility of storing CO_2 in the Utsira formation as part of a long term Dutch CCS strategy: An evaluation based on a GIS/MARKAL toolbox. Int J Greenh Gas Con 4: 351-366

Broek, M vd, Brederode E, Ramirez A, Kramers L, van der Kuip M, Wildenborg T, Turkenburg W, Faaij A, 2010b. Designing a cost-effective CO_2 storage infrastructure using a GIS based linear optimization energy model. Environ Modell Softw 25(12): 1754-1768

Browning E., 1976. The Marginal Cost of Public Funds. The Journal of Political Economy 84(2): 283-298.

Burness, S., 1976. On the Taxation of Nonreplenishable Resources. Journal of Environmental Economics and Management 3:289-311.

Bushnell J., Mansur E.T., Saravia C., 2005. Market Structure and Competition: a cross-market analysis of U.S. electricity deregulation. CSEM WP 126, University of California Energy Institute.

Campbell, H. F. and Lindner, R. K., 1985. A Model of Mineral Exploration and Resource Taxation. The Economic Journal 95(377): 146-160.

Capros P, Mantzos L, Tasios N, De Vita A, Kouvaritakis N (2010) EU energy trends to 2030 - Update 2009. Publications Office of the European Union, Luxembourg

Cardell J. B., Hitt C.C., Hogan W. W., 1997. Market power and strategic interaction in electricity networks. Resource and Energy Economics 19: 109-137.

Carlton, D., and J. Perloff, 2000. *Modern Industrial Organization, Third Edition*, Reading, Massachusetts, Addison-Wesley.

Carr, P., H. Geman, and D.B. Madan, 2001. Pricing and hedging in incomplete markets. Journal of Financial Economics 62: 131-167.

Cass, D. and A. Citanna, 1998. Pareto improving financial innovation in incomplete markets. Economic Theory 11: 467-494.

Centrica, 1997-2001. Press releases: "Extraordinary General Meeting: British Gas Shareholders Approve Demerger" (12/2/1997), "British Gas signs up its first electricity customers" (1/12/1997), "Centrica delivers generation for growing electricity customer base" (29/5/2001), "Centrica acquires further electricity generation" (24/8/2001). Centrica Media Relations, Centrica corporate website (`http://www.centrica.co.uk`).

Coase, R., 1960. The Problem of Social Cost. Journal of Law and Economics 3(1): 1-44.

CO2Sense, 2010. A Carbon Capture and Storage network for Yorkshire and the Humber, Pre-Front End Engineering Study – Executive summary. CO2Sense Yorkshire, www.co2sense.org.uk, last consulted on 26 November 2010

Conrad, R. and Hool, B., 1981. Resource Taxation with Heterogeneous Quality and Endogenous Reserves. Journal of Public Economics 16:17-33.

Contreras, J. and Wu, F. (2000). A Kernel-Oriented Algorithm for Transmission Expansion Planning. IEEE Transactions On Power Systems 15(4): 1434-1440.

CREG – Commissie voor de Regulering van de Elektriciteit en het Gas / Commission de Régulation de l'Electricité et du Gaz, 2010a. Studie over de kostenstructuur van de elektriciteitsproductie door de nucleaire centrales in België. Studie (F)20100506-CDC-968. CREG, Brussels.

CREG – Commissie voor de Regulering van de Elektriciteit en het Gas / Commission de Régulation de l'Electricité et du Gaz, 2010b. Studie over het nucleair akkoord in Duitsland en de toepassing ervan op België. Studie (F)101014-CDC-999. CREG, Brussels.

CREG – Commissie voor de Regulering van de Elektriciteit en het Gas / Commission de Régulation de l'Electricité et du Gaz, 2011. Rente nuclaire : la CREG analyse les explications dElectrabel et confirme sa position. Press release, March 1 2011. CREG, Brussels.

Csercsik, D. and Koczy, L. (2011). Externalities in the games over electrical power transmission networks. IEHAS Discussion Papers.

Dahl, C., 1993. *A Survey of Energy Demand Elasticities in Support of the Development of the NEMS*, Paper prepared for United States Department of Energy, Contract De-AP01-93EI23499.

Dasgupta, P. and Heal, G., 1979. Economic Theory and Exhaustible Resources. Cambridge University Press, Cambridge, UK.

Dasgupta, P., Heal, G. and Stiglitz, J., 1980. The Taxation of Exhaustible Resources. NBER Working Paper No. 436, January 1980.

Davis, G., 1995. Learning to Love the Dutch Disease: Evidence from the Mineral Economies. World Development 23(10): 1765-1779.

Demsetz, H., 1968. Why regulate utilities? The Journal Of Law And Economics 11: 55-66.

De Tijd. Uitgeversbedrijf Tijd NV, Brussels.

Deneckere, R., and D. Kovenock, 1992. Price Leadership. The Review of Economic Studies 59: 143-162.

Dixit, A., 1980. The Role of Investment in Entry-Deterrence. The Economic Journal 90(357):95-106.

Dixit, A. and R. Pindyck, 1994. *Investment Under Uncertainty*, Princeton: PrincetonUniversity Press.

Downs, A., 1957. An Economic Theory of Political Action in a Democracy. Journal of Political Economy 65(2): 135-150.

Duffie, D. and R. Rahi, 1995. Financial Market Innovation and Security Design - an Introduction. Journal of Economic Theory 65: 1-42.

Duffie, D. and W. Shafer, 1985. Equilibruim in incomplete markets I: Basic model of generic existence. Journal of Mathematical Economics 65: 1-42.

Duffie, D. and W. Shafer, 1986. Equilibrium in incomplete markets II: Generic existence in stochastic economies. Journal of Mathematical Economics 15: 199-216.

E.ON, 2009. Financial Report 2009. E.ON AG, Düsseldorf.

Easterly, W., 2001. *Global Development Network Growth Database*. The World Bank, Washington D.C.

Elia, 2011. Corporate website `www.elia.be`. Elia System Operator NV, Brussels. Last consulted on July 12th 2011.

Elul, R., 1995. Welfare Effects of Financial Innovation in Incomplete Markets Economies With Several Consumption Goods. Journal of Economic Theory 65: 43-78.

Elul, R., 1999. Welfare-improving financial innovation with a single good. Economic Theory 13: 113-131.

E-PRTR (European Pollutant Release and Transfer Register), 2010. Database available from the European Environment Agency, Copenhagen, www.eea.europa.eu, update of 8 June 2010

EURELECTRIC – Union of the Electricity Industry, 2010. Power Choices, Pathways to Carbon-Neutral Electricity in Europe by 2050. Eurelectric, Brussels

European Commission, 2007. *An Energy Policy for Europe*, Communication to the European Council and the European Parliament.

European Commission, 2008. Energy Sources, Production Costs and Performance of Technologies for Power Generation, Heating and Transport. Commission Staff Working Document accompanying the Second Strategic Energy Review, An EU Energy Security and Solidarity Action Plan. SEC(2008) 2872. Brussels.

European Commission, 2010a. Energy 2020 - A strategy for competitive, sustainable and secure energy. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM(2010) 639 final, Brussels 10.11.2010

European Commission, 2010b. Energy infrastructure priorities for 2020 and beyond - A Blueprint for an integrated European energy network. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM(2010) 677 final, Brussels, 17.11.2010

European Commission, 2011. Energy Roadmap 2050, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. COM(2011) 885/2 final, Brussels, 19.12.2011.

Eurostat, 2011. Statistics downloaded from `epp.eurostat.ec.europa.eu`.

Evans, F., Zolezzi, J. and Rudnick, H., 2003. Cost assignment model for electrical transmission system expansion: an approach through the Kernel theory. IEEE Transactions on Power Systems 18: 625-632.

Eydeland, A. and H. Geman, 1998. Pricing power derivatives. Risk Magazine October 1998, 71-73.

Fabrizio, K., Rose, N. and Wolfram, C., 2007. Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on US Electric Generation Efficiency. American Economic Review 97(4): 1250-1277.

Financial Times. The Financial Times Ltd, London.

Freedom House, 2011. *Freedom in the World 2011*. Freedom House, Washington DC.

Fridolfsson, S. and Tangers, T., 2011. Nuclear Capacity Auctions. IFN Working Paper No. 892.

Gately, D. (1974). Sharing the Gains from Regional Cooperation: A Game Theoretic Application to Planning Investment in Electric Power. International Economic Review 15: 195-208.

Gaudet, G. and M. Moreaux, 1990. Price versus Quantity Rules in Dynamic Competition: The Case of Nonrenewable Natural Resources. International Economic Review 31(3): 639-650.

Golombek, R., M. Hoel and J. Vislie (eds.), 1987. *Natural Gas Markets and Contracts*, North Holland, Elsevier Science Publishers.

Golombek, R. and E. Gjelsvik, 1995. Effects of Liberalizing the Natural Gas Markets in Western Europe. The Energy Journal 16(1).

Golombek, R., E. Gjelsvik and K. Rosendahl, 1998. Increased Competition on the Supply Side of the Western European Natural Gas Market. The Energy Journal 19(3): 1-18.

Grais, W. and K. Zheng, 1996. Strategic Interdependence in European East-West Gas Trade: A Hierarchical Stackelberg Game Approach. The Energy Journal 17(3): 61-84.

Green, R., 2003. *Electricity contracts and retail competition.* University of Hull (December 2003).

Green, R., 2007. Carbon Tax or Carbon Permits: The Impact on Generators' Risks. Conference on "The Economics of Energy Markets", Toulouse (January 15-16, 2007).

Hall, R. and Jones, C., 1999. Why Do Some Countries Produce So Much More Output Per Worker Than Others? Quarterly Journal of Economics 114, 83-116.

Hart, O., 1975. On the optimality of equilibruim when the market structure is incomplete. Journal of Economic Theory 11: 418-835.

Hartley, P. and K. Medlock III, 2009. Potential Futures for Russian Natural Gas Exports. The Energy Journal 30(Special Issue): 73-95.

Heaps, T., 1985. The Taxation of Nonreplenishable Natural Resources Revisited. Journal of Environmental Economics and Management 12:14-27.

Heston, A., Summers, R. and Aten, B., 2011. *Penn World Table Version 7.0*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, May 2011.

Himona, I., 2005. *Oil & Gas: Nigerian Oil & Gas: Vital Statistics*, Morgan Stanley Report 10718962.

von Hirschhausen, C., B. Meinhart and F. Pavel, 2005. Transporting Russian Gas to Western Europe – A Simulation Analysis. The Energy Journal 26(2): 49-68.

Hobbs, B. and Kelly, K. (1992): Using game theory to analyze electric transmission pricing policies in the United States. European Journal of Operational Research 56: 154-171.

Hobbs B.F., Pang J.S., 2007. Nash-Cournot equilibria in electric power markets with piecewise linear demand functions and joint constraints. Operations Research 55(1): 113-127.

Holz F., C. von Hirschhausen and C. Kemfert, 2008. A strategic model of European gas supply (GASMOD). Energy Economics 30: 766-788.

Hoster, F., 1996. *Auswirkungen des europischen Binnenmarktes fr Energie auf die deutsche Elektrizittswirtschaft.* Munich: Oldenbourg.

Hotelling, H., 1929. Stability in Competition. The Economic Journal 39(153): 41-57.

Hotelling, H., 1931. The Economics of Exhaustible Resources. The Journal of Political Economy 39(2): 137-175.

Hubert, F. and S. Ikonnikova, 2004. *Holdup, Multilateral Bargaining, and Strategic Investment: The Eurasian Supply Chain for Natural Gas*, Discussion paper Humboldt University Berlin.

Hubert, F, Ikonnikova S, 2009. Investment Options and Bargaining Power the Eurasian Supply Chain for Natural Gas. MPRA Paper No. 17854, September 2009.

Hugonnier, J. andE. Morellec, 2007. Corporate control and real investment inincomplete markets. Journal of Economic Dynamics and Control 31: 1781–1800.

International Energy Agency, 2008. *World Energy Outlook 2008*, Paris, OECD/IEA.

Ikonnikova, S. and G. Zwart, 2009. Strengthening buyer power on the EU gas market: Import caps and supply diversification. Presentation at KULeuven Energy Institute Seminar, May 2009.

Ikonnikova, S. and G. Zwart, 2010. Reinforcing Buyer Power: Trade Quotas and Supply Diversification in the EU Natural Gas Market. TILEC Discussion Paper No. 2010-018.

IPCC, 2005. Special Report on Carbon Dioxide Capture and Storage. Prepared by Working Group III of the Intergovernmental Panel on Climate Change [Metz, B., O. Davidson, H.C. de Coninck, M. Loos, and L.A. Meyer (eds.)]. Cambridge University Press, Cambridge

Jackwerth, J. and M. Rubinstein, 1996. Recovering probability distributions from option prices. Journal of Finance 51(5): 1611-1632.

Janeba E., 2000. Tax Competition When Governments Lack Commitment: Excess Capacity as a Countervailing Threat. The American Economic Review 90(5): 1508-1519.

Jepma, C.J., 2001. Gaslevering onder druk: invloed van de Richtlijnen van de DTe op de Nederlandse gasstromen. April 2001. Groningen/Paterswolde, Stichting JIN.

Johnson N, Ogden J, 2010. Detailed spatial modeling of carbon capture and storage (CCS) infrastructure deployment in the southwestern United States. Proceedings of the 10th International Conference on Greenhouse Gas Control Technologies (GHGT-10), 19-23 September 2010, Amsterdam, The Netherlands. Energy Procedia, forthcoming

Johnston, D., 1994. Global petroleum systems compared by contractor take. Oil & Gas Journal Dec 12 1994: 47-50.

Johnston, D., 1999. Maturing planet, tougher terms change upstream landscape. Oil & Gas Journal Dec 13 1999: 16-20.

Johnston, D., 2003. *International Exploration Economics, Risk, and Contracts Analysis.* PennWell Books, Tulsa, Oklahoma.

Karl, T. L., 1997. *The Paradox of Plenty: Oil Booms and Petro-States*, Berkeley: University of California Press.

Kaufmann, D., Kraay, A. and Mastruzzi, M., 2010. *Governance Matters IX: Aggregate and Individual Governance Indicators 1996-2009*, Washington, D.C.: The World Bank.

Kleindorfer, P., Wu, D.-J. and Fernando, C. (2001). Strategic gaming in electric power markets. European Journal of Operational Research 130:156-168.

Krugman, P., 1987. The narrow moving band, the Dutch disease, and the competitive consequences of Mrs. Thatcher: notes on trade in the presence of dynamic scale economies. Journal of Development Economics 27: 41-55.

Lal, D., 1995. Why growth rates differ. The political economy of social capability in 21 developing countries. In: B. Koo and D. Perkins (eds.), *Social Capability and Long-Run Economic Growth*, Basingstoke: Macmillan: 288-309.

Lise W., B. Hobbs and F. van Oostvoorn, 2008. Natural gas corridors between the EU and its main suppliers: Simulation results with the dynamic GASTALE model. Energy Policy 36: 1890-1906.

Mankiw, G. and M. Whinston, 1986. Free Entry and Social Inefficiency. The RAND Journal of Economics 17(1): 48-58.

Marshall, M. and Jaggers, K., 2010. *Polity IV Project: Political Regime Characteristics and Transitions, 1800-2010*, Maryland: Center for Systemic Peace (CSP).

Maskin, E., 1999. Uncertaintyand entry deterrence. Economic Theory 14: 429-437.

Mathiesen, L., K. Roland and K. Thonstad, 1987. The European natural gas market: degrees of market power on the selling side, in Golombek et al. (1987).

Matthes F.C., 2010. Auswertungsaktualisierung des am 5. September 2010 ausgehandelten Modells für die Laufzeitverlängerung der deutschen Kernkraftwerke. Öko-Institut e.V., Berlin.

Mendelevitch R, Herold J, Oei P-Y, Tissen A, 2010. CO_2 Highways for Europe - Modeling a Carbon Capture, Transport and Storage Infrastructure for Europe. DIW Discussion Paper 1052, September 2010. Deutsches Institut für Wirtschaftsforschung, Berlin

Miaoa, J. and N. Wang, 2007.Investment, consumption, and hedging under incomplete markets. Journal of Financial Economics 86: 608-642.

Middleton R., Bielicki J., 2009. A scalable infrastructure model for carbon capture and storage: SimCCS. Energ Policy 37: 1052-1060

Milne, F. and H. M. Shefrin, 1987. Information and securities: A note on Pareto dominance and second best. Journal of Economic Theory 43: 314-328.

Modigliani, F., 1958. New Developments on the Oligopoly Front, *The Journal of Political Economy* 66(3): 215-232.

Morbee, J., Serpa J., Tzimas E., 2010. The Evolution of the Extent and the Investment Requirements of a Trans-European CO2 Transport Network. JRC Scientific and Technical Report Series, EUR 24565 EN. Publications Office of the European Union, Luxembourg

Morbee, J., Serpa, J. and Tzimas, E., 2012. Optimised Deployment of a European CO2 Transport Network. International Journal of Greenhouse Gas Control 7: 48-61.

Mulder, M. and G. Zwart, 2006. *Government involvement in liberalised gas markets*, CPB document No 110, Centraal Planbureau, Netherlands.

Mulligan, C., Gil, R. and Sala-i-Martin, X., 2004. Do Democracies Have Different Public Policies than Nondemocracies? The Journal of Economic Perspectives 18(1): 51-74.

Neuhoff, K. and C. von Hirschhausen, 2005. *Long-term contracts vs. short-term trade of natural gas – A European perspective*, Working paper University of Cambridge.

Newberry D.M., European Deregulation: Problems of liberalising the electricity industry, 2002, European Economic Review, 46, pp 919-927.

Niskanen, W., 1997. Autocratic, Democratic, and Optimal Government. Economic Inquiry 35:464-479.

Nordhaus, W., 1974. The 1974 Report of the President's Council of Economic Advisers: Energy in the Economic Report. American Economic Review 64(4): 558-565.

Norwegian Petroleum Directorate, 2011. The Norwegian Petroleum Directorate's FactPages. `http://factpages.npd.no/factpages/`. Last consulted on August 22nd 2011.

Oh, G., 1996. Some results in the CAPM with nontraded endowments. Management Science 42: 286-293.

OME – Observatoire Méditerranéen de l'Energie, 2002. *Assessment of internal and external gas supply options for the EU, evaluation of the supply costs of new natural gas supply projects to the EU and an investigation of related financial requirements and tools*, Study for the European Commission.

Oum, Y., Oren, S., and Deng, S., 2006. Hedging Quantity Risks with Standard Power Options in a Competitive Wholesale Electricity Market. March 24, 2006, Eleventh Annual POWER Research Conference on Electricity Regulation and Restructuring, UCEI, Berkeley.

Palsson, A.-M., 1996. Does the degree of relative risk aversion vary with household characteristics? Journal of Economic Psychology 17: 771-787.

Peltzman, S., 1976. Toward a more general theory of regulation. Journal of Law and Economics, 19:211-240.

Perrakis, S. and G.Warskett, 1983. Capacity and Entry Under Demand Uncertainty. The Review of Economic Studies 50(3): 495-511.

Persson, T. and Tabellini, G., 2000. Political Economics: Explaining Economic Policy. MIT Press, Cambridge, Massachusetts.

Persson, T. and Tabellini, G., 2003. The Economic Effects of Constitutions: What do the data say? MIT Press, Cambridge, Massachusetts.

Peterson, F., 1978. A Model of Mining and Exploring for Exhaustible Resources. Journal of Environmental Economics and Management 5: 236-251.

Pindyck, R., 1978. The Optimal Exploration and Production of Nonrenewable Resources. Journal of Political Economy 86(5):841-861.

Platts, 2010. PowerVision database, version 2010. Platts (The McGraw-Hill Companies), New York.

Various specialized and general press sources, 2006. Several articles from January to August 2006, and January 2009 from *International Herald Tribune* (Paris, The New York Times Company), *De Tijd* (Brussels, Uitgeversbedrijf Tijd NV).

Riordan, M., 1984. On Delegating Price Authority to a Regulated Firm. The RAND Journal of Economics 15(1): 108-115.

Robinson, J. A., Torvik, R. and Verdier, T., 2006. Political foundations of the resource curse. Journal of Development Economics 79(2): 447-468.

Rubinstein, M., 1994. Implied binomial trees. Journal of Finance 49: 771-818.

Sachs, J. and Warner, A., 1995. *Natural resource abundance and economic growth*, NBER Working Paper 5398.

Salanié B., 2003. The Economics of Taxation. The MIT Press, Cambridge, Massachusetts.

Schmalensee, R., 1981. Economiesof Scale and Barriers to Entry. The Journal of Political Economy: 89(6): 1228-1238.

Schmoker, J. and Klett, T., 2000. U.S. Geological Survey Assessment Model for Undiscovered Conventional Oil, Gas, and NGL Resources – The Seventh Approximation. In USGS (2000).

Schröter, J., 2004. *Auswirkungen des europäischen Emissionshandelssystems auf den Kraftwerkseinsatz in Deutschland*, Diploma Thesis, Berlin University of Technology, Institute of Power Engineering.

Shapley, LS, 1953. A Value for n-person Games. In Contributions to the Theory of Games, volume II, by H.W. Kuhn and A.W. Tucker, editors. Annals of Mathematical Studies 28: 307-317.

Siddiqui, A. S., 2003. Managing electricity reliability risk through the forward markets. Networks and Spatial Economics 3: 225-263

Singh, N. and X. Vives, 1984. Price and quantity competition in a differentiated duopoly. RAND Journal of Economics 15(4): 546-554.

Solow, R. and Wan, F., 1976. Extraction Costs in the Theory of Exhaustible Resources. The Bell Journal of Economics 7(2): 359-370.

Spence, M., 1977. Entry, capacity,investment and oligopolistic pricing. The Bell Journal of Economics 8(2): 534-544.

Staum, J., 2008. Incomplete Markets. In: Birge, J.R. and V. Linetsky, eds., *Handbooks in Operations Research and Management Science* 15: 511-563.

Stigler, G., 1971. The theory of economic regulation. Bell Journal of Economics and Management Science 2:3-21.

Stiglitz, J., 1976. Monopoly and the Rate of Extraction of Exhaustible Resources. The American Economic Review 66(4): 655-661.

Stock, J., Wright, J. and Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business & Economic Statistics 20:518-529.

Stutzer, M., 1996. A simple nonparametric approach to derivative security valuation. Journal of Finance 51: 1633-1652.

Sylos Labini, P., 1969. *Oligopoly and Technical Progress*, Cambridge: Harvard University Press.

The Economist. The Economist Newspaper Limited, London.

Tordo, S., 2006. *Fiscal Systems for Hydrocarbons: Design Issues*, World Bank Working Paper No. 123.

Torvik, R., 2002. Natural resources, rent seeking and welfare. Journal of Development Economics 67: 455-470.

Train, K., 1991. Optimal Regulation: The Economic Theory of Natural Monopoly Kenneth Train. Cambridge, The MIT Press.

Ulph, A. and G. Folie, 1980. Economic Implications of Stackelberg and Nash-Cournot Equilibria. Journal of Economics 40(3-4): 343-354.

USGS World Energy Assessment Team, 2000. U.S. Geological Survey World Petroleum Assessment 2000 – Description and Results. U.S. Geological Survey Digital Data Series - DDS-60. USGS Information Services, Denver, CO.

van der Ploeg, F., 2008. *Why do many resource-rich countries have negative genuine saving? Anticipation of better times, or rapacious rent-seeking*, CEPR Discussion Paper No. 7021.

van der Ploeg, F., 2011. Natural Resources: Curse or Blessing? Journal of Economic Literature 49(2): 366-420.

Van Meurs, A.P. and Seck, A., 1997. Government takes decline as nations diversify terms to attract investment. Oil & Gas Journal May 26 1997: 35-40.

van Wijnbergen, S., 1984. The 'Dutch disease': A Disease After All? The Economic Journal 94(373): 41-55.

Vangkilde-Pedersen, T, Kirk K, Smith N, Maurand N, Wojcicki A, Neele F, Hendriks C, LeNindre Y-M, Antonsen KL, 2009. D42 GeoCapacity Final Report. EU GeoCapacity project - Assessing European Capacity for Geological Storage of Carbon Dioxide. Available from www.geology.cz/geocapacity.

Ventosa M., Baillo A., Ramos A., Rivier M., 2005. Electricity market modeling trends. Energy Policy 33: 897-913.

VGE, 2006. *Jahrbuch der europäischen Energie- und Rohstoffwirtschaft 2006*. Essen: Verlag Glckauf GmbH.

von Weizsäcker, C., 1980.A Welfare Analysis of Barriers to Entry. The Bell Journal of Economics 11(2): 399-420.

Weber A., 1909. *Über den Standort der Industrie.*

Wei J.Y., Smeers Y., 1999. Spatial oligopolistic electricity models with cournot generators and regulated transmission prices. Operations Research 47(1): 102-112.

Willems, B., 2006. *Virtual divestitures, will they make a difference? Cournot competition, option markets and efficiency.* CSEM Working Papers, May 2006, Berkeley.

Willems, B. and J. Morbee, 2010. Market completeness: How options affect hedging and investmentsin the electricity sector. Energy Economics 32: 786–795.

Willems B., Rumiantseva I., Weigt H., 2009. Cournot versus supply functions: what does the data tell us? Energy Economics 31: 38-47.

Willen, P., 2005. New financial markets: Who gains and who loses. Economic Theory 26: 141-166.

Wolfram, C., 1999. Measuring duopoly power in the British electricity spot market. American Economic Review 89: 805-826.

The World Bank, 2011. *World Development Indicators & Global Development Finance.* May 2011, The World Bank, Washington D.C.

Zero Emissions Platform (ZEP), 2011. The Costs of CO_2 Storage. Brussels, European Technology Platform for Zero Emission Fossil Fuel Power Plants.

Propositions

1. Electrabel will charge and should charge the price of freely obtained CO_2 emission allowances to its customers.

2. There is no shortage of engineers and scientists.

3. Belgium's very high (75%) marginal tax rate on labor (including VAT) may improve happiness.

Doctoral dissertations from the Faculty of Business and Economics, see:
`http://www.kuleuven.ac.be/doctoraatsverdediging/archief.htm`.