Essays on market design and strategic behaviour in energy markets

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2017

vorgelegt
von

Diplom-Wirtschaftsinformatiker Stefan Lorenczik

aus

Oberhausen

Referent:	Prof. Dr. Felix Höffler
Koreferent:	PD Dr. Dietmar Lindenberger
Tag der Promotion:	13.11.2017

Acknowledgements

I want to thank my supervisor Prof. Dr. Felix Höffler for his support and feedback over the last years. I also thank PD Dr. Dietmar Lindenberger for co-referring my thesis. Furthermore, my gratitude goes to Prof. Dr. Johannes Münster for chairing the examination committee.

Financial support of the 'Deutsche Forschungsgesellschaft' (German Research Foundation) through research grant HO 5108/2-1 is gratefully acknowledged.

I thank my co-authors Raimund Malischek and Johannes Trüby for their endurance in successfully rewriting and improving our common paper; Timo Panke for the efficient and pleasant collaboration; Joachim Bertsch, Christian Growitsch and Stephan Nagl from whom I learned a lot when writing my first paper.

Special thanks go to my brother Christian for his helpful comments and remarks concerning my work. For their constant support I want to thank my parents. Gabriele Förderer deserves credit for her patience and support in many ways.

Finally, I want to thank all my former and present colleagues at ewi Energy Research & Scenarios and the EWI Institute of Energy Economics at the University of Cologne for creating an inspiring environment.

Stefan Lorenczik August 2017

Contents

Introduction

Overview

Today's energy markets face great challenges. The liberalisation of the European electricity markets started at the end of the 1990s. At about the same time the use of intermittent renewable energies, i. e., wind and solar power, started to increase significantly. This led to serious doubts concerning the security of supply for two reasons: First, concerns regarding the ability of the electricity system to cope with the variable feed-in of renewable energies. And second, doubts that the liberalised (energy only) market will send sufficient investment incentives for dispatchable capacity – especially since full load hours of conventional power plants have decreased driven by renewables. Until today, the debate concerning the appropriate design of electricity markets is controversial.

Although great efforts are undertaken to increase the share of renewable energies in all sectors, large parts of the global economy still depend on natural resources. The markets for these resources, e. g., coal, oil and gas, are characterised by a high concentration on the supply side. This poses the threat of an abuse of market power. Regulatory bodies try to address this thread by competition law but face various challenges: First of all, it is difficult to evaluate market behaviour and detect violations against competition law. Furthermore, the effects of market regulations are difficult to anticipate due to the frequently complex market structures involving demand and supply sides across the world.

This thesis deals with both areas – market design in electricity markets and strategic behaviour in natural ressource markets – and consists of the following four essays:

- Flexibility in Europe's power sector – A necessary good or an inevitable complement? (based on Bertsch et al., 2016)[1]
- On the interaction effects of market failure and capacity payments in interconnected electricity markets

[1] Available here: https://doi.org/10.1016/j.eneco.2014.10.022.

- Assessing market structures in resource markets – An empirical analysis of the market for metallurgical coal using various equilibrium models (based on Lorenczik and Panke, 2016)[2]
- Modeling strategic investment decisions in spatial markets (based on Lorenczik et al., 2017)[3]

The four essays are divided into two parts. In the first part, consisting of the first two essays, market design issues in electricity markets are discussed. Part two deals with strategic behaviour in spatial natural resource markets. In all papers that where developed with co-authors the researchers contributed in equal parts.

Part 1: Market design in electricity markets

The first part of this dissertation addresses market design issues in electricity markets. More precisely, it deals with the two earlier sketched concerns regarding security of supply: First, the concerns regarding the availability of sufficiently flexible power plants. And second, the consequences of insufficient investments signals in energy only markets in interconnected electricity markets.

The first essay (chapter 1) focuses on the flexibility requirements. The central finding is that the flexibility requirements of an electricity system with an increasing share of variable renewable energies – more specifically the ramping capabilities and balancing power provision – can be dealt with by the changing mix of conventional capacity that evolves from the changing residual load pattern. Additional targeted mechanisms do not appear to be required – provided that either an energy only market exists which provides sufficient investments signals or alternatively some kind of complementary capacity remuneration mechanism (CRM) is in place.

The second essay (chapter 2) is complementary to the first one – this time the absolute level of installed capacity is addressed. More precisely, the effects of insufficient price signals on welfare in single markets, but also in particular in interconnected markets are analysed. The main finding is that insufficient price signals in one market have only a very limited effect on neighbouring markets if those do not have price restrictions themselves. The negative domestic effects of distorted prices is amplified by trade. This renders capacity mechanisms to counter market failure the more necessary.

In the following both essays are outlined in greater detail.

[2] Available here: https://doi.org/10.1016/j.eneco.2016.07.007.
[3] Available here: https://doi.org/10.1016/j.ejor.2016.06.047.

Chapter 1: Flexibility in Europe's power sector – A necessary good or an inevitable complement?

The European Union has ambitious targets concerning the reduction of greenhouse gases. To achieve these targets the share of renewable energies is supposed to increase significantly in the next decades. Given a large deployment of wind and solar capacities, there are two major impacts on electricity systems: First, the electricity system has to be flexible enough to cope with the volatile renewables generation, i. e., ramp up or down demand or supply on short notice. Second, sufficient back-up capacities are needed for times of low intermittent renewables generation.

This paper analyses the question if the expected future increase of intermittent renewable energy capacities imposes special requirements on the market design. More specifically, is there a need for additional investment incentives for flexible system components? For this purpose, the development of the European electricity markets up to the year 2050 is simulated by deploying a linear investment and dispatch optimisation model. Flexibility requirements are implemented in the model via ramping constraints and requirements for the provision of balancing power.

The existing literature on the necessity for taking into account flexibility in the design of capacity mechanisms is rather scarce – the predominant focus in the debate concerning capacity mechanisms is on the totally installed capacity. Nonetheless, some papers identify the need for a market with products for flexibility. The analysis in this essay extends the previous literature by considering a broader range of flexibility options in the electricity system including demand side response. Additionally, the ambitious long-term renewable targets of the EU are taken into account which pose further challenges concerning flexibility on the electricity system.

The analyses show that the increase in intermittent renewables has a significant impact on the volatility of the residual load. Consequently, the demand for flexibility increases. However, least cost generation capacity investments result in a sufficiently flexible power plant fleet. Additional incentives for flexibility are not needed. The main trigger for investing in flexible resources are the achievable full load hours and the need for backup capacity. Due to a steeper residual load curve more power plants with low investment but high variable costs are included in the cost-efficient technology mix. Those technologies, e. g., gas-fired power plants, provide flexibility as a by-product. Thus, flexibility never poses a challenge in a cost-minimal capacity mix.

Chapter 2: On the interaction effects of market failure and capacity payments in interconnected electricity markets

The second essay of this dissertation adds to the ongoing discussion concerning capacity remuneration mechanisms (CRMs) in electricity markets. It builds on the prevailing view that energy only markets do not provide sufficient investments incentives as wholesale electricity markets are characterised by certain market failures (and regulatory inefficiencies) that render some form of external investment incentive necessary.

This paper focuses on the interaction effects of market failures in adjacent markets. For reference, we start by analysing the equilibrium market outcome assuming that no market failure and no capacity mechanism is in place. Subsequently, at first a price cap representing market failure and then an adjacent market is introduced. The analyses are complemented by the introduction of capacity payments. The comparison of the effects in isolated as well as in connected markets enables the assessment of cross-border effects as well as incentives for free-riding.

The discussion in the literature mainly focuses on the necessity of CRMs and discusses the properties of market design options on a national basis. Cross-border effects are frequently neglected. The essay adds to the literature by paying special attention to the influence of trade on the effects of market failure. In particular, it deals with the following questions: What effect does an insufficient level of capacity in one market have on neighbouring markets? Do markets with insufficient domestic capacity incentives benefit from CRMs in neighbouring markets? And what effects do exaggerated efforts to increase domestic capacity have on interconnected markets?

The results show that the negative implications of price caps in energy only markets worsen in interconnected markets: Installed capacity and ultimately welfare decrease to a larger extent. In addition, capacity payments are less efficient in countering these effects than in isolated markets. Price caps do not appear to have a (significantly) harmful effect on adjacent markets. But as domestic market failures have only little effect on neighbouring markets, so do capacity mechanisms: Capacity payments in one market do not appear to significantly support neighbouring markets and thus provide no incentive for free-riding.

Part 2: Strategic behaviour in spacial natural resource markets

Strategic behaviour and the exertion of market power have always been a matter of concern in energy markets, especially in natural resource markets. The exertion of market power can result in deadweight losses – regulatory bodies try to address this by market regulations aiming for a welfare maximising market outcome. The first problem is to detect collusive behaviour as available data is frequently limited. The second question is how regulatory decisions may influence the market outcome. This is especially relevant in resource markets as demand and supply regions are usually located in different parts of the world and therefore subject to conflicting interests.

The third essay deals with the question how collusive behaviour and underlying market structures can be detected. The most accessible data in most cases is historic market data – which also this essay relies on in the analysis. Various mathematical models are developed and applied to the metallurgical coal market. Thereby, two new market structures are identified which appear at least as likely as the cases that were previously considered in the literature.

The fourth essay complements and expands the previous analysis in two aspects: First, the focus shifts from the analysis of different varieties of oligopolies to the analysis of alternative market designs. More specifically, the effects of a switch from trade based on short-term contracts to long-term contracts is analysed. Second, investment decisions in production capacity are included in the market models. Again, the models are applied to the metallurgical coal market.

In the following both essays of part 2 are outlined in greater detail.

Chapter 3: Assessing market structures in resource markets –
An empirical analysis of the market for metallurgical coal using various equilibrium models

Resource markets are frequently characterised by a high concentration on the supply side and a low demand elasticity. This raises concerns that market results may be the product of explicit collusion between producers. But the actual market structure is usually unknown. Common models used to investigate the market structure try to replicate the market outcomes by applying economic models representing competitive markets, strategic Cournot competition and Stackelberg structures (usually limited to one leader).

For this essay three mathematical models including the mentioned Cournot and Stackelberg market representations were developed. Additionally, a multi-leader multi-follower market model was included in the analysis. The models are used to simulate four spatial market set-ups with varying market conduct of the individual players. The models are applied to the international market for metallurgical coal using input data for the years 2008 until 2010. Using several statistical measures the most likely market structures are identified.

The essay contributes to the literature on applied industrial organisation and, more specifically, the analysis of the international market for metallurgical coal. Previous studies are expanded by applying an Equilibrium Problem with Equilibrium Constraints (EPEC), a mathematical programme used to model multi-leader-follower settings, to a spatial market, i. e., a market with multiple, geographically disperse supply and demand nodes. This way, the essay adds to the literature by extending the scope of possible market structures under scrutiny.

The essay demonstrates the multiplicity of underlying market structures that result in the observed outcomes concerning trade flows and market prices. By analysing additional data, more distinctive conclusions were drawn. This demonstrated the need for comprehensive market analyses in order to achieve reliable conclusions. Especially omitting specific market configurations – unknowingly due to a lack of market insights or knowingly due to analytical restrictions – might result in premature and false assessments.

Chapter 4: Modeling strategic investment decisions in spatial markets

In oligopolistic markets production capacities are often a key factor for the strategic interaction between oligopolists. This essay expands the scope ot the previous essay by adding an investment phase to the models. Additionally, the focus shifts from various oligopoly set-ups to the analysis of different market designs. It thereby takes into account the recent steps taken towards a spot market based trade instead of long-term contracts.

Different market structures and designs influence oligopolistic capacity investments and thereby affect supply, prices and rents. The models used in the analysis comprise an investment stage and a supply stage in which players compete in quantities. Three models are compared: A perfect competition and two Cournot models. In the Cournot models, the product is either traded through long-term contracts or on spot markets. The models are applied to the international metallurgical coal market.

The essay adds to the literature by explicitly taking into account the separation between long-term investment and short-term production decisions. This multi-stage market representation is new to the analysis of spatial resource markets. It takes into account a market design with spot market based trade and enables the comparison of spot market based trade and long-term contract. In contrast to the prevalent literature, the implicit assumption of simultaneous investment and production decisions is lifted. Applied to the metallurgical coal market, possible consequences of the ongoing regime switch from long-term contracts to a more spot market based trade are analysed.

The essay demonstrates the importance of an appropriate market representation: The frequently used one-stage approach has more convenient mathematical properties compared to a bi-level approach and can therefore be solved more easily. But, depending on the actual market structure, the results might be misleading. In the application at hand

– although the total welfare is only slightly affected – the distribution between producer and consumer surplus differs substantially. The current shift from long-term contracts to spot market trade benefits consumers, to the disadvantage of the production companies.

Numerical modeling: Limitations and further research

Although addressing different research questions and dealing with separate markets, in all essays numerical approaches are used to solve the economic models. This illustrates the wide range of topics and the flexibility of today's energy market models. None of the analyses would have been possible twenty years ago, either due to limited computational power (in case of the linear optimisation problems), a lack of readily available solvers (in case of MPECs) or any general solution strategies for reasonable sized problems at all (in case of EPECs).

The advancements in solving economic models by using numerical methods – rather than solving them analytically – have made them increasingly interesting to economists. But even given these improvements, different areas for improvements remain. They can broadly be categorised into three groups: First, the dependency on accurate input data. Second, the ability to represent real world systems. And third, the space for misinterpretations. Some limitations will be outlined in the following – more certainly exist.

In contrast to closed-form solutions, numerical model results provide less general insights. But real world models can hardly be solved analytically given the complex nature of many economic problems. Thus, numerical approaches have to be applied. As these approaches rely on specific data, the quality of the model results hinges on the accuracy (and availability) of the input data. This is especially true for data that cannot be observed directly but has to be estimated. In this dissertation, this was in particular relevant for the analysis of market conduct in the metallurgical coal market (chapter 3). As the aim of the analysis was to replicate historic market outcomes, accurate input data was crucial. Although having an extensive dataset that has proven itself to provide reasonable model results in previous analyses, key parameters remained uncertain. This was first of all the case for the elasticity of demand – especially as the literature provides arguments for a wide range of realisations. We addressed this uncertainty by using a variety of different values. As this did not lead to conflicting conclusions, additional confidence in the findings was built. The results did not allow for a definite conclusion concerning the market structure. Additional analysis is needed, but hampered by a lack of available data. Given that economic research provides valuable insights, a stronger engagement for more transparency by regulatory authorities is desirable.

Even more difficult than a retrospective analysis is the task to model a scenario for a future development. Although economists (usually) do not claim to make predictions of the future, scenarios are meant to provide guidance by quantifying possible developments given alternative circumstances. It is vitally important to design consistent scenarios in order to get meaningful results. This frequently confronts economists trying to model real world scenarios with two contradicting objectives: First, paying attention to a detailed model representation of particular aspects. And second, taking into account interaction effects with adjacent markets. The first objective usually requires a partial equilibrium model in which certain variables are treated as being exogenous and therefore fix. The second objective demands for general equilibrium models which usually lack a detailed representation of individual aspects. More work is needed in the area of combining both approaches in order to identify consistent scenarios, interactions of parameters and to gain a better understanding of model results. A connection of General Equilibrium Models (CGE) with detailed partial energy market models promises additional insight in future research.

Despite the recent progress in computational power, today's economic models have some (common) limitations which have to be kept in mind in order to draw meaningful conclusions. The causes for most limitations and simplifications can be categorised into two groups. They are either chosen to limit the computational burden in terms of calculation speed or simplified models due to the absence of ready to use solution techniques. This is not only true for the more sophisticated MPEC and EPEC models used in this dissertation (chapters 2, 3 and 4) but also for relatively – in terms of available and ready to use solvers – simple linear models (chapter 1). Two of the most prominent limitations even in linear models are the assumptions of perfect foresight and perfect information (which also applies to all models in this dissertation). The issue of perfect foresight has frequently been addressed – especially in linear models – by using stochastic modelling. But still a trade-off has to be made between modelling accuracy, e. g., concerning the technical properties of power plants, and relieving the perfect foresight assumption. Perfect information has very rarely been bypassed in numerical models but is – depending on the specific set-up – certainly a strong assumption.

When it comes to more sophisticated models with respect to the market behaviour of individual players, the sacrifices that have to be made become bigger. Simple Cournot competition models for a medium sized number of players, time steps and markets that are formulated as mixed linear complementarity problems (MLCPs) can reliably be solved with standard tools. If the question at hand requires a bi-level representation of market interaction (e. g., in Stackelberg games as in chapter 3 or with separate investment and dispatch decisions as in chapter 4), even solving medium sized problems can be challenging. If only one optimisation problem has to be solved at the first level the model constitutes a so called Mathematical Programs with Equilibrium Constraints (MPECs). As the feasible region is not necessarily convex, usually non-linear solvers have to be applied. In most cases, only local optima can be guaranteed. Alternatively, mixed-integer approaches are common in the literature which approximate the optimal solution (as has been done in this dissertation). The most challenging models to solve are multi-level models with each stage requiring an equilibrium outcome. Although occasionally examples exist in the literature, reasonable sized models are usually restricted to two stages. For equilibrium problems with equilibrium constraints (EPECs) no easily accessible solution approach yet exists. In addition, there may not exist an equilibrium solution or there might as well be multiple solutions. In general, it cannot be determined how many solutions exist. Early approaches try to systematically search for different equilibria, but further research is needed to increase their reliability.

Even if the challenging tasks of obtaining reliable data as well as formulating and solving the appropriate model have been completed, the interpretation of model results may not be straightforward. At first, economic models that are solved numerically are frequently very complex and take into account a large number of influencing factors. Due to this complexity, the disentangling of cause and effect and the deduction of universal conclusions is challenging. In addition, and especially in complex and more realistic multi-stage models, multiple solutions might exist that differ significantly from each other. Additional research is needed concerning equilibrium selection mechanisms in the context of numerical models.

Numerical models tend to seduce their users into possibly misleading conclusions – especially if the results appear to match expectations or historical data. In the dissertation at hand this can be observed in the analysis of the metallurgical coal market (see chapter 3): The market structure allows several assumptions concerning the market conduct of producers. Previous analyses concluded that – based on the numerical models that represented some of the alternatives – some set-up are (more) likely (than others). By adding further (and more complex) set-ups to the potpourri of economic models, additional likely cases were identified – probably more exist. This highlights the necessity to explore and expand the range of economic models that can reliably be solved in order not to miss reasonable conclusions. Additionally, this illustrates the bias towards models (and therefore also conclusions) that are more easy accessible. This is especially true for investment models as simplifying single-level models are still predominant in real world applications – which can significantly influence results and conclusions (see chapter 4).

In summary, the improvements in the field of numeric models do not come without stumbling blocks. Economists need to keep an open mind and should constantly seek to improve and question their models and model results.

Part I

Market design in electricity markets

Flexibility in Europe's power sector – A necessary good or an inevitable complement?

By 2050, the European Union aims to reduce greenhouse gases by more than 80 %. The EU member states have therefore declared to strongly increase the share of renewable energy sources (RES-E) in the next decades. Given a large deployment of wind and solar capacities, there are two major impacts on electricity systems: First, the electricity system must be flexible enough to cope with the volatile RES-E generation, i. e., ramp up supply or ramp down demand on short notice. Second, sufficient back-up capacities are needed during times with low feed-in from wind and solar capacities.

This paper analyzes whether there is a need for additional incentive mechanisms for flexibility in electricity markets with a high share of renewables. For this purpose, we simulate the development of the European electricity markets up to the year 2050 using a linear investment and dispatch optimization model. Flexibility requirements are implemented in the model via ramping constraints and provision of balancing power. We find that an increase in fluctuating renewables has a tremendous impact on the volatility of the residual load and consequently on the flexibility requirements. However, any market design that incentivizes investments in least (total system) cost generation investment does not need additional incentives for flexibility. The main trigger for investing in flexible resources are the achievable full load hours and the need for backup capacity. In a competitive market, the cost-efficient technologies that are most likely to be installed, i. e., gas-fired power plants or flexible CCS plants, provide flexibility as a by-product. Under the condition of system adequacy, flexibility never poses a challenge in a cost-minimal capacity mix. Therefore, any market design incentivizing investments in efficient generation thus provides flexibility as an inevitable complement.

1.1 Introduction

By 2050, the European Union aims to reduce greenhouse gases by more than 80 %. The EU member states have therefore declared to strongly increase the share of renewable energy sources (RES-E) in the next decades. The vast majority of renewable energy is expected to come from wind and photovoltaics (PV). These sources, however, depend on local weather conditions, leading to an increase in stochastic electricity generation. Given a large deployment of wind and PV capacities, weather uncertainty results in two major impacts on electricity systems: First, the capacity mix must be flexible enough to cope with the volatile RES-E generation, i. e., ramp up supply or ramp down demand on short notice. Second, sufficient back-up capacities are needed to provide secure supply during times with low feed-in from wind and solar capacities. Otherwise, sharp decreases or increases in renewable production may lead to price spikes on the wholesale market and, if supply and demand do not meet, to potential black-outs. The provision of back-up capacity has been intensely discussed in the literature in recent years (for instance Cramton and Stoft, 2008, Joskow, 2008). Concerning flexibility, the discussion is rather new and previous literature is scarce. Lamadrid et al. (2011), an exception, argue that as volatility increases, additional incentives to invest in flexible resources should be implemented in market design. Meanwhile, the Californian system operator (CAISO) has already started to implement ramping products in market design to ensure flexibility (Xu and Threteway, 2012).

This paper analyzes whether there is a need for additional incentive mechanisms for flexibility in electricity markets with a high share of renewables.[4] One challenge of analyzing the role of flexibility in electricity markets is accounting for the possible contributions of all parts of an electricity system. First, the supply side is able to complement volatile RES-E generation with highly flexible gas-fired power plants or upcoming technologies such as power plants with a detachable carbon capture and storage unit. Second, the demand side can contribute flexibility by improving demand side management. Third, storages can restrain the volatility of the residual load for both the demand and supply side. Therefore, an integrated analysis of all flexibility possibilities is needed to answer the question of how an electricity system can adapt to an increasing share of renewables. From that, one can deduce whether flexibility requirements necessitate a special market design.

For this purpose, we simulate the development of the European electricity markets up to the year 2050 using a linear investment and dispatch optimization model accounting for all mentioned flexibility options. We assume investments in renewable energies

[4] The discussion concerning the necessity of capacity mechanisms is beyond the scope of this paper.

lead to an 80 % renewable share of total electricity generation in Europe in 2050. The model determines the cost-efficient capacity mix, ensuring adequate capacity and fulfillment of flexibility requirements.[5] These requirements result from load variation and the provision of balancing power, which are necessary due to the stochastic in-feed from renewable generation. Flexibility of power plants, however, is restricted by minimum load and start-up constraints. Due to the importance of flexibility provision on short notice, the calculations are supplemented by using a dispatch model for 8760 hours for selected years (2020, 2030, 2040 and 2050). CO_2 emission costs may have effects on installed capacity (or generation) of base or peak load and storage capacities. Thus, impacts on the optimal capacity mix, flexible resources and flexibility provision are further analyzed by calculating an alternative scenario differing in CO_2 emission costs serving as a sensitivity analysis. The model results can be interpreted whether additional incentives for flexibility will be required or if flexibility will come as a complement given a competitive system.

Previous literature on integrated analyses of flexibility in electricity systems can be divided into static (dispatch only) and dynamic (dispatch and investment) analyses. In a static analysis, Denholm and Hand (2011) use a reduced-form dispatch model to analyze the effects of higher flexibility requirements on the capacity mix. They state that in an isolated system, flexible resources, i. e., elimination of must-run technologies, are crucial for the utilization of fluctuating renewable generation. A unit-commitment approach, focusing on the operational integration, is chosen in Ummels et al. (2006). These authors find that flexibility (in terms of ramp rates) does not pose a problem for the Netherlands in 2012. However, they identify the need for wind curtailment due to minimum load restrictions. Lamadrid et al. (2011) conclude from their analysis of an optimal dispatch with varying capacities and ramping cost configurations that there is a need for a market for ramping products. In a dynamic analysis, Möst and Fichtner (2010), Nicolosi (2010) and De Jonghe et al. (2011) analyze investment decisions under operational constraints to determine an optimal capacity mix. They find that operational constraints tend to change the optimal capacity mix compared to when only considering achievable full load hours from base-load to mid- or peak-load capacities. By comparing model runs with and without operational constraints, Nicolosi (2012) states that utilization rather than operational constraints determines the investments of peak load capacities. However, previous research neglects the ambitious renewable targets of the EU, especially in the long term when flexibility becomes a greater issue for the electricity system. Moreover, demand side reactions to high wholesale prices in case of low renewable production or to

[5]The objective of the model is to minimize total system costs of the electricity supply for the exogenously defined electricity demand. Total system costs include investment costs, fixed operation and maintenance costs, variable production costs (which comprise fuel and CO_2 costs) as well as costs due to the start ups of thermal power plants.

volatile wholesale prices in general have not yet been analyzed. We therefore contribute to this literature by considering the long-term developments in transitioning to a mostly renewable electricity system in Europe, especially with regard to a renewable-dependent provision of balancing power. Furthermore, previously not considered flexibility options on the supply (flexible CCS plants) and demand side (demand side management) are considered.

We find that an increase in fluctuating renewables has a tremendous impact on the volatility of the residual load and therefore on flexibility requirements. However, any market design that incentivizes investments in least (total system) cost generation does not need additional incentives for flexibility. Under the assumption of perfect competition the challenges of volatility and therefore flexibility are met by an increase in peak-load and a reduction in mid- and base-load capacities. Neither hourly load changes nor the provision of balancing power poses a challenge. Moreover, at every point in time of the simulation, the provision of balancing power is never a binding constraint, indicating excess flexibility provision. Therefore, the main trigger for investing in flexible resources are the achievable full load hours and the need for backup capacity. In a competitive market, the cost-efficient technologies most likely to be installed, i. e., gas-fired power plants or flexible CCS plants, provide flexibility as a by-product. Under the condition of system adequacy, flexibility never poses a challenge in a cost-minimal capacity mix.

As renewable support is currently discussed and partly reduced in various EU countries, the future development of renewable deployment is rather uncertain. Assuming a realization of the EU 2050 goals, however, can be seen as an upper bound of flexibility demand in a very high RES-E share energy system. Our results show that even in such an optimistic RES-E scenario, flexibility does not become an issue of system adequacy.

The remainder of this paper is organized as follows: Section 1.2 defines the used concept of flexibility and flexibility options in electricity systems, Section 1.3 presents the applied methodology and underlying assumptions. In Section 1.4, results with regard to the change in flexibility requirements and the adaption of the electricity system are analyzed. Section 1.5 concludes and discusses policy implications.

1.2 Flexibility in electricity systems

In electricity systems demand and supply have to be balanced at any time. Flexibility on the supply side was in previous decades mainly necessary, because inelastic demand was subject to fluctuations, following daily, weekly and seasonal patterns. Recently and with increasing importance the source-dependent volatile electricity generation by renewable energies becomes relevant for the evaluation of needed flexibility.

One can – depending on the considered time period – distinguish two kinds of flexibility.[6] On the one hand, variability relates to longer time frames (larger than 1 h) and especially to the need of thermal power plants to adapt to changing residual load (i. e., demand minus generation by renewable energies such as wind and solar).[7] Renewables do not cause variable generation costs and are thus usually dispatched prior to thermal plants (and depending on market regulations even required to do so). With the increasing share of fluctuating electricity generation from renewables, the demand served by thermal plants is thus subject to a higher variation.

On the other hand, in the shorter time span up to about 1 h, the need for flexibility options mainly arises from the deviation between forecast renewable generation and actual outcome (forecast errors of demand are of minor magnitude). As this deviation occurs on short notice, the electricity system has limited options to adapt, as for instance older power plants need more time to adapt their electricity output, especially if they have to start up first.

1.3 Methodology and assumptions

Due to the expected structural changes in electricity systems, historical data cannot be used to analyze the effects of a high share of renewables on the optimal capacity mix and on the future role of flexible resources. This renders an econometric analysis impossible. Nevertheless, an integrated analysis is necessary due to the possible contribution from all parts of the electricity system to flexibility. For this analysis, we apply the electricity market model (DIMENSION) of the Institute of Energy Economics at the University of Cologne, as presented in Richter (2011).[8]

The model optimizes investments and dispatch of conventional, nuclear, storage and renewable technologies up to 2050 via cost minimization. Demand is assumed to be fixed (excepted for the option to shift demand of DSM-processes). Moreover, competitive markets are assumed. Investment and generation decisions are based on perfect foresight. The model balances demand and supply in every considered market for every hour of the year. Imports and exports can contribute fully to the balancing but only partly

[6] A more detailed elaboration on the definition of flexibility in electricity systems can be found in, e. g., IEA (2011a).

[7] The electricity grid also provides some kind of flexibility as it spreads demand and renewable power generation over a broader geographical area. Grid extensions are not part of the optimization model but exogenous, however their contribution is incorporated through the trade between markets.

[8] See also Fürsch et al. (2011), Nagl et al. (2011), Fürsch et al. (2012) or Jägemann et al. (2012).

to the peak demand constraint. In addition to the need to cover demand at all times, a peak-load constraint has to be fulfilled. This constraint requires sufficient (secured) capacity to be available to cover a historic peak-demand (including a security margin), with interconnector capacities being partially credited.

Further equations include constraints on electricity generation and technologies (such as general availability due to revisions or existing nuclear construction restrictions), storage level restrictions and net transfer capacities. All technologies are subject to an hourly availability, which allows us to model a fluctuating feed-in structure of renewable wind and solar technologies. For every hour and region, there is maximum feed-in derived from solar irradiation and wind speeds. The model therefore can decide not to use the full amount of RES-E generation available, i. e., curtail RES-E generation. The available feed-in of RES-E is calculated for every market via underlying subregions (47 for onshore, 42 for offshore and 28 for photovoltaics) to account for geographical patterns.[9] The regional focus of the model in Europe is due to the expected integrated European market. Given the expected integration, changes concerning the electricity system in one country have high influence on neighboring countries. This is especially relevant for the deployment of a large amount of RES-E since this produces significant changes in the supply structure.

Within the investment model a typical day approach is used, capturing seasonal, weekly and daily patterns for demand and RES-E generation. In the detailed dispatch calculation, a 8760 h time series is used. The investment model and the dispatch calculation are linked via capacities. The capacities of the investment model are fed into the high resolution dispatch model, wherein time series effects and more possible dispatch situations can be modeled. The equations used for the dispatch are the same in both models, only the parameters of the time series differ.

Stochastic influences (short-term) are accounted for by the procurement of balancing power for the adjustment of forecast errors of renewables generation. The development of installed renewable energy capacities is exogenous as it is mainly driven by political will rather than a result of market dynamics. To account for long-term uncertainty we analyze two scenarios as described in Section 3.6. As we focus on flexibility options we restrict the display of more detailed modeling aspects in the following to renewables, start-up restrictions of thermal power plants, storages (including DSM) and detachable CCS units.[10] These options are relevant for short-term balancing of supply and demand and comprise the major model improvements.

[9] Cf. Fürsch et al. (2011) for more detail.

[10] Flexible CCS power plants may not contribute much to the needed flexibility in absolute terms, however, observing when this ability is used, can be highly relevant for the interpretation of flexibility within the modeling approach.

1.3.1 Model description

The following table provides an overview of the most important model sets, parameters and variables:[11]

Table 1.1: Model abbreviations including sets, parameters and variables

Abbreviation	Dimension	Description
Model sets		
$a \in A$		Technologies
$k \in A$	Subset of a	Technologies starting-up within 1 h
$l \in A$	Subset of a	Technologies starting-up in more than 1 h
$s \in A$	Subset of a	Storage technologies
$r \in A$	Subset of a	RES-E technologies
$f \in A$	Subset of a	CCS technologies with attached CCS unit
$g \in A$	Subset of a	CCS technologies with detached CCS unit
$w \in A$	Subset of a	Wind technologies
$m \in M$		DSM processes
$c \in C$		Countries
$e \in C$	Subset of c	Subregions
$d \in D$		Days
$h \in H$		Hours
$y \in Y$		Years
Model parameters		
ac_a	€$_{2010}$/MWh	Attrition costs for ramp-up operation
an_a	€$_{2010}$/MW	Annuity for technology specific investment costs
$\mathrm{av}^{d,h}_{c,a}$	%	Availability
$\mathrm{de}^{d,h}_{y,c}$	MW	Demand
dr_y	%	Discount rate
ef_a	t CO_2/MWh$_{th}$	CO_2 emissions per fuel consumption
fc_a	€$_{2010}$/MW	Fixed operation and maintenance costs
$\mathrm{fu}_{y,a}$	€$_{2010}$/MWh$_{th}$	Fuel price
cp_y	€$_{2010}$/t CO_2	Costs for CO_2 emissions
$\mathrm{pd}^{d,h}_{y,c}$	MW	Peak demand (increased by a security factor)
η_a	%	Net efficiency
$\tau^{d,h}_{y,c,a}$	%	Capacity factor
$\mathrm{du}^{d,h}_{y,c}$	MW	Acquired positive balancing power
$\mathrm{dd}^{d,h}_{y,c}$	MW	Acquired negative balancing power
dv	%	Maximum deviation between RES feed-in and forecast
$\mathrm{ll}^{d,h}_{y,m}$	MW	Lower limit of demand of DSM process
$\mathrm{ul}^{d,h}_{y,m}$	MW	Upper limit of demand of DSM process
ml_a	%	Minimal load
su_a	hours	Inverse of start-up time
pr_a	hours	Precision of start-up representation
Model variables		
$\mathrm{AD}_{y,c,a}$	MW	Commissioning of new power plants
$\mathrm{CU}^{d,h}_{y,c,a}$	MW	Online capacity
$\mathrm{CUP}^{d,h}_{y,c,a}$	MW	Capacity switched-on
$\mathrm{CDO}^{d,h}_{y,c,a}$	MW	Capacity switched-off
$\mathrm{GE}^{d,h}_{y,c,a}$	MW	Electricity generation
$\mathrm{IM}^{d,h}_{y,c,c'}$	MW	Net imports
$\mathrm{IN}_{y,c,a}$	MW	Installed capacity
$\mathrm{LVL}^{d,h}_{y,c,a}$	MWh	Storage level
$\mathrm{ST}^{d,h}_{y,c,s}$	MW	Consumption in storage operation
TCOST	€$_{2010}$	Total system costs (objective value)

[11] If not stated otherwise, MW are MW$_{el}$.

The objective of the model (Eq. 1.1) is to minimize discounted total system costs while meeting demand at all times:

$$\min \ TCOST = \sum_{y\in Y}\sum_{c\in C}\sum_{a\in A}\Bigg[dr_y \cdot \Bigg(AD_{y,c,a}\cdot an_a + IN_{y,c,a}\cdot fc_a$$

$$+\sum_{d\in D}\sum_{h\in H}\Bigg(GE^{d,h}_{y,c,a}\cdot\left(\frac{fp_{y,a}+cp_y\cdot ef_a}{\eta_a}\right)+CU^{d,h}_{y,c,a}\cdot\left(\frac{fp_{y,a}+cp_y\cdot ef_a}{\eta_a}+ac_a\right)\Bigg)\Bigg)\Bigg] \tag{1.1}$$

Total system costs include investment, fixed operation and maintenance, variable production and thermal power plant start up costs. Investment costs are annualized with a 5 % interest rate for the technology-specific depreciation time. Fixed costs occur for staff, insurance and maintenance. Variable production costs consist of costs for fuel and CO_2, and depend on the emission factor and net efficiency of the several technologies. Start up costs include costs of attrition and co-firing. Combined heat and power plants (CHP) are able to generate revenues from heat production and therefore reduce total costs.

1.3.2 Renewable-dependent provision of balancing power

Due to a high share of fluctuating RES-E, balancing power (i. e., essentially tertiary reserve) must be available to quickly balance electricity supply and demand if necessary. In contrast to the current common practice to contract a fixed amount of balancing power for whole days, weeks or even months, we assume the required balancing power to be dependent on the hourly changing requirements and thus adjust the demand according to wind and solar generation. That way, we can deduce the actual scarcity of flexibility options rather than a shortage due to an (arbitrarily) high magnitude of balancing power acquisition. We assume that the system has to be able to balance potential forecast errors of at least 10 % of expected wind and solar generation at all times. The quality of short-term prediction of wind and solar feed-in has improved in recent years due to improved forecast models. As stated in Giebel et al. (2011), relative forecast errors were reduced on average from about 10 % in 2000 to 6 % in 2006. However, in order to be able also to balance large forecast errors, high flexibility is nevertheless still needed (cf. Holttinen, 2005, Holttinen and Hirvonen, 2005).

Constraint (Eq. 1.2) was added to the model to ensure sufficient short term flexibility in order to increase production at all times. The parameter du represents the potential need for positive balancing power, which is set to 10 % of available wind and photovoltaic feed-in for each hour in every country.

$$\begin{aligned} du_{y,c}^{d,h} \leq & \sum_{l \in A} \left(CU_{y,c,l}^{d,h} - GE_{y,c,l}^{d,h} \right) + \sum_{k \in A} \left(av_{c,k}^{d,h} \cdot IN_{y,c,k} - GE_{y,c,k}^{d,h} \right) + \sum_{s \in A} ST_{y,c,s}^{d,h} \\ & + \sum_{w \in A} \left((1 - dv) \cdot \left(av_{c,w}^{d,h} \cdot IN_{y,c,w} - GE_{y,c,w}^{d,h} \right) \right) \\ & + \sum_{f \in A} \left(GE_{y,c,f}^{d,h} \cdot \left(\frac{\eta_g}{\eta_f} - 1 \right) \right) + \sum_{m \in M} \left(ll_{y,c,m}^{d,h} - GE_{y,c,m}^{d,h} + ST_{y,c,m}^{d,h} \right) \end{aligned} \tag{1.2}$$

Table 1.2 gives an overview of available options to provide positive as well as negative balancing power in the electricity system.

TABLE 1.2: Overview of flexibility options

Positive flexibility	Negative flexibility
• Ramping of thermal power plants in part load operation • Open cycle gas turbines able to start operation within 15-20 minutes • Switching off CCS unit to increase power output • Utilization of stored energy or stop of storage • Shifting through demand side management (reduction) • Utilization of previously curtailed wind power	• Thermal power plants in operation (ramping down) • Storage technologies • Curtailment of wind power • Shifting through demand side management (increase)

Positive balancing power can be provided by thermal power plants in several ways: Technologies that need more than 1 h to start-up (l) are limited to increase their production by the amount of capacity currently in part load, i. e., online capacity minus current production. Highly flexible technologies, especially open cycle gas turbines, can start within 1 h (k) and thus can potentially provide balancing power even if currently not in operation. They can increase production until reaching installed and available capacity. Advanced CCS plants can, by switching off their CCS unit and thus accepting higher emissions, increase generation quickly. The maximum additional production can be calculated by multiplying the fraction of both efficiencies (with and without CCS) minus one with the current power generation of plants with applied CCS (cf. Section 3.3; see Davison, 2009; Martens et al., 2011).

Storage units (e. g., pump storage, compressed air storage or batteries) are in general very flexible and can by either increasing generation or reducing storage operation provide positive balancing power. Demand side management acts very similar from a system perspective as a reduction in electricity demand is comparable to generation and increased demand to consumption. Compared to classic storage, DSM processes usually have a limited time span in which shifts in demand have to be compensated and are restricted by the minimum power demand (GE indicates decreasing and ST increasing regular demand).

The last option for balancing electricity generation is withdrawing curtailment of wind production.[12] The available capacity is restricted to 90 % of the expected and therefore curtailed power (to account for forecast errors). As this is not associated with any costs (neglecting transaction costs) this should be usually the first option taken.

The following constraint represents the need for negative flexibility where dd is equal to 10 % of expected feed-in by photovoltaics:[13]

$$\begin{aligned} dd_{y,c}^{d,h} \leq & \sum_{l \in A} \left(GE_{y,c,l}^{d,h} - ml_l \cdot CU_{y,c,l}^{d,h} \right) + \sum_{k \in A} GE_{y,c,k}^{d,h} \\ & + \sum_{s \in A} \left(av_{c,s}^{d,h} \cdot IN_{y,c,a}^{s} - GE_{y,c,a}^{d,h} \right) + \sum_{w \in A} GE_{y,c,w}^{d,h} \\ & + \sum_{g \in A} GE_{y,c,g}^{d,h} \cdot \left(\frac{\eta_f}{\eta_g} \right) + \sum_{m \in A} \left(ul_{y,c,m}^{d,h} - ST_{y,c,m}^{d,h} + GE_{y,c,m}^{d,h} \right) \end{aligned} \tag{1.3}$$

The options for providing negative balancing power, i. e., in the case of excess electricity generation, are ramping down thermal power plants, increasing storage or decreasing turbine operation of storage (including DSM) and curtailing wind generation. Running power plants that cannot shut-down operation on short notice are only able to reduce production to minimum load. Highly flexible plants (e. g., gas turbines), on the contrary, can stop production completely. Storage can, in addition to reducing production, increase power consumption. Measures taken by thermal plants as well as storage usually reduce system costs (by fuel cost savings and the option to use (free of cost) stored electricity at other times respectively) and are thus chosen previously to wind curtailment. Flexible CCS power plants that are not using their CCS units can switch-on CO_2 segregation and thus reduce efficiency and production. DSM processes can increase consumption until their maximum demand is reached.

[12] By assumption other renewable energy sources – for instance photovoltaic panels – are not considered for withdrawing of curtailment as the installations are usually small in size, no technical steering possibilities are installed and thus the effort to activate sufficient capacity comparably high. But as wind generation is generally sufficiently available in case any curtailment took place, this does not change any results.

[13] An underestimation of wind feed-in can be balanced by wind curtailment, thus there is no additional need for negative flexibility.

1.3.3 Power plants with a detachable CCS unit

CCS technologies may become an important technology in the capacity mix in the future. Technologies with (captured) high emissions, i. e., lignite- and coal-fired power plants, might therefore still be supplying electricity in a low-carbon electricity system. Lignite- and coal-fired are considered base and mid-load and lack the flexibility of, e. g., open cycle gas turbines. Hence, an additional flexibility option for these power plants, might give insights into the deployment of base-load technologies which are simultaneously able to provide flexibility.

Flexible CCS plants have the option to switch off their capture unit and thereby increase power output, while simultaneously emitting more CO_2 into the atmosphere. This can be done on short notice and is thus suitable for the provision of short-term flexibility. CSS plants are thus able to provide both, short-term flexibility and required backup capacity to serve peak-demand. Depending on their dispatch they therefore allow some insights concerning the tightness of both requirements.

These units were modeled with the same constraints as conventional power plants, but with the possibility to switch between operation modes within 1 h. This was implemented by adding a new technology g for every power plant f with a CCS unit, where g represents the share of capacity f whose CCS unit is switched off.[14]

The following constraint ensures that the total online capacity of technology f (CCS switched on) and its counterpart g (same technology with CCS switched off) does not exceed the total available capacity. By multiplying the ramped-up capacity of g with the fraction of the efficiencies of f and g, the increased net efficiency of power plants with switched-off CCS can be taken into account by:

$$CU_{y,c,f}^{d,h} + CU_{y,c,g}^{d,h} \cdot \frac{\eta_f}{\eta_g} \leq av_{c,f}^{d,h} \cdot IN_{y,c,f} \tag{1.4}$$

Additional modifications for power plants with a detachable CCS unit have to be made when modeling start-up behavior. These will be pointed out in the following subsection.

1.3.4 Start-up of thermal power plants

The maximum and minimum operational capacities in one point in time are dependent on the plants' statuses of the previous hours. Time periods are freely selectable and by considering more points in a given time period more realistic start-ups of power plants

[14]The technical details of this process were taken from Davison (2009) and Finkenrath (2011).

can be modeled. Starting up capacity (variable CUP) is considered in the objective by a cost parameter which approximates co-firing costs, attrition costs etc.[15]

Equation 1.5 makes use of the variables CUP and CDO, which symbolize capacity that was started and shut-down from the previous hour to the current one:

$$CU_{y,c,a}^{d,h} = CU_{y,c,a}^{d,h-1} + CUP_{y,c,a}^{d,h-1} - CDO_{y,c,a}^{d,h-1} \tag{1.5}$$

The restriction on the maximum online capacity of technologies with a flexible CCS unit (represented by f if CCS switched on, and g if CCS switched off) is similar to an ordinary power plant:

$$CU_{y,c,f}^{d,h} + CU_{y,c,g}^{d,h} = CU_{y,c,f}^{d,h-1} + CU_{y,c,g}^{d,h-1} + CUP_{y,c,f}^{d,h-1} - CDO_{y,c,f}^{d,h-1} \tag{1.6}$$

As the online capacities of f and g belong to the same technology and since switching the CCS unit on and off can be done within 1 h, the two can be combined.

The maximum start-up of capacities from 1 h to the next depends on the overall available capacity, the capacity already in operation and the technology's start-up time (inverse of su_a). The model is taking into account the capacity that was started-up in previous hours:

$$CUP_{y,c,a}^{d,h} \leq \left(av_{c,a}^{d,h} * IN_{y,c,a} - CU_{y,c,a}^{d,h} + \sum_{i=h-pr_a}^{i<h} CUP_{y,c,a}^{d,t} \right) \cdot su_a \tag{1.7}$$

If power plants of one technology were starting-up in the previous hour, e. g., after all plants had been shut-down completely, then, under the assumption of a linear start-up trajectory, all plants are able to start-up with the same magnitude in all hours until reaching their maximum online capacity. On the contrary, if there was not any ramping activity in the previous hours, only the capacity currently not in operation is able to start-up. Parameter pr represents the precision of the modeling of start-up behavior ($0 \leq pr_a \leq \frac{1}{su_a}$).

The constraint has to be altered slightly for technologies with a detachable CCS unit by linking technology f with its counterpart g:

$$CUP_{y,c,f}^{d,h} \leq \left(av_{c,f}^{d,h} * IN_{y,c,f} - CU_{y,c,f}^{d,h} - CU_{y,c,g}^{d,h} + \sum_{i=h-pr_f}^{i<h} CUP_{y,c,f}^{d,t} \right) \cdot su_f \tag{1.8}$$

[15] Depending on the different technologies, these costs are set to about 1/5 of the variable cost.

The restriction for shutting-down technologies can be enhanced analogously by replacing the original constraint with the following:

$$CDO_{y,c,a}^{d,h} \leq \left(CU_{y,c,a}^{d,h} + \sum_{i=h-pr_a}^{i<h} CDO_{y,c,a}^{d,t} \right) \cdot su_a \tag{1.9}$$

And equivalently for technologies with detachable CCS unit:

$$CDO_{y,c,f}^{d,h} \leq \left(CU_{y,c,f}^{d,h} + CU_{y,c,g}^{d,h} + \sum_{i=h-pr_f}^{i<h} CDO_{y,c,f}^{d,t} \right) \cdot su_f \tag{1.10}$$

1.3.5 Storages and DSM

Storages (e. g., pumped-storage plants, reservoirs and CAES) and demand side management (DSM) processes can be modeled similarly with the latter possessing additional restrictions to account for the maximal time span demand can be shifted.

The basic equation for storages keeps track of the current storage level which is computed from the electricity consumption (taking into account efficiency loss), production and, if applicable, natural inflow (respectively outflow) with generation and consumption being restricted to the available capacity:

$$LVL_{y,c,a}^{d,h} = LVL_{y,c,a}^{d,h-1} + ST_{y,c,s}^{d,h} \cdot \eta_a - GE_{y,c,a}^{d,h} + inf_{y,c,a}^{d,h} \tag{1.11}$$

In the case of DSM, these limits are not solely depending on the installed capacity but also depend on the specific hour. Therefore, we account for variations in the utilization of the processes. In the considered period total consumption and production have to be balanced. For DSM processes this constraint can be be more restrictive (depending on the kind of process):

$$\sum_{y,d,h}^{d=d',h=h',y=y'} ST_{y,c,s}^{d,h} \cdot \eta_a - GE_{y,c,a}^{d,h} + inf_{y,c,a}^{d,h} = 0 \tag{1.12}$$

d', h', y' indicate the relevant time span in which consumption and production have to be balanced. This way, e. g., cooling systems can be forced to catch up on electricity consumption within 4 h to avoid any damage to cooled goods (see Table 1.9).

We take into account 28 different DSM processes for each region, grouped by sectors (see Table 1.8 in the Appendix). Technical specifications include the balancing interval,

i. e., the time a deviation from original electricity consumption has to be recovered, efficiency, which represents losses due to rescheduling, and maximum reduction and increase of demand (see Table 1.9 in the Appendix). The latter figures represent limits for demand adjustments due to the time dependent consumption of the processes. Table 1.10 (Appendix) displays assumed DSM capacities for each considered year.

1.3.6 Assumptions and scenario setting

Assumptions for the simulation include the regional electricity demand development, net transfer capacities between regions, capacities of existing power plants, technical and economic parameters for power plant investments as well as fuel and CO_2 prices.

Installed capacities of renewable energies are exogenous. The development and allocation between different technologies reflect current national and European policies as well as regional investment costs and potentials. Where available, national targets have been considered (e. g., national renewable energy action plans (NREAP) for EU member states). As official long-term plans are not available, assumptions based on available studies have been made for the subsequent time span (see Table 1.7 in the Appendix for installed capacities in 2050). Electricity generation from renewables reaches 75 % over all countries in 2050 (see also section 1.4.1). RES-curtailment does not impose any costs.

The setting chosen for this analysis is only one possible development and should not be interpreted as a forecast. The assumptions are based on several sources such as Capros et al. (2010), Prognos/EWI/GWS (2010), IEA (2011), ENTSO-E (2011) and Fürsch et al. (2011) and represent a trade-off between their projections. The underlying assumptions used in the scenario analysis can be found in the Appendix A.

The analyzed scenarios A and B only differ regarding assumed CO_2 prices. Thus they reflect and aggregate divergent expectations for influencing factors like the ambitiousness of CO_2 reduction, economic growth and emissions in other sectors. Since CO_2 emission costs may have effects on installed capacity (or generation) of base or peak load and storage capacities, a sensitivity analysis with a higher CO_2 price is performed. The underlying assumption is that in Scenario A CO_2 prices increase up to 50 EUR_{2010}/t CO_2 in 2050 and in Scenario B up to 100 EUR_{2010}/t CO_2 in 2050. In addition to CO_2 emission costs no additional restrictions (i. e. CO_2 targets) are imposed. Table 1.3 depicts the assumed CO_2 emission prices from 2020 to 2050.

TABLE 1.3: Assumed CO_2 emission prices [EUR_{2010}/t CO_2]

	2020	2030	2040	2050
CO_2 price in Scenario A	22.6	31.8	40.9	50.0
CO_2 price in Scenario B	35.1	56.8	78.4	100.0

1.4 Results

In this section, the results from the analysis are presented. First, the impacts of an increasing share of RES-E on the residual load are discussed with examples of selected European electricity systems up to 2050. Note, that these examples rather reflect a possible development of electricity systems with certain characteristics (e. g., integration into the European electricity grid or renewables mix) than a projection of future developments for the shown countries. Second, the adaption of the system to the increasing share of renewables with regard to capacity and generation mix focusing on flexible resources is discussed. Third, aspects of the different flexibility requirements in terms of modeling constraints are analyzed in detail. Finally, an overview of the implications for market design is given.[16]

1.4.1 Impacts of an increasing share of RES-E

Due to the negligible variable costs of RES-E, they can be integrated on the left-hand side of the merit order. This means they are usually dispatched before other supply technologies. The impact of an increasing share of renewables can thereby best be discussed by analyzing the residual load to be covered by other technologies. The impact is twofold, on the one hand the (residual) load duration curve is affected and the achievable full load hours for other technologies reduced. On the other hand, the hourly changes of residual load possibly impose additional flexibility of the other supply technologies. Furthermore the provision of balancing power becomes more relevant due to possibly increasing absolute forecast errors.

Based on simulation assumptions, the RES-E share on gross electricity demand in Europe increases from 34 % in 2020 to 54 % in 2030, and to 75 % in 2050. In the short term (until 2020), hydro-power (39 % of RES-E generation) and onshore wind (26 % of RES-E generation) are the most deployed renewable energy sources. Due to the assumed large deployment of on- and offshore wind turbines, more than 50 % of the renewable energy is provided by wind power in 2050. Solar technologies – mainly deployed in southern

[16]Detailed numerical data can be found in the Appendix. This section highlights the developments relevant to the topic of the paper.

Europe – generate about 22 % of the renewable energy. We illustrate the effects of such a high share of renewables with the examples of Germany and the UK. Both countries are chosen due to their geographical position within Europe and the assumptions on renewable deployment. The examples should be thought of as illustrative case studies rather than forecasts for the development of the two countries. While Germany is well-connected to its neighboring countries, the UK only has few interconnections and is closer to an insular system. For Germany, the renewable technologies, i. e., wind and photovoltaics, are by assumption diversified, whereas the renewable capacities in the UK consist mostly of on- and offshore wind capacities, which lead to greater challenges due to the fluctuating nature of wind. In 2050, Germany has a renewable generation share of 61 % of gross electricity consumption, of which about 64 % is wind and 20 % PV. The UK has a renewable share of about 76 % with over 90 % wind.

1.4.1.1 Residual load

The high share of renewables has significant effects on the residual load as shown for Germany and the UK in Figure 1.1.[17]

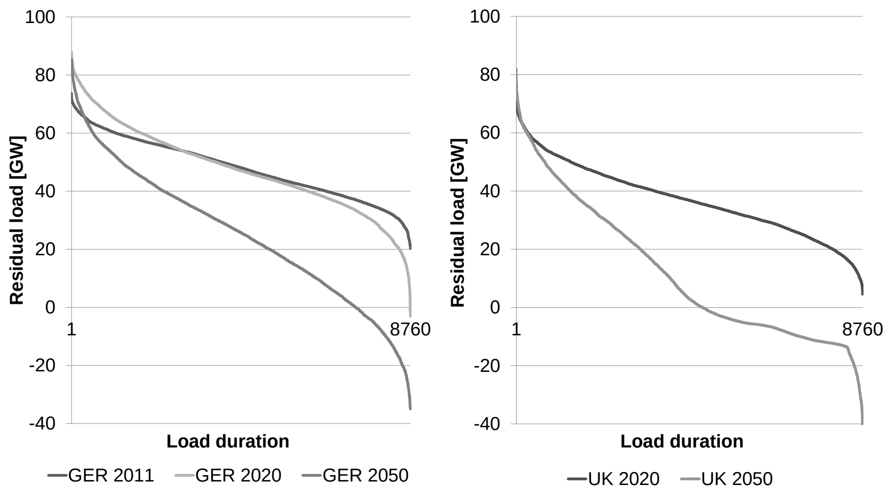


FIGURE 1.1: Residual load duration curve for Germany(left) and UK (right)

From the historical 2011 data to the assumed feed-in in 2020, the residual load duration curve for Germany changes slightly due to the assumed increase in electricity consumption and in deployed renewables. The residual load duration curves for Germany and the UK are steeper in 2050. The number of hours with negative residual load increases and occurs for nearly half the hours in the UK, where renewable electricity generation exceeds actual demand by up to 40 GW. Despite these developments, hours with high

[17]Data source for 2011 load in Germany is ENTSO-E. Wind and photovoltaic generation data for 2011 is from the European Energy Exchange (EEX). For the UK, no data for the renewable feed-in was available.

load levels remain. This means that achievable full load hours for conventional generation are reduced, but backup capacities for hours with high levels of residual load are still needed. The effects on the residual load depend on the installed renewable technology. In Italy and the Iberian Peninsula, for example, the shape of the residual load curve in 2050 is similar to the curves in 2020 due to the high shares of CSP plants with integrated thermal storages. CSP smoothes residual load by using its thermal storage unit and reduces the effects of fluctuating generation.

1.4.1.2 Volatility of residual load

The volatility of residual load is analyzed on an hourly basis. Figure 1.2 depicts the boxplots for Germany in 2011, 2020 and 2050 and for the UK in 2020 and 2050.

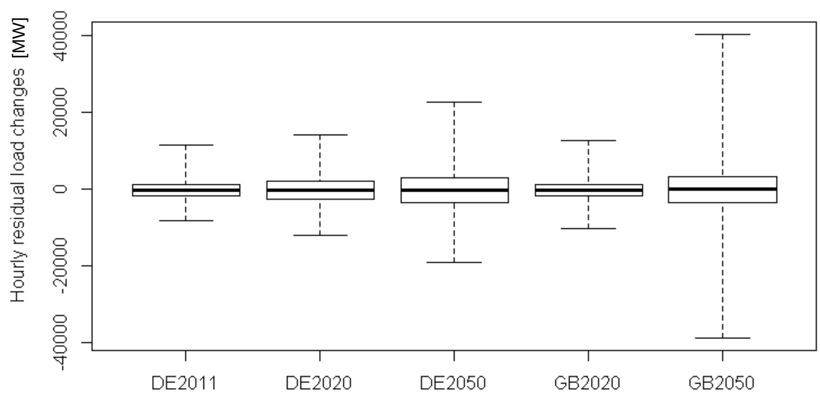


FIGURE 1.2: Box plot of hourly changes for Germany and the UK 2020 and 2050

Two main developments can be identified. First, the extreme values grow larger with a higher share of fluctuating renewables. Still, in 2020, only a few hours with an absolute change of more than 10,000 MW occur in any country. In 2050, all countries with a residual load of more than 40 GW face hourly changes (positive and negative) greater than 10,000 MW. In countries with high demand and high penetration of renewables, hourly fluctuations up to 40,000 MW (UK) in residual load occur more often. The power systems in Germany, France, Scandinavia and the Iberian Peninsula still face hourly load changes of around 20,000 MW. Smaller countries, like Denmark, may have to deal with smaller changes in absolute amounts but experience extreme hourly changes relative to the residual load level. For the electricity system, large changes in times of low or negative residual load are especially challenging. Due to a high share of renewable generation in these hours, no conventional capacity is running and must therefore be started up. This requires sufficient flexible resources that are able to start up quickly. The second development is that there is a more widespread distribution of hourly changes. While in Germany the quartiles increase by about 50 %, in the UK these values double. Absolute hourly changes therefore increase tremendously, indicating an increased need for flexible resources able to adapt generation rapidly. These developments are also confirmed by analyzing means and standard deviation of positive and negative hourly changes as shown in Table 1.4. The means change in the same manner as the analyzed quartiles. The standard deviation changes significantly, indicating more widely distributed hourly changes.

Table 1.4: Mean, maximum and standard deviation of hourly load changes for Germany and the UK [MW]

	Germany			UK	
	2011	2020	2050	2020	2050
Mean positive	2,242	3,083	4,105	2,345	4,619
Standard deviation positive	2,148	2,572	3,373	2,229	4,739
Max positive	11,396	14,106	22,775	12,545	40,286
Mean negative	-1,853	-2,604	-3,656	-1,977	-4,661
Standard deviation negative	1,420	1,922	2,727	1,724	4,891
Max negative	-8,016	-12,069	-18,984	-10,186	-38,631

1.4.1.3 Provision of balancing power

Together with the higher feed-in of fluctuating renewables, forecast errors and therefore balancing power increase in absolute amounts as long as prediction is not improved. Figure 1.3 shows the duration curve of balancing power for renewables when 10 % of renewable generation must be provided as balancing power.[18]

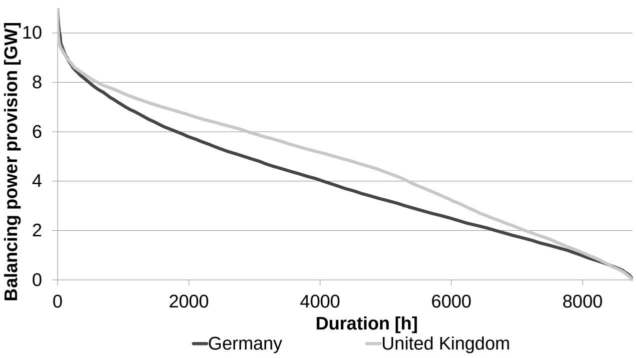


Figure 1.3: Required positive balancing power in Germany and the UK in 2050

[18]It seems reasonable to assume a variable balancing power provision, because the requirements for the conventional utilization of balancing power is assumeably not going to change. Therefore, only the renewables forecast errors have to be balanced, for which it seems reasonable to implement a balancing power market with a renewable-dependent quantity relatively close to physical dispatch. The introduction of quarter-hourly intra-day markets fulfills more or less this functionality.

For Germany and the UK, up to 10,000 MW is needed as provision of balancing power only for renewables. Compared to current values this is significantly higher (e. g. for Germany with around 2,000 MW for the minute reserve). Therefore, the requirement for flexible resources to provide balancing power to backup forecast errors or failures of RES-E increases.

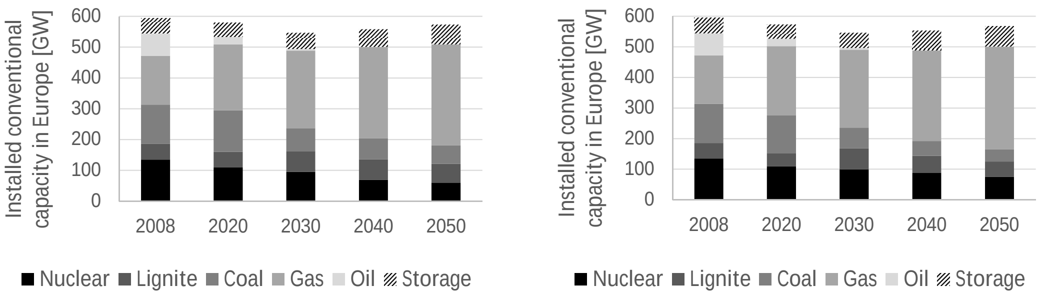
1.4.2 Adaptation of the electricity system

The changing residual load leads to changes in the electricity system. This section describes the development of capacity and generation mix with special focus on flexible resources.

1.4.2.1 Development of the capacity mix

The capacity mix changes significantly in both scenarios up to 2050 due to the large deployment of renewables. By assumption, RES-E capacities are primarily increased by onshore wind until 2020/2030, offshore wind from 2030 onwards and solar plants after 2030. Due to the low secured capacity of intermittent renewable technologies and an assumed increase in electricity demand, total gross capacity more than doubles by 2050. Renewables capacity amounts to 1.5 TW in 2050.[19]

Figure 1.4 depicts the gross electricity conventional capacities in Scenario A for the years 2020, 2030, 2040 and 2050 on the left side and for Scenario B on the right side. As can be seen, the difference of the two scenarios is rather small.



(A) Scenario A (B) Scenario B

FIGURE 1.4: European gross conventional capacity mix up to 2050

The overall conventional capacity in both scenarios remains relatively constant, but the share of base- and mid-load capacities decreases from 64 % in 2008 to 36 % in 2050. At the same time the share of gas-fired capacities (open and combined cycle) increased from

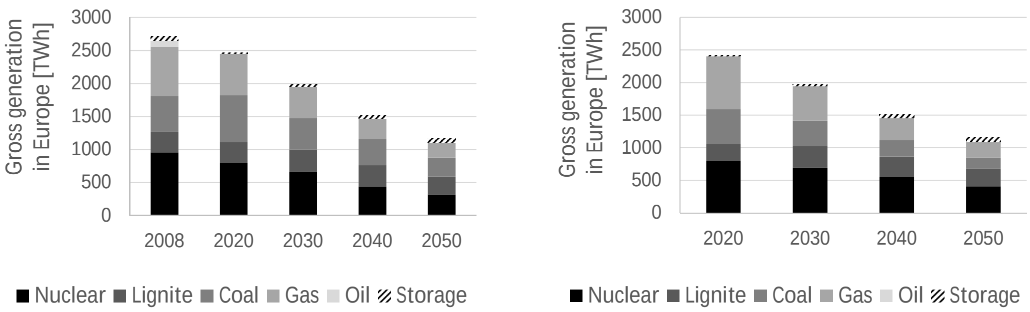
[19]For detailed figures on different technologies please consult the Appendix.

36 % to 64 %. Higher CO_2 prices in Scenario B lead to a small increase of nuclear and CCS capacities. However, this has little effect on the general mix between base/mid (33 %) and peak load (67 %) capacities. Storage is mainly deployed in countries with high amounts of negative residual load. An additional 22 GW of storage capacities are installed in Scenario A. In Scenario B, wind and solar curtailment is associated with higher costs due to higher costs of fossil fuel generation, making additional storage (4 GW) cost-efficient.

Hence, we observe that flexible resources, namely gas-fired power plants (and to a certain extent storage), contribute largely to a cost-efficient capacity mix with a high share of renewables.

1.4.2.2 Development of the electricity generation

The electricity generation from all conventional power plants decreases with the additional RES-E generation in both scenarios. The RES-E generation for whole Europe amounts to about 3,000 TWh.[20] Figure 1.5 depicts the gross conventional electricity generation in Scenario A for the years 2020, 2030, 2040 and 2050 on the left side and for Scenario B on the right side.



(A) Scenario A (B) Scenario B

FIGURE 1.5: European gross conventional electricity generation up to 2050

Higher CO_2 prices in Scenario B lead to a coal-to-gas switch. In Scenario B, about 200 TWh of electricity are generated in combined and open-cycle gas turbines instead of hard coal and lignite power plants. This includes 60 TWh of electricity generation from gas-fired CHP plants. More than 470 TWh of electricity is generated in coal and gas-fired power plants equipped with CCS units in 2050. Due to CO_2 prices of 100 EUR/t CO_2 in Scenario B in 2050, almost all conventional generation takes place in nuclear or fossil power plants equipped with CCS in the long term. CO_2 emissions in 2050 account to about 400 million tons in Scenario A and 152 million tons in Scenario B. For

[20]See the Appendix for detailed figures on the different technologies.

the emissions of the electricity system this means a reduction of 68 % (88 %) compared to 1990. More than 140 TWh of possible wind and solar generation, which represents about 7 % of total wind and solar generation, is curtailed in both scenarios in 2050. The main reason is the excessive feed-in due to the increased generation capacities. Load is already covered and it is not cost-efficient to build more storage capacities. This can also be seen by comparing the two scenarios: Storage capacities are slightly higher in Scenario B because of the additional opportunity costs of curtailment. Conventional generation is more expensive and therefore more RES-E generation is stored. However, the utilization rates of neither storage nor DSM technologies are significantly different in the scenarios. This indicates that smoothing of residual load due to storage rather depends on the costs of conventional generation rather than the capacity and generation mix.

By looking at the utilization rates of the conventional technologies in Table 1.5 , it can be seen that conventional generation decreases more than capacity, which can be explained by the sunk costs of capacities.

Table 1.5: Full load hours of conventional technologies [h]

		2020	2030	2040	2050
Scenario A	Nuclear	7271	6987	6373	5307
	Lignite	6252	4926	4688	4337
	Coal	5277	6365	5873	4837
	Gas	2908	1889	1037	678
Scenario B	Nuclear	45	-7	-77	131
(compared to A)	Lignite	5	-87	894	1083
	Coal	-1021	-643	-671	-549
	Gas	685	171	87	29

The utilization rate of gas-fired power plants decreases more than the rates of base and mid-load power plants. This is remarkable, especially given that the installed capacities are increasing. The reason for this could be that gas-fired peak load plants are used as backup capacity. However, the observation of low utilization rates alone does not allow a final conclusion that gas-fired capacities as flexible resources are installed mainly as backup capacities. Even with low utilization rates, the increased amount of gas-fired and storage capacities might be part of the cost-efficient capacity mix due to their high flexibility. To analyze this issue further, the modeling and the fulfillment of flexibility requirements in the model have to be considered.

1.4.3 Fulfilling flexibility requirements

In the analysis of the system adaption, the most remarkable result regarding flexibility options in the electricity system, was the increase of gas-fired capacities (especially open-cycle turbines) with a simultaneous, disproportionately high decrease of utilization rates. From the development of storage and DSM utilization, no conclusion about the importance of flexibility in the system could be drawn. However, the development of the gas-fired capacities deserves a closer look.

To analyze this effect further, some modeling aspects have to be considered. First, recall that flexibility requirements in the model stem from the flexible resources needed for load-following, i. e., mainly start-up constraints, and from the provision of balancing power. As Nicolosi (2012) already showed, ramping constraints alter the capacity mix to a certain extent towards more flexible resources. However, if we compare the two scenarios, we see that the conventional capacity mix changes towards CCS-technologies due to the high CO_2 prices. Wind curtailment becomes more costly, so additional storage is built to prevent wind curtailment from excessive generation and to smooth load following. The amount of additional storage however is small compared to the effects of the switch to CCS-technologies. We even see an increase of utilization rates of base-load (lignite) and a decrease in mid-load (coal) while the peak load (gas) utilization rates remain relatively constant. So even with a strong increase in the opportunity costs of curtailing wind and load following, the conventional capacity mix does not become more flexible due to ramping constraints, since the amount of gas-fired capacities does not change significantly.

Another possibility for the expansion of gas-fired capacities could be the provision of balancing power. From conventional generation, only online capacity or quickly started up capacity (open cycle gas turbines) can contribute to the provision. However, the constraint for providing balancing power never becomes relevant. Even in the peak load hours where there is nearly no renewables feed-in and all available conventional capacity is running (under consideration of the security margin). The marginal costs of providing an additional megawatt of balancing power are zero and hence there are always flexible resources available. So, why is this balancing power provision not important in terms of flexibility? For illustration purposes Fig. 1.6 shows the availability of negative balancing power for a summer week in Germany in 2020. The black line symbolizes the renewable-dependent provision, i. e., 10 % of the feed-in from wind and photovoltaics.

As can be seen, the provision requirement corresponds to the RES-E feed-in. Hence, if RES-E generation can be curtailed, the provision of negative flexibility resulting from the RES-E feed-in is easily fulfilled. In times with low available capacities to provide

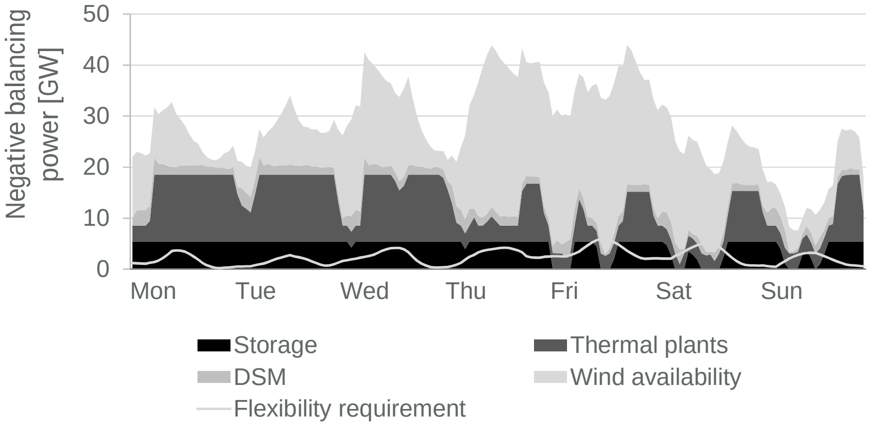


Figure 1.6: Availability of negative balancing power in June 2020 in Germany

negative flexibility, the demand for provision is also low due to the low renewables feed-in. In times with low feed-in, the availability is low, but barely any negative flexibility is needed since strong negative deviations cannot occur. With the rising share of wind capacities and the possible curtailment, sufficient negative balancing power is always available.

Figure 1.7 shows the availability of positive balancing power in the same week as demonstrated before in Germany 2020.

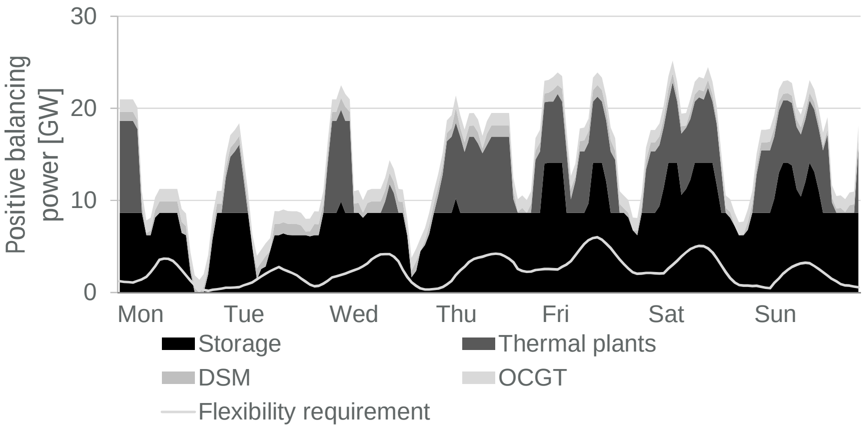


Figure 1.7: Availability of positive balancing power in June 2020 in Germany

Again, the available capacity corresponds to the provision requirement. During hours with high requirements, sufficient capacity for providing additional generation on short notice is sufficiently available. This is due to the fact that conventional generation is replaced by renewable generation, therefore capacities are idle and contribute to the availability of balancing power. Up to 2050, the availability of positive balancing power changes according to the source, i. e., there are more gas-fired power plants (OCGT) and less capacities in part-load.

Since positive and negative balancing power requirements never pose any challenge in any country considered, we conclude that the backup effect of the gas-fired capacities is significantly more relevant. Apparently, the gas-fired capacities are built because they are needed as a backup capacity rather than for providing flexibility during other hours. This finding is backed up by a closer look at the behavior of flexible CCS power plants. These power plants have the ability to shortly increase output by detaching their CCS unit. However, the only use of this ability is found in hours with scarce capacity, not in hours with high balancing power provision requirements or in hours with steep changes in residual load. This means that in a cost-efficient capacity mix the realizable full load hours matter more than flexibility issues. The main reason is surely the fact, that conventional capacity which is cost-efficient with low realizable full load hours, i. e., open cycle gas turbines with low capital costs and high variable costs, is highly flexible. Therefore, flexibility comes as a complement of a cost-efficient capacity mix in electricity systems with a high share of RES-E.

1.4.4 Implications for market design

For a proper discussion of the implications for market design on base of the performed analysis, three major points deserve closer consideration.

First, we assumed an exogenous deployment of RES-E, which may or may not reflect future European policies. However, this assumption is more in favor of our argumentation since it imposes stronger distortion for the electricity system, especially regarding the requirements for flexibility. The purpose of this paper was to take a closer look on flexibility in systems with a high share of renewables. Of course, all results are subject to the current information, e. g. the cost structure of renewables. However, as long as the renewables, which are deployed in the electricity system are not dispatchable, the full-load hour argument still holds.

Second, the scenario assumption about CO_2 prices is rather high compared to current trends. We do not know how the CO_2 prices develop and hence, we cannot finally conclude whether current price levels are a good estimation for future price levels. However, we can estimate the impacts of lower CO_2 prices for the model results. The differences of the scenarios with prices of 50 €/t emitted CO_2 (2050 in Scenario A) and 100 €/t in 2050 in Scenario B led to a fuel-switch from coal to gas and to more deployment of CCS capacities. If we assume lower prices for CO_2, there would be no significant changes in the capacity mix, since the merit order of the generation technologies does not change compared to Scenario A.[21] The general conclusions from the results will not be altered. What will change, however, is the relative value of renewables curtailment. Conventional generation will be cheaper and hence, less storage will be built. But since the difference of storage investments between the two scenarios is low, it will presumably have little impact if the CO_2 price is lower than in Scenario A.

Third, the result of the model due to the chosen approach is always an adequate, cost-efficient electricity system, implicitly assuming perfect competition and neglecting other distortions. The capacities built in the model are per definition profitable, and hence the discussion about Energy-Only-Market and capacity markets is irrelevant. For our approach, the only relevant issue is whether the market design produces a cost-efficient capacity mix. It does not matter, how this capacity mix is achieved. Hence, we can neglect the discussion of suitability of Energy-Only-Markets.

At the same time, the dependency of our results on a cost-efficient capacity mix means that any policies distorting this capacity mix could alter our results. For example, guaranteed payment for nuclear in the UK could increase the base load capacities and therefore change the capacity mix. As a consequence this could influence the availability of flexible resources. The other way round this implies that distorting the cost-efficient capacity mix by forcing certain technologies into the market, might create negative externalities with respect to available flexibility.

What we can draw as a conclusion for market design – with the stated limitations – is therefore, that assuring system adequacy is more important than introducing any additional incentives for flexibility. If even balancing power does not pose any challenges with a high share of renewables, additional instruments for incentivizing investments into flexible resources despite the existing spot and balancing market are certainly not necessary.

1.5 Conclusions

Electricity systems with a high share of renewables are confronted with an increasing requirement for flexibility. If the market does not provide sufficient flexibility and requires additional incentives, market design may be affected. In this paper, we analyzed this issue for the European electricity system. In an integrated system analysis, a linear investment and dispatch model is used to simulate the development of electricity markets in Europe up to 2050. The model was extended by including constraints for the provision of balancing power provision depending on renewable feed-in, demand-side

[21] We counter-checked this result by analyzing the fuel-switch induced by lower CO_2 prices than 50 €/t.

reactions, start-up processes of conventional power plants and flexible CCS power plants with a detachable CCS unit.

The results of the integrated analysis show that achievable full load hours of conventional capacities are reduced as renewable generation increases. Depending on the fluctuating renewable share, the volatility of the residual load increases and significantly impacts the electricity system. In 2050, when, e. g., for Germany and the UK with 50 % and 70 % of fluctuating renewables respectively, the spread of hourly changes increase by 50 % in Germany and doubles in the UK. Extreme values of hourly changes occur more often and reach up to 40,000 MW in the UK due to the high wind penetration. In other countries with a more balanced renewable portfolio, values around 20,000 MW still occur. Provision of balancing power for forecast errors increases and, given a 10 % provision of renewable feed-in, reaches over 10,000 MW in some hours.

The system adapts to the reduced achievable full load hours by adding more peak-load capacities, i. e., gas-fired power plants. Due to the relatively low investment costs, they serve as cost-efficient backup technologies. With higher CO_2 prices, the general case does not change: only more conventional capacity is equipped with CCS. Due to different storage investments in Scenario A and B, storages seem mainly to be built to prevent renewable curtailment, rather than to provide flexibility. This conjecture is confirmed by the fact that the provision of balancing power is never a binding constraint throughout the whole simulation. Therefore, at every point in time, excess capacity is able to ramp up within 15 minutes, allowing the electricity system to deal with any flexibility requirement. This finding is supported by the analysis of the utilization of flexible CCS power plants. The ability of these plants to provide generation in short time is only beneficial if renewable feed-in is low during peak-times – but not for the purpose of providing flexibility in hours with high volatility. Therefore, we conclude that the main trigger for investments in flexible resources such as gas-fired power plants or flexible CCS plants is system adequacy. Flexibility is a by-product of the cost-efficient adaptation to the reduced achievable full load hours under system adequacy.

Under the condition of system adequacy, flexibility never poses a challenge in a cost-minimal capacity mix. Therefore, any market design incentivizing investment in efficient generation thus provides flexibility as an inevitable complement.

Our results, however, depend on our assumption on current costs and technological progress. Disruptive innovations and changes in national energy policies might alter the future demand of flexibility. It seems, however, rather unlikely that such changes will increase the demand for flexibility beyond the level assumed in our analysis.

1.6 Appendix

1.6.1 Appendix A: Model assumptions

Table 1.6: Net electricity demand [TWh_{el}] and potential heat generation in CHP plants [TWh_{th}]

		2020		2030		2040		2050	
Austria	(AT)	65.3	41.2	70.0	41.5	74.3	41.8	78.5	42.0
BeNeLux	(LU)	221.6	129.9	237.6	130.8	252.2	131.5	266.5	132.3
Czech Republic	(CZ)	69.9	55.1	78.8	55.7	88.3	56.4	98.5	57.0
Denmark	(DK)	40.5	54.7	43.4	55.1	46.0	55.4	48.6	55.7
Eastern Europe	(EE)	151.9	132.6	171.1	134.2	191.8	135.7	214.0	137.2
France	(FR)	480.0	31.6	514.6	31.8	546.4	32.0	577.2	32.2
Germany	(DE)	567.0	192.4	584.2	192.9	584.2	192.9	584.2	192.9
Iberian Peninsula	(IB)	354.5	72.9	409.4	73.9	470.5	75.0	538.0	76.0
Italy	(IT)	362.9	169.2	419.1	171.7	481.6	174.1	550.7	176.5
Poland	(PL)	140.0	93.3	157.8	94.4	176.9	95.5	197.3	96.6
United Kingdom	(UK)	415.5	68.1	445.6	68.6	473.0	69.0	499.7	69.3
Scandinavia	(SK)	365.4	98.1	391.8	98.8	415.9	99.4	439.4	99.9
Switzerland	(CH)	65.4	3.0	70.1	3.0	74.5	3.0	78.7	3.0

Table 1.7: Gross installed capacities of renewable energies in 2050 [GW]

	Biomass	Biomass-CHP	Wind onshore	Wind offshore	PV	CSP	Geothermal
Austria	1.7	0.7	4.4	0.0	9.0	0.0	0.4
BeNeLux	5.8	3.1	14.0	35.7	2.2	0.0	1.3
Czech Republic	3.4	1.5	16.7	0.0	10.1	0.0	0.6
Denmark	3.7	1.6	4.4	10.5	0.0	0.0	0.3
Eastern Europe	9.8	4.3	31.2	0.0	31.1	0.0	0.4
France	6.7	2.9	71.5	62.0	62.4	27.3	0.6
Germany	11.4	5.0	63.4	48.9	91.8	0.0	2.2
Iberian Peninsula	4.7	2.1	74.1	5.4	28.0	49.9	1.4
Italy	6.4	2.8	53.2	19.3	52.1	49.9	4.1
Poland	5.7	2.5	42.1	26.9	11.9	0.0	0.4
United Kingdom	9.8	4.3	53.4	93.8	2.7	0.0	1.3
Scandinavia	6.5	2.8	19.4	36.6	0.0	0.0	1.8
Switerland	3.5	1.5	1.2	0.0	3.8	0.0	0.3

Table 1.8: Considered demand side management processes

Sector	Processes
Industry	aluminum-electrolysis, cement mills, paper machine, pulp refining, paper coating / calendaring, recycled paper treatment, electric arc furnace, chlorine-alkali-electrolysis (membrane), ventilation, compressed air
Service	medium and large water heaters (>30 l), air conditioning, ventilation, cold storage houses, walk-ins / chillers / freezers
Domestic	refrigerator, freezer, washing machine, dryer, dish washer, medium and large water heaters (>30 l), air conditioning, night storage heating, circulation pumps
Transport	e-mobility
Municipal	pumping, aeration
Others	heat pumps

Table 1.9: Technical specifications for demand side management processes

Technologies	Balancing interval [h]	Efficiency [%]	Max. demand reduction [%]	Max. demand increase [%]
ventilation, compressed air, circulation pumps, heat pumps, air conditioning	2	95	24-90	75-90
medium and large water heaters (>30 l), cold storage houses, freezer, pumping	4	95	90	50-90
dish washer	12	100	90	90
washing machine, dryer, night storage heating, e-mobility, aeration	24	100	25-90	25-90
aluminium-electrolysis, cement mills, paper machine, paper coating / calendaring, pulp refining, recycled paper treatment, electric arc furnace, chlorine-alkali-electrolysis (membrane)	8760	100	15-90	50

Table 1.10: Development of DSM-capacities in Europe until 2050 [MW]

	2020		2030		2040		2050	
	Developed	**Existing**	**Developed**	**Existing**	**Developed**	**Existing**	**Developed**	**Existing**
Industry	11,268	15,565	13,226	16,111	14,531	16,657	15,996	17,254
Service	681	13,619	2,151	14,338	4,269	14,229	7,167	14,334
Domestic	1,891	63,020	6,177	61,772	11,455	57,273	18,648	53,281
Transport	589	589	1,342	1,342	3,019	3,019	5,717	5,717
Municipal	106	1,056	310	1,035	558	1,014	745	994
Other	173	3,466	494	3,292	941	3,137	1,499	2,998

Table 1.11: Overnight investment costs [EUR_{2010}/kW]

	2020	2030	2040	2050
Nuclear	3,157	3,157	3,157	3,157
Lignite	1,850	1,850	1,850	1,850
Lignite CHP	2,350	2,350	2,350	2,350
Lignite CCS	-	2,896	2,721	2,652
Lignite CCS (flexible)	-	3,041	2,842	2,764
Lignite - innovative	1,950	1,950	1,950	1,950
Lignite - innovative CCS	-	2,996	2,821	2,752
Lignite - innovative CCS (flexible)	-	3,145	2,945	2,867
Lignite - innovative CHP and CCS	-	3,396	3,221	3,152
Hard coal	1,500	1,500	1,500	1,500
Hard coal CHP	2,650	2,342	2,135	2,030
Hard coal CCS	-	2,349	2,207	2,152
Hard coal CCS (flexible)	-	2,459	2,298	2,236
Hard coal - innovative	2,250	1,904	1,736	1,650
Hard coal - innovative CCS	-	2,753	2,443	2,302
Hard coal - innovative CCS (flexible)	-	2,894	2,560	2,410
Hard coal - innovative CHP and CCS	-	3,191	2,842	2,682
CCGT	700	700	700	700
CCGT - CHP	1,000	1,000	1,000	1,000
CCGT - CCS	-	1,127	1,057	1, 030
CCGT - CCS (flexible)	-	1,189	1,109	1,078
CCGT - CHP and CCS	-	1,409	1,341	1,314
OCGT	400	400	400	400
Compressed air storage	850	850	850	850
Biomass gas	2,398	2,395	2,393	2,390
Biomass gas CHP	2,597	2,595	2,592	2,590
Biomass liquid	1,700	1,700	1,700	1,700
Biomass solid	3,297	3,293	3,290	3,287
Biomass solid CHP	3,497	3,493	3,490	3,486
Concentrated solar power	3,989	3,429	3,102	2,805
Geothermal (hot dry rock)	10,504	9,500	9,035	9,026
Geothermal (high enthalpy)	1,050	950	904	903
PV ground	1,796	1,394	1,261	1,199
PV roof	2,096	1,627	1,471	1,399
Wind onshore	1,221	1,161	1,104	1,103
Wind offshore	2,615	2,365	2,249	2,247

Table 1.12: Techno-economic figures for generation technologies [as indicated]

	Net efficiency [%]	Availability [%]	FOM costs [EUR_{2010}/kWa]	Lifetime load[a]	Minimum [%]	Ramp-up times [h]
Nuclear	33	84.5	96.6	60	45	48
Lignite	43	86.3	43.1	45	30	3 - 12
CHP	22.5	86.3	62.1	45	30	3 - 12
CCS	33.5	86.3	70.3	45	30	3 - 12
CCS (flexible)	32.9	86.3	71.6	45	30	3 - 12
innovative	46.5	86.3	43.1	45	30	3 - 12
innovative CCS	37	86.3	70.3	45	30	3 - 12
innovative CCS (flexible)	36.4	86.3	71.6	45	30	3 - 12
innovative CHP and CCS	20	86.3	89.3	45	30	3 - 12
Hard coal	46	83.8	36.1	45	30	1 - 6
CHP	22.5	83.8	55.1	45	30	1 - 6
CCS	36.5	83.8	59	45	30	1 - 6
CCS (flexible)	35.9	83.8	60.2	45	30	1 - 6
innovative	50	83.8	36.1	45	30	1 - 6
innovative CCS	40.5	83.8	59	45	30	1 - 6
innovative CCS (flexible)	39.9	83.8	60.2	45	30	1 - 6
innovative CHP and CCS	20	83.8	78	45	30	1 - 6
CCGT	60	84.5	28.2	30	40	0.75 - 3
CHP	36	84.5	40	30	40	0.75 - 3
CCS	52	84.5	46	30	40	0.75 - 3
CCS (flexible)	51.6	84.5	50.5	30	40	0.75 - 3
CHP and CCS	33	84.5	57.9	30	40	0.75 - 3
OCGT	40	84.5	17.2	25	20	0.25
Biomass gas	40	85	120	30	30	
Biomass gas CHP	30	85	130	30	30	
Biomass liquid	30	85	85	30	30	
Biomass solid	30	85	165	30	30	
Biomass solid CHP	22.5	85	175	30	30	
Concentrated solar power	-	-	120	25		
Geothermal (HDR)	22.5	85	300	30		
Geothermal	22.5	85	30	30		
PV ground	-	-	30	25		
PV roof	-	-	35	25		
Run-off-river hydropower	-	-	11.5	100		
Wind onshore	-	-	41	25		
Wind offshore	-	-	128	25		

TABLE 1.13: Fuel costs [EUR_{2010}/MWh_{th}]

	2008	2020	2030	2040	2050
Uranium	3.6	3.3	3.3	3.3	3.3
Lignite	1.4	1.4	1.4	1.4	1.4
Hard coal	17.3	13.4	13.8	14.3	14.7
Oil	44.6	99	110	114	116
Natural gas	25.2	28.1	31.3	33.2	35.2
Hydrogen	-	46.7	47.4	48.2	48.9
Bioliquid	53.2 - 94.3	57.1 - 101.1	61.8 - 109.4	61.8 - 109.4	61.8 - 109.4
Biogas	0.1 - 70.0	0.1 - 67.2	0.1 - 72.9	0.1 - 78.8	0.1 - 85.1
Biosolid	15.0 - 27.7	15.7 - 34.9	16.7 - 35.1	17.7 - 35.5	18.8 - 37.5

1.6.2 Appendix B: Detailed scenario results

Please note that the numbers for storage contain pump storage as well as Compressed Air Energy Storage (CAES).

TABLE 1.14: Gross installed capacities in Europe in [GW] (%)

		Scenario A				**Scenario B**			
	2008	**2020**	**2030**	**2040**	**2050**	**2020**	**2030**	**2040**	**2050**
Nuclear	135 (17)	109 (10)	95 (7)	69 (4)	60 (3)	109 (10)	100 (7)	88 (5)	75 (4)
Lignite	51 (7)	45 (4)	29 (2)	15 (1)	9 (0)	36 (3)	20 (1)	5 (0)	0 (0)
Lignite-CHP	0 (0)	6 (1)	3 (0)	1 (0)	0 (0)	6 (1)	3 (0)	1 (0)	0 (0)
Lignite-CCS	0 (0)	0 (0)	32 (2)	48 (3)	48 (2)	0 (0)	42 (3)	46 (3)	46 (2)
Lignite-CHP-CCS	0 (0)	0 (0)	3 (0)	4 (0)	4 (0)	0 (0)	4 (0)	4 (0)	4 (0)
Coal	128 (16)	83 (8)	19 (1)	3 (0)	0 (0)	82 (7)	19 (1)	3 (0)	0 (0)
Coal-CHP	0 (0)	50 (5)	39 (3)	38 (2)	32 (2)	40 (4)	23 (2)	14 (1)	7 (0)
Coal-CCS	0 (0)	2 (0)	17 (1)	27 (2)	27 (1)	2 (0)	26 (2)	32 (2)	32 (2)
Coal-CHP-CCS	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Gas	158 (20)	177 (16)	239 (17)	296 (17)	328 (16)	189 (17)	232 (17)	285 (16)	324 (15)
Gas-CHP	0 (0)	37 (3)	13 (1)	0 (0)	0 (0)	38 (3)	13 (1)	0 (0)	0 (0)
Gas-CCS	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Gas-CHP-CCS	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	9 (1)	12 (1)	12 (1)
Oil	72 (9)	19 (2)	5 (0)	0 (0)	0 (0)	19 (2)	5 (0)	0 (0)	0 (0)
Oil-CHP	0 (0)	5 (0)	0 (0)	0 (0)	0 (0)	5 (0)	0 (0)	0 (0)	0 (0)
Storage	50 (6)	47 (4)	52 (4)	57 (3)	64 (3)	47 (4)	50 (4)	64 (4)	68 (3)
Hydro	92 (12)	172 (16)	172 (12)	172 (10)	172 (8)	172 (16)	172 (12)	172 (10)	172 (8)
Biomass	9 (1)	29 (3)	39 (3)	55 (3)	79 (4)	29 (3)	39 (3)	55 (3)	79 (4)
Biomass-CHP	0 (0)	12 (1)	17 (1)	24 (1)	35 (2)	12 (1)	17 (1)	24 (1)	35 (2)
Wind onshore	49 (6)	160 (14)	284 (21)	368 (21)	449 (21)	160 (15)	284 (21)	368 (21)	449 (21)
Wind offshore	1 (0)	51 (5)	123 (9)	218 (13)	339 (16)	51 (5)	123 (9)	218 (13)	339 (16)
PV	5 (1)	82 (7)	138 (10)	222 (13)	305 (14)	82 (7)	138 (10)	222 (13)	305 (15)
CSP	0 (0)	7 (1)	38 (3)	91 (5)	127 (6)	7 (1)	38 (3)	91 (5)	127 (6)
Geothermal	1 (0)	2 (0)	11 (1)	13 (1)	15 (1)	2 (0)	11 (1)	13 (1)	15 (1)
Others	26 (3)	11 (1)	11 (1)	11 (1)	11 (1)	11 (1)	11 (1)	11 (1)	11 (1)

Table 1.15: Gross electricity generation in Europe [TWh] (%)

		Scenario A				Scenario B			
	2008	**2020**	**2030**	**2040**	**2050**	**2020**	**2030**	**2040**	**2050**
Nuclear	955 (28)	794 (21)	665 (16)	442 (10)	321 (7)	799 (21)	695 (17)	552 (12)	409 (9)
Lignite	315 (9)	310 (8)	130 (3)	64 (1)	37 (1)	251 (7)	16 (0)	4 (0)	0 (0)
Lignite-CHP	0 (0)	7 (0)	5 (0)	1 (0)	0 (0)	9 (0)	0 (0)	0 (0)	0 (0)
Lignite-CCS	0 (0)	2 (0)	180 (4)	231 (5)	208 (4)	2 (0)	291 (7)	284 (6)	246 (5)
Lignite-CHP-CCS	0 (0)	0 (0)	19 (0)	24 (1)	25 (1)	0 (0)	25 (1)	25 (1)	25 (1)
Coal	543 (16)	407 (11)	104 (3)	12 (0)	1 (0)	308 (8)	57 (1)	5 (0)	0 (0)
Coal-CHP	0 (0)	290 (8)	252 (6)	232 (5)	167 (4)	207 (6)	145 (4)	47 (1)	16 (0)
Coal-CCS	0 (0)	15 (0)	120 (3)	157 (4)	120 (3)	15 (0)	186 (5)	200 (4)	154 (3)
Coal-CHP-CCS	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Gas	747 (22)	512 (14)	444 (11)	308 (7)	223 (5)	643 (17)	401 (10)	275 (6)	194 (4)
Gas-CHP	0 (0)	111 (3)	35 (1)	0 (0)	0 (0)	174 (5)	66 (2)	0 (0)	0 (0)
Gas-CCS	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Gas-CHP-CCS	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	60 (1)	61 (1)	44 (1)
Oil	91 (3)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Oil-CHP	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Storage	67 (2)	17 (0)	36 (1)	55 (1)	75 (2)	15 (0)	36 (1)	62 (1)	77 (2)
Hydro	466 (13)	497 (13)	491 (12)	487 (11)	464 (10)	497 (13)	492 (12)	487 (11)	462 (10)
Biomass	88 (3)	26 (1)	66 (2)	92 (2)	99 (2)	44 (1)	89 (2)	110 (2)	116 (2)
Biomass-CHP	0 (0)	67 (2)	94 (2)	117 (3)	134 (3)	78 (2)	101 (2)	125 (3)	138 (3)
Wind onshore	117 (3)	335 (9)	599 (15)	728 (16)	821 (17)	335 (9)	599 (15)	732 (16)	823 (17)
Wind offhore	0 (0)	180 (5)	444 (11)	751 (17)	1107 (23)	180 (5)	444 (11)	753 (17)	1110 (23)
PV	7 (0)	86 (2)	152 (4)	247 (6)	335 (7)	86 (2)	152 (4)	248 (6)	334 (7)
CSP	0 (0)	26 (1)	138 (3)	319 (7)	427 (9)	26 (1)	138 (3)	319 (7)	427 (9)
Geothermal	6 (0)	10 (0)	58 (1)	68 (2)	75 (2)	10 (0)	58 (1)	68 (2)	74 (2)
Others	67 (2)	57 (2)	57 (1)	57 (1)	57 (1)	57 (2)	57 (1)	57 (1)	57 (1)
DSM	(0)	14 (0)	22 (1)	32 (1)	49 (1)	14 (0)	22 (1)	33 (1)	50 (1)

Table 1.16: Renewable curtailment [TWh] (%)

		2020	2030	2040	2050
Scenario A	Wind onshore	0.7 (0.2)	5.0 (0.8)	44.3 (6.1)	103.0 (12.6)
	Wind offshore	0.0 (0.0)	1.7 (0.4)	15.3 (2.0)	46.7 (4.2)
	Solar power	0.1 (0.1)	0.2 (0.1)	2.6 (1.1)	10.0 (3.0)
Scenario B	Wind onshore	0.7 (0.2)	5.4 (0.9)	40.1 (5.5)	101.1 (12.3)
	Wind offshore	0.0 (0.0)	1.4 (0.3)	13.3 (1.8)	43.9 (4.0)
	Solar power	0.1 (0.1)	0.3 (0.2)	2.4 (1.0)	10.2 (3.0)

On the interaction effects of market failure and capacity payments in interconnected electricity markets

The ongoing debate on the necessity of capacity remuneration mechanisms (CRMs) to ensure sufficiency of generation capacity primarily focuses on a national perspective. The research concerning possible spill-over effects, positive or negative, in adjacent markets is lagging behind. This is the case for the effects of CRMs as well as for the effects of market failures. We address both topics in this paper.

Specifically, we analyse the effects of price caps in two interconnected markets. Additionally, we analyse the effects of capacity payments meant to counter the deadweight losses triggered by the price restrictions.

Although we find no indication that price caps or capacity payments in one market have (serious) negative effects on neighbouring markets, being connected to other markets can worsen the deadweight losses induced by market inefficiencies. Also capacity mechanisms might be less effective than in isolated markets. Finally, in the analysed set-up we find no indication that capacity efforts in one market support neighbouring markets with insufficient generation capacity.

2.1 Introduction

The question about whether or not capacity remuneration mechanisms (CRMs) are necessary to ensure security of supply and maximise welfare as well as their concrete design has been widely discussed for many years.[22] The prevailing view is that energy only markets do not provide sufficient investment incentives as wholesale electricity markets are characterised by certain market failures (and regulatory inefficiencies) that render some form of external investment incentive necessary. The most dominant arguments in favour of market interventions are the insufficient elasticity of demand in times of scarcity of supply, regulatory interventions preventing sufficient price signals and risk-averse investment behaviour.

Although many countries in Europe have already implemented some form of capacity support scheme (or several, as a recent sector inquiry on capacity mechanisms by the European Commission (EC) points out)[23] the debate concerning the appropriate market design continues. The German government for example only recently renewed its commitment to free price formation in the electricity market – assuming that this would trigger sufficient investments. Nonetheless, a capacity reserve for unexpected events is planned.[24]

Whereas the prevalent literature on capacity remuneration mechanisms mainly focuses on single markets, against the sketched background of numerous and various regulations and market designs the question arises if and how markets with different capacity levels, market interventions and CRMs interact. Correspondingly, concerns have been raised, e. g., by the EC, that CRMs may distort price signals in domestic as well as foreign markets which could have a negative impact on investment decisions (see European Commission, 2016a) and potentially endanger the functioning of the European internal market (see ACER, 2013, European Commission, 2013).

On the other hand, national regulators pursue market interventions to support capacity investments with diverse mechanisms and varying intensity. That is why suspicions aroused that some markets may free-ride on the efforts of others.

Given the outlined background and concerns we address the following research questions: First, do market failures in isolated markets have different effects than in interconnected markets? Additionally, what effect does an insufficient level of capacity in one market,

[22]See, e. g., Borenstein and Holland (2005), Cramton and Ockenfels (2012), Cramton and Stoft (2005, 2006), Hogan (2005), Joskow and Tirole (2007), Joskow (2006, 2008). A recent overview is provided by Cramton et al. (2013).

[23]See European Commission (2016a)

[24]As stated in the electricity market law 'Gesetz zur Weiterentwicklung des Strommarktes (Strommarktgesetz)', 26th of July 2016, BGBl. 2016 I, nr. 37, p. 1786, 1796-1797.

caused by a malfunctioning energy only market, have on adjacent markets? Second, do CRMs, in our case capacity payments (CP), behave differently in interconnected than in isolated markets? And third, do markets with insufficient domestic capacity incentives benefit from CRMs in neighbouring markets, i. e., do incentives to free-ride exist?

The existing literature about CRMs is mostly focusing on a national perspective. This is the case for the negative welfare effects of market failures rendering CRMs necessary in the first place as well as the efficiency and effectiveness of mechanisms designed to counter those effects.

Rare examples analysing the interaction of capacity mechanisms in adjacent countries are Cepeda and Finon (2011), Elberg (2014) and Meyer and Gore (2014). Cepeda and Finon (2011) deploy a system dynamics model to simulate varying market designs in two markets. They conclude that a CRM in one market can result in negative externalities by impeding the performance of adjacent energy only markets. They state that harmonised approaches (whether energy only markets or CRMs) are mutually beneficial. Elberg (2014) as well as Meyer and Gore (2014) consider the presence of market failure as given and therefore focus on the comparison of mechanisms. The former concludes, using an analytical model, that capacity payments, although equally efficient as a strategic reserve in isolated markets, are a dominant strategy in interconnected markets. Meyer and Gore (2014), using a simulation model, conclude that unilateral capacity mechanisms can have negative cross-border effects worsening the problem of insufficient investments in neighbouring markets. However, if both regions introduce a mechanism, welfare is increased in the given example.

In this paper we contribute to the previous research by analysing the effects of market failure and CRMs in interconnected markets using a equilibrium model. Thus, we can identify market equilibria rather than relying on simulation results as in Cepeda and Finon (2011). Additionally, we model the complex interaction of adjacent electricity market using real world data (in contrast to Elberg (2014) and Meyer and Gore (2014) who assume equal demand at all times). This way, a more realistic estimation of the magnitude of effects can be made. Our results contradict the previous literature as our findings do not indicate any serious negative cross-border effects of CRMs.

For the analysis of cross-border effects in electricity markets the joint distribution of demand as well as the generation of renewable energies are critical. This complex and market-specific dependency renders a theoretical analysis difficult. We thus develop a mathematical model, more specifically a Mixed Complementarity Problem (MCP), to calculate equilibrium market outcomes. We use 2015 data for the demand and renewable generation patterns for Germany and France. As generators we take into account one base and one peak load technology. Market failure is represented by a cap on electricity prices. Such a cap does not necessarily require an explicit (regulatory) cap but may also be inflicted on the market by undue market interventions meant to secure security of supply. To analyse the effects of CRMs, we focus on capacity payments as one possible alternative.[25] We assume that those payments are made for all (conventional) capacity in the market and paid as a lump sum.

The main findings are as follows: First, and in line with the existing literature, our results show that – in an isolated market – the disadvantages of insufficient price signals induced by a cap can be cured by capacity payments (leaving problems concerning the determination of their appropriate level aside).

Second, the negative implications of price caps worsen in interconnected markets: Installed capacity and ultimately welfare decrease to a larger extent. The artificially low prices result in additional exports during peak demand and thus do not (fully) benefit domestic consumers. In addition, more capacity recedes as imports partly replace the missing domestic generation which hinders prices to reach the equilibrium level. A higher interconnector capacity worsens the negative effects of price caps.

The third finding is that a price cap in one market does not appear to have a negative impact on welfare in neighbouring markets. On the contrary, the price cap results in imports from the market having the cap which benefits the market without (or a higher) cap. These benefits exceed the disadvantages of decreasing generation capacity due to the additional imports that are dampening market prices.

Forth, capacity payments are less efficient in interconnected than in isolated markets. This is due to price distortions that are not targeted by CP: Inefficient price signals persist and result in inefficient trade. But although some welfare losses persist, our results indicate that most of the losses can be recuperated.

Finally, capacity payments in one market do not appear to greatly support neighbouring markets that have a price cap: Losses might be slightly reduced given less restrictive caps. For more restrictive caps CP might even be harmful for adjacent markets. The additional capacity triggered by the support scheme in one market pushes back capacity in the other market which outweighs the benefits of imports. If neighbouring markets both have a price cap, the optimal welfare level can almost be restored if both markets introduce capacity payments. Remaining deadweight losses are due to inefficient price signals caused by the price caps which result in an inefficient allocation of capacity and supply.

[25] Assuming an omniscient regulator, capacity payments and capacity auctions are equivalent with respect to the resulting market outcome.

A few things have to be mentioned concerning our findings: We assume that the markets and therefore also trade are not limited or regulated other than by – depending on the scenario – price caps. This means that exports are not restricted when the price cap is reached. If caps are explicitly intended in market regulations exports might be limited in order to strengthen domestic consumption. But as this would interfere with the free trade in the European internal market we refrain from considering this option in the analysis at hand.

The second assumption underlying the analyses is that demand is flexible: A power outage is not a problem caused by (foreseeable) imbalances of demand and supply. In fact, over the last five years no reliability problems occurred in continental Europe in the ten EU member countries mentioned in the EU Commission's sector inquiry – although some expect problems in the future.[26] Assuming that demand is actually (at least partially) inflexible could result in an inability of the market to match demand and supply. In this case the European transmission system operators (TSOs) are required to execute countermeasures to restore system stability – including (involuntary) load shedding as a last resort. This emergency measures might cause additional welfare losses which are not taken into account in this paper. Concerning our findings the general trend remains unchanged but may have a greater order of magnitude. Also the effects on neighbouring countries would only slightly change as they are driven by a change in trade flows rather than the influence on capacity.

The remainder of this paper is structured as follows: In section 2.2 the mathematical programme for the subsequent analyses is presented. Section 2.3 first outlines the sample data followed by the quantitative analyses of the effects of market failure and capacity payments. Section 2.4 concludes.

2.2 Market model

2.2.1 Model formulation

The model represents interconnected electricity markets m.[27] Markets are regulatory independent regions connected by limited transmission capacity. We analyse investments $y_{m,n}$ in production technology n with variable costs v_n and annual investment (plus other fixed) costs k_n. In the context of this paper, technologies always refer to conventional power plants; renewable energies are treated separately. A single investment period is followed by several production periods t with production $x_{m,n}^t$ in each time period.

[26] See European Commission (2016a).

[27] The terms market, region and country are used interchangeably throughout this paper.

Production of renewable energies (wind and photovoltaics) r_m^t is curtailed if production exceeds demand. We model each year individually, i. e., we do not account for investments over time. Investment and production decisions are made simultaneously. Perfect competition of generators is assumed, i. e., generation capacity is utilised until market prices equal variable costs or capacity is exhausted. Investment in generation capacity takes place until profits are zero. Countries are connected by a transmission line with limited capacity $\bar{l}_{m,-m}$. Actual trade is denoted by $l_{m,-m}$ and indicates the electricity trade from region m to another region $-m$. Transmission capacity is utilised to maximise total welfare, i. e., until prices in both markets are equal or the capacity is exhausted.

Market interventions in terms of capacity mechanisms take place prior to investment and generation decisions. They influence investment costs via capacity payments c_m which are subtracted from the investment costs.

The basic notations are listed in Table 2.1, additional symbols are explained where applicable.

Table 2.1: Model sets, parameters and variables

Abbreviation		Description
Model sets		
m		Markets
$n \in N$		Technologies
$t \in T$		Time index
Model parameters		
k_n	$\mathbb{R}^+$	Marginal investment costs
v_n	$\mathbb{R}_0^+$	Variable production costs
a_m^t	$\mathbb{R}_0^+$	y-intercept of the demand curve
b	$\mathbb{R}^+$	Linear slope of the demand curve
K_r	$\mathbb{R}_0^+$	Total investment costs of wind and photovoltaics
r_m^t	$\mathbb{R}_0^+$	Wind and photovoltaics production
$\bar{l}_{m,-m}$	$\mathbb{R}_0^+$	Transmission capacity
Model variables		
$y_{m,n}$	$\mathbb{R}_0^+$	Capacity investments
$x_{m,n}^t$	$\mathbb{R}_0^+$	Production
p_m^t	$\mathbb{R}_0^+$	Market price
c_m	$\mathbb{R}_0^+$	Capacity payments per unit of capacity
$l_{m,-m}^t$	$\mathbb{R}_{[-\bar{l},\bar{l}]}$	Trade

In the following we present the mathematical programme that is used to solve the model sketched above. We start with the equilibrium model representing the competitive market. Afterwards, we present the calculation of welfare based on the results of this model.

We start by formulating the market outcome as an optimisation problem which is separated into investment and production decisions of a generation company and trade conducted by a transmission system operator (TSO).

The objective function of the generation company represents a profit maximising investor who maximises revenues minus variable and investment costs. The optimisation problem is as follows:

$$\max_{x^t_{m,n}, y_{m,n}} \sum_{t,m,n} (p^t_m - v_{m,n}) x^t_{m,n} - \sum_{m,n} y_{m,n}(k_n - c_m) \tag{2.1}$$

subject to

$$y_{m,n} - x^t_{m,n} \geq 0 \; (\lambda^t_{m,n}) \tag{2.2}$$

$$p^t_m = a^t_m - b \left(\sum_n x^t_{m,n} - \sum_{-m} (l^t_{m,-m} - l^t_{-m,m}) + r^t_m \right) \tag{2.3}$$

$$y_{m,n} \geq 0, \; x^t_{m,n} \geq 0. \tag{2.4}$$

The first constraint limits the hourly electricity generation to the installed capacity. Equation (2.3) represents the assumed linear demand function. Finally, non-negativity constraints are listed. Related dual variables are listed in brackets next to each equation. The influence of capacity mechanisms is represented by capacity payments c_m. This optimisation problem for competitive generators is equivalent to the following Karush-Kuhn-Tucker (KKT) conditions:[28]

$$v_{m,n} + \lambda^t_{m,n} - p^t_m \geq 0 \tag{2.5}$$

$$k_n - c_m - \sum_t \lambda^t_{m,n} = 0 \tag{2.6}$$

$$0 \leq y_{m,n} - x^t_{m,n} \perp \lambda^t_{m,n} \geq 0 \tag{2.7}$$

$$(2.3) - (2.4) \tag{2.8}$$

with the perp operator ($\perp$) meaning that the product of the expressions to the left and to the right has to equal zero. The first equation reflects the first order condition for the optimal choice of production and links the market price, variable production costs and the marginal value of capacity.

The second equation states the first order condition for investments: Annual investment costs for generation companies, i. e., nominal investment costs k minus capacity payments c, have to equal the sum of the marginal values of capacity λ. Equation (2.7) limits production to installed capacity and links the associated dual variable.

[28] As we assume competitive generators the conjectural variation parameter is set to zero (see, e. g., Murphy and Smeers, 2005).

We assume a profit maximising, competitive TSO. The objective function of the TSO consists of the profits from trade calculated as the price difference between both regions times the trade volume:

$$\max_{l^t_{m,-m}} \sum_{t,m,-m} l^t_{m,-m}(p^t_{-m} - p^t_m) \tag{2.9}$$

subject to:

$$\bar{l}_{m,-m} - l^t_{m,-m} \geq 0 \; (\pi^t_{m,-m}) \tag{2.10}$$
$$\bar{l}_{m,-m} + l^t_{m,-m} \geq 0 \; (\mu^t_{m,-m}) \tag{2.11}$$
$$(2.3) \tag{2.12}$$

The first and second constraint represent the upper and lower bound of transmission between regions. The linear demand function is equal to the one for the generation company.

This optimisation problem is equivalent to the following KKT conditions:

$$p^t_m - p^t_{-m} + \pi^t_{m,-m} - \mu^t_{m,-m} = 0 \tag{2.13}$$
$$0 \leq \bar{l}_{m,-m} - l^t_{m,-m} \perp \pi^t_{m,-m} \geq 0 \tag{2.14}$$
$$0 \leq \bar{l}_{m,-m} + l^t_{m,-m} \perp \mu^t_{m,-m} \geq 0 \tag{2.15}$$
$$(2.3) \tag{2.16}$$

The first order condition (2.13) postulates that the price difference between two regions equals the marginal value of the transmission capacity. The following two complementarity conditions (2.14) and (2.15) enforce the transmission limits and link the associated dual variables.

Simultaneously solving equations (2.5) - (2.8) and (2.13) - (2.15) provides the competitive market outcome. Due to the quasi-concave objective function and the convexity of restrictions, the KKT conditions are necessary and sufficient for an optimal solution.

A cap on electricity prices is implemented as an additional generation technology $cap \in N$ with zero investment costs ($k_{cap} = 0$) and variable costs v_{cap} equal to the desired cap.

2.2.2 Welfare calculation

Total welfare W in each region consists of the sum of producer surplus for conventional power plants PS, consumer surplus CS and congestion rent CoR minus the costs for the capacity mechanism CM plus the net revenues of renewable energies RES (see equation

2.17). As the markets are assumed to be competitive, producers invest until profits are zero.

$$W_m = PS_m + CS_m + CoR_m - CM_m + RES_m \tag{2.17}$$

The calculation of consumer surplus is presented in equation 2.18: The first part represents the surplus if all generation was provided by regular capacity. The second part corrects this value by deducing the generation by capacity representing the price cap (x_{cap}).[29]

$$CS_m = \sum_t \left(\frac{(a_m^t - p_m^t)^2}{2\ b} - \frac{b\ (x_{m,cap}^t)^2}{2} \right) \tag{2.18}$$

The congestion rent is assumed to be split evenly between trading countries (see equation 2.19). The costs of the capacity mechanism, the capacity payments, simply consist of the fixed capacity payments per unit of capacity c times the installed capacity (see equation 2.20).

Net revenues of renewable energies are calculated as revenues minus investment costs K_r (see equation 2.21). The investments in renewable energies are exogenous and constant for all scenarios. Thus, they cancel out in the delta analyses.[30]

$$CoR_m = \frac{1}{2} \sum_{t,-m} \left(l_{m,-m}^t (p_{-m}^t - p_m^t) + l_{-m,m}^t (p_m^t - p_{-m}^t) \right) \tag{2.19}$$

$$CM_m = \ y_{m,n}\ c_m \tag{2.20}$$

$$RES_m = \sum_t r_m^t\ p_m^t - K_r \tag{2.21}$$

2.3 Quantitative analysis

2.3.1 Data

We apply our model to the case of Germany and France. We use projections concerning demand and the installed capacity of renewable energies for the year 2020 based on the EU Reference Scenario 2016 by the European Commission.[31] According to the EC's

[29] If neighbouring markets have the same price cap the hourly market result may be ambiguous. If both market prices actually reach the cap we assume that capacity is used to serve domestic demand first – this way the outcome is unique.

[30] Given today's investment costs the net revenues of renewables are rather negative than positive and thus affect the welfare negatively. However, using the net revenues is more convenient in the context of this paper as additional revenues result in an increase of welfare.

[31] See European Commission (2016b) and the data provided in Appendix A.

sector inquiry on capacity mechanisms from 2016, European TSOs expect reliability issues in the upcoming five years.

Annual demand is set to 530 TWh and 452 TWh for Germany and France respectively according to the 2020 value of the EU Reference Scenario. The demand structure is based on hourly values and corresponding prices for 2015 based on ENTSO-E (2016) and the European Energy Exchange (EEX)[32] respectively. The demand and other hourly input values are scaled with a constant factor for each region to meet the scenario assumptions concerning the yearly values.

A variety of estimates for the elasticity (denoted as η) of demand exists in the literature (see, for example, Lijesen (2007), Knaut and Paulus (2016) and the literature overview provided therein). Time-of-use elasticity, i. e., real-time or wholesale price elasticity, is mostly estimated to have very low values ranging between -0.002 and -0.16, depending on the date and time. In the base case we assume an elasticity of -0.01, which is rather at the lower end of estimated values.[33] We do this as the effects of price caps are mostly relevant during high prices which in turn are more likely if demand elasticity is low.

The electricity generation structure of wind and photovoltaics is based on data provided by the French and German TSOs for 2015.[34]

Modelled residual load, i. e., demand minus electricity generation by renewable energies, is illustrated in figure 2.1. The average residual load is about 40 and 44 GW for Germany and France respectively; the correlation coefficient between residual load levels is 0.62 with an average absolute difference of about 10.3 GW.

To limit the computational burden, we only include the first out of every four weeks of the whole year. Investment costs are scaled accordingly.

As conventional generation technologies we take into account one base and one peak load technology. Overnight investment costs are assumed to be 1,500 €/kW and 500 €/kW for the base and the peak load technology respectively with variable production costs of 25 €/MWh and 50 €/MWh.

To analyse the influence of the interconnector we vary the capacity using 3,000 MW or 6,000 MW.[35]

[32]See www.eex.com.

[33]As we use linear demand functions we are referring to point elasticities. The slope of the demand curve b is assumed to be constant and calculated as $\frac{1}{\eta}\frac{p^*}{d^*}$ with reference price $p^* = 100$ €/MWh and reference demand $d^* = 50$ GW (thus $b = 0.2$). We use hourly y-intercepts a that result in demand curves which meet historic demand and prices combinations.

[34]For France, see RTE (2016), for Germany, see Amprion (2016), 50Hertz (2016), TenneT (2016) and TransnetBW (2016).

[35]The Ten-Year Network Development Plan 2016 (version November 2015) assumes 3,000 MW for the cross-border capacity between DE and FR in 2020.

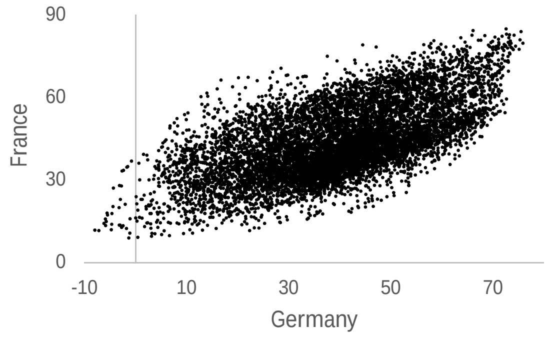


FIGURE 2.1: Modelled residual load in Germany and France [GW]

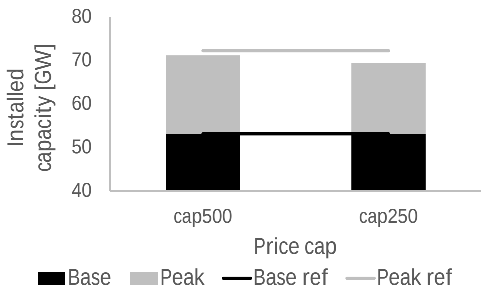
We take into account three levels of price caps: unrestricted prices, a cap at 500 €/MWh and a cap at 250 €/MWh.

2.3.2 Price caps and capacity payments in isolated markets

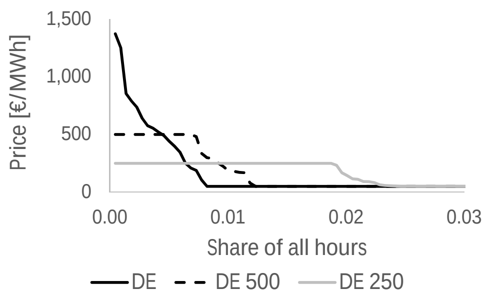
2.3.2.1 The effects of price caps

We start by analysing the effects of insufficient price signals and capacity payments in an isolated market. This forms the basis for the analysis of interconnected markets that follows afterwards. Here, we present the results for Germany. Similar effects can be observed for France (see Appendix B).

Figure 2.2 illustrates the effects of price caps on installed capacity and the price duration curve.



(A) Installed capacity



(B) Price duration curve (excerpt)

FIGURE 2.2: Effects of capacity payments in DE

Compared to the reference case, a price cap reduces the revenues of power plants in times of scarcity of supply. All technologies are affected to the same extent as the full capacity is utilised. Assuming capacity would be identical with and without a (binding) cap, this results in losses for the generators. Thus, capacity has to recede to restore an equilibrium state. As can be observed in Figure 2.2a, although a price cap affects the revenues of all capacities, exclusively the capacity of peak load plants decreases. The decrease of peak capacity results in additional revenues for all types of technology whereas a decrease in base capacity would benefit technologies unevenly. Thus, in order to restore an equilibrium state, only peak capacity recedes.

Compared to the reference case in which prices are not restricted, the peak capacity at a price cap of 500 €/MWh decreases by about 1.1 GW. In the case of a price cap of 250 €/MWh the capacity decrease amounts to about 2.8 GW.

The resulting effect on prices is illustrated by the excerpt from the price duration curve shown in Figure 2.2b. It shows the 3 % highest prices in descending order for varying price caps and the reference case. The price caps and the resulting capacity decrease result in an increasing number of hours with prices above variable costs of the peak load technology. The revenues in these hours balance the losses resulting from the missing peak prices. In total, the power plants' revenues per unit of capacity as well as the average market price $\tilde{p}$ remain constant. This follows directly from the generators' KKT conditions. From equation 2.5 follows:

$$\tilde{p} = \frac{\sum_t p^t}{T} = v_n + \frac{\sum_t \lambda_n^t}{T} \tag{2.22}$$

Adding equation 2.6 results in:

$$\tilde{p} = v_n + \frac{k_n - c}{T} \tag{2.23}$$

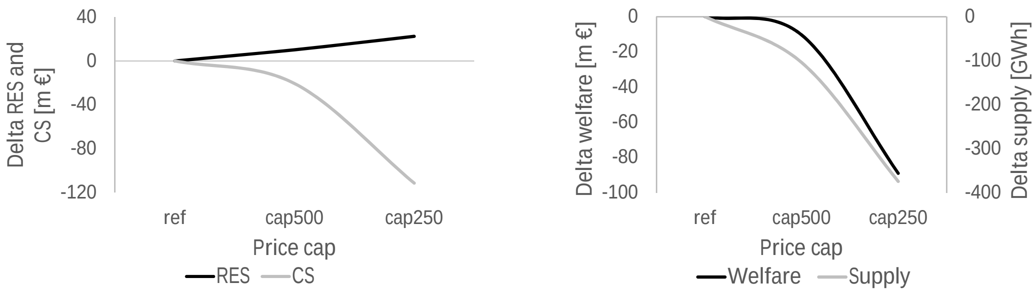
Thus, the average market price solely depends on the annual fixed and marginal generation costs as well as capacity payments. A price cap – and in the later scenarios trade with other markets – does not influence the average price. In the current scenario, base load plants have to recover their annual fixed costs which amount to about 86.2 €/kW.[36] With variable costs of 25 €/MWh this translates into an average market price of 34.8 €/MWh.

The effects on welfare and supply are illustrated in Figure 2.3. As seen before, the lower the price cap the more capacity recedes and the more often prices exceed the variable costs of peak plants. This benefits the revenues of renewable energies: Without a price cap, high prices only occur during a few hours of scarcity of supply. These hours do not coincide with renewable generation. With a cap and due to the receding conventional capacity, renewables also benefit from peak prices.

The gain in revenues is overcompensated by losses in consumer surplus: With decreasing generation capacity also the served demand decreases which in sum results in deadweight

[36] We assume an interest rate of 5 % and a technical lifetime of 25 and 40 years for the peak and base load technology respectively.

losses. Due to the exponential effect of an increasingly binding price cap on consumer surplus (see equation 2.18), welfare is decreasing more severely with a decreasing cap.



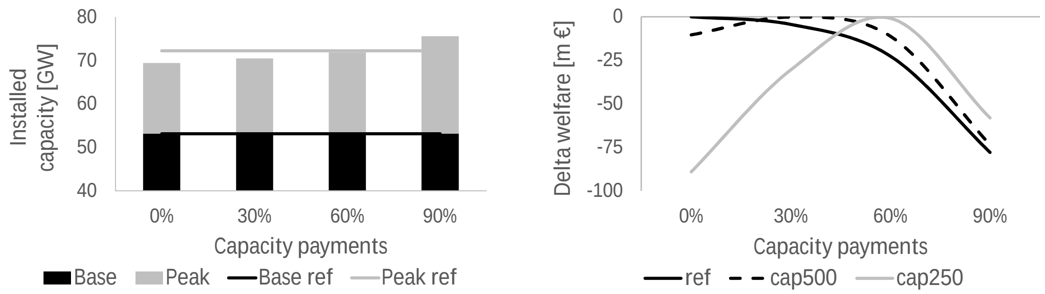
(A) Renewables' revenue and consumer surplus (B) Welfare and supply

FIGURE 2.3: Effects of price caps in DE

A price cap of 500 €/MWh has a comparably small effect on welfare – the effect of a cap of 250 €/MWh is about nine times higher.

2.3.2.2 The effects of capacity payments

Capacity payments are an option to support generation capacity and counter the negative effects of a price cap. Figure 2.4a illustrates the increase of capacity for the case of a price cap at 250 €/MWh. The capacity payments are indicated as the share of investment costs of the peak load technology.



(A) Installed capacity ($\bar{p} = 250$ €/MWh) (B) Welfare

FIGURE 2.4: Effects of capacity payments in DE

At about 60 % of investment costs the installed capacity reaches the same level as it does without price restrictions (indicated by the lines). Not only the total level of installed capacity is restored but also the capacity mix: Although all capacities receive the same subsidies per unit of capacity, only the peak load capacity increases. The economic intuition is similar to the previous argument explaining why only peak capacity recedes while all capacity is affected by a price cap: With CP short-run profits have to decrease in order to restore the zero profit condition of investment. In equilibrium the sum of reduced profits equals the capacity payments. If base load capacity increases, profits of base and peak capacities would be affected unevenly in times in which base capacity is setting the price. Whereas if peak capacity increases, all capacities are affected evenly by a decreasing price. Therefore, in order to compensate for the capacity payments to the same extent and thus reach an equilibrium state, only peak capacity increases.

At higher levels of CP the capacity exceeds those of the reference case. CP can fully compensate for the foregone revenues induced by a price cap and restore the optimal welfare (see Figure 2.4b). If the capacity support exceeds the optimal level, i. e., more than compensates for foregone revenues, the welfare drops again as this results in excess capacity.

Capacity payments also influence the market prices by increasing generation capacity: Prices are still limited by the price cap but due to the additional capacity, prices reach the cap less frequently (see Figure 2.5).

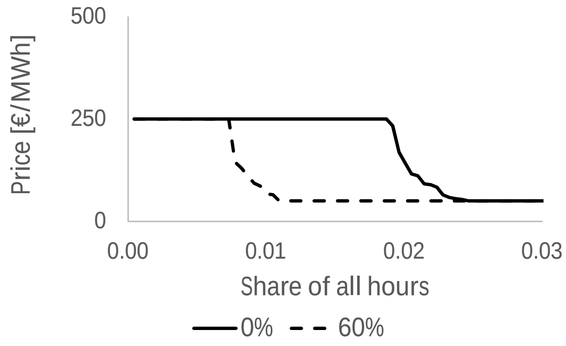


FIGURE 2.5: Price duration curves (excerpt) for varying CP ($\bar{p} = 250$ €/MWh)

As discussed earlier, average prices remain unchanged with price caps. The same is not true for capacity payments. As CP directly influence investment costs they also alter the equilibrium condition for investments: Less short run profits have to be earned in order to cover long run costs. This results in decreasing average market prices with increasing CP.

In summary, our results show that the negative effects of price caps can be countered by capacity payments in isolated markets. The optimal level of generation capacity as well as welfare are fully restored with an appropriate level of support.

2.3.3 Market failure and capacity payments in interconnected markets

2.3.3.1 Reference case

We start by presenting the market outcome of a base case with no price restrictions. This serves as a reference for the subsequent analyses. The main outcomes of the simulations are summarised in the following table:

TABLE 2.2: Results of the base case

	$\bar{l} = 3{,}000\,\text{MW}$		$\bar{l} = 6{,}000\,\text{MW}$	
	DE	FR	DE	FR
Installed capacity [GW]	Base: 52.6 Peak: 18.5	Base: 50.4 Peak: 22.9	Base: 51.6 Peak: 19.9	Base: 50.6 Peak: 22.0
Trade volume [TWh]	5.9		10.0	
Average interconnector utilisation [%]	22.5		19.0	
Average market price [€/MWh]	34.84	34.84	34.84	34.84
Maximum price [€/MWh]	1372	1372	1403	1403

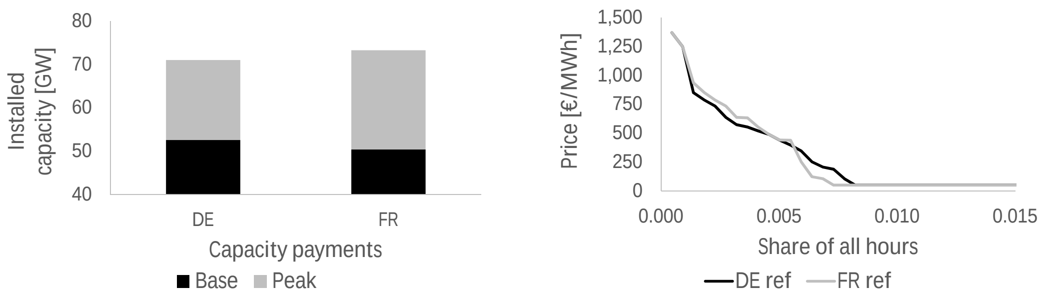
The total installed capacity in both markets decreases only slightly by about 300 MW with the higher trade capacity. Although the total base load capacity is reduced by about 800 MW, the peak capacity compensates some of the reduction by increasing by about 500 MW: Base load plants can be used more efficiently indicated by an increasing utilisation factor whereas peak capacity is needed to cover peak demand.

The average electricity price in both markets is identical – and independent from the interconnector capacity – at 34.8 €/MWh. This results directly from the equilibrium conditions of investment as discussed earlier: Average prices solely depend on investment and variable costs as well as capacity payment, but are independent from trade.

Figure 2.6 illustrates the generation capacity mix and price duration curve for peak load hours for a transmission capacity of 3,000 MW. The higher total but lower base load capacity in France compared to Germany indicates a steeper residual demand curve in France. The price duration curves for Germany and France for the 1 % highest prices (see Figure 2.6b) display a similar pattern for both markets, with a slightly steeper trajectory for France.

2.3.3.2 The effects of price caps

Market failures are often regarded as being harmful not only to domestic welfare but also as negatively affecting neighbouring markets. We investigate this hypothesis by simulating price caps in the model. We start by applying price caps to Germany whereas



(A) Installed capacity (B) Price duration curve (excerpt)

FIGURE 2.6: Installed capacity and price duration curve ($\bar{l} = 3{,}000\,\text{MW}$)

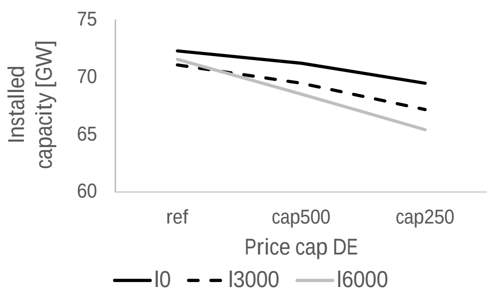
market prices in France are not restricted. Later on, we analyse the case of a cap only in France.

We start by looking at the effect of a price cap on capacity. The capacity of the peak technology in Germany recedes similar to the earlier case when Germany was assumed to be isolated. This time the decrease of capacity is stronger: Whereas a price cap of 250 €/MWh previously resulted in a decrease of about 2.8 GW this more than doubles with an interconnector capacity of 6,000 MW (see Figure 2.7a). This is driven by the price dampening effect of imports: In order to counter the profit loss induced by the price cap, capacity recedes to restore the previous price level. But unlike the single market case, imports partly replace the receding domestic capacity. In total, to restore the previous average prices, capacity has to decrease to a higher extent. Here, the same principle applies as in the isolated case: The average market price remains unchanged as long run investment costs have to be covered by short run prices.

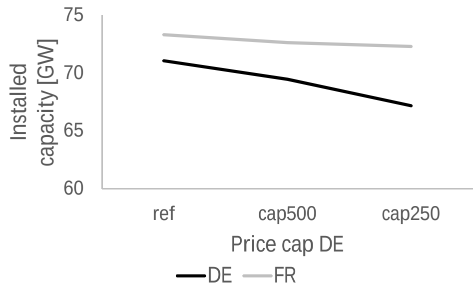
But although capacity is mainly affected in Germany, also capacity in France decreases (see Figure 2.7b): The price cap in Germany results in additional exports from Germany to France during peak demand. This decreases profits of power generators in France. Accordingly, also French capacity recedes in order to balance short run profits and investment costs.

The effect of a German price cap on prices in France is illustrated in Figure 2.8a. The German price cap also hinders price peaks in France. This ultimately forces French capacity to recede as it would otherwise be unprofitable. Although trade volumes remain fairly stable, the congestion rent increases (see Figure 2.8b). This is due to an increasing price difference triggered by the price cap.

Consumers in Germany are effected to a higher extend by a price cap if the German market interacts with its neighbour France (see Figure 2.9a). Similar to the isolated case, less demand is served in Germany. But this time, due to the on average higher prices in France, even less demand is served as electricity is exported during high demand. Thus,

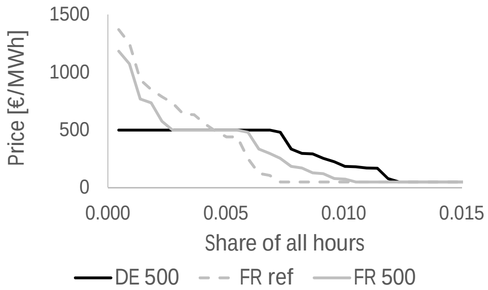


(A) Installed capacity in DE

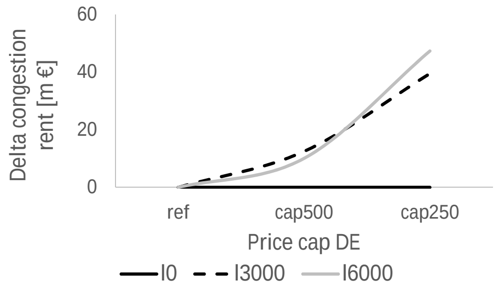


(B) Capacity in Germany and France ($\bar{l}$ = 3000)

Figure 2.7: Installed capacity with a price cap in Germany



(A) Price duration curves ($\bar{l} = 3,000$; excerpt)



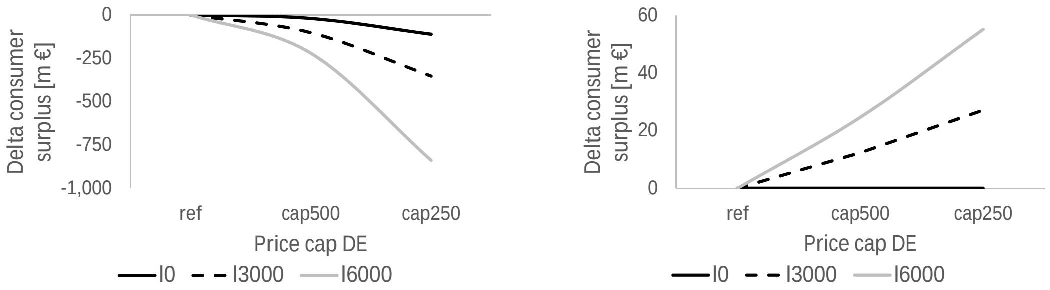
(B) Congestion rent

Figure 2.8: Price duration curves and congestion rent with a price cap in Germany

the benefits of lower prices during peak demand due to the cap are partly cancelled out by exports instead of domestic consumption. Given a price cap of 500 €/MWh this results in about five and eleven times higher losses in consumer surplus compared to Germany being isolated with an interconnector capacity of 3,000 MW and 6,000 MW respectively.

For France the effect on consumer surplus is positive, but less distinct (see Figure 2.9b). There are two opposing effects: On the one hand France benefits from imports attracted by higher domestic prices. On the other hand these imports force French capacity to recede due to their price decreasing effect (as discussed earlier). The sum of served demand does not change but the structure of supply does: During peak demand more demand is served whereas supply decreases during low demand (this is also reflected in the changing price price pattern as depicted in Figure 2.8a). In total, this increases the consumer surplus.

The higher the interconnector capacity the stronger the benefit for consumers in the country without (or a lower) price cap: Less restrictive interconnector capacity results in higher imports during peak demand which dampens peak prices.

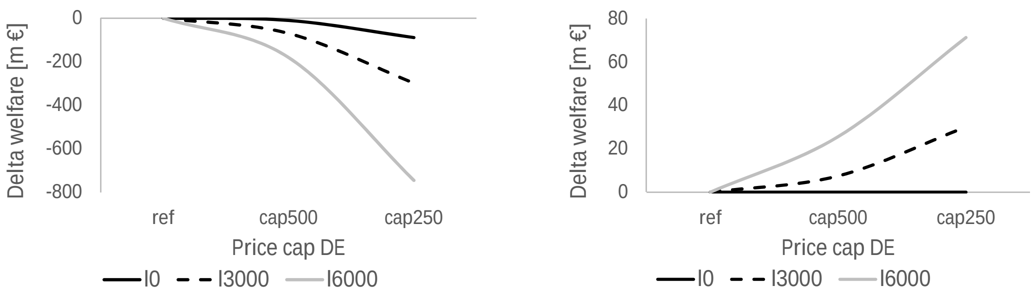


(A) Germany (B) France

FIGURE 2.9: Consumer surplus with a price cap in Germany

Although some of the losses in consumer surplus are compensated by (small) gains in congestion rent and revenues of renewables, for Germany a price cap results in all analysed cases in higher welfare losses compared to the case of an isolated market (see Figure 2.10a). The losses are so significant that having no connection may be better altogether.

The total welfare effect in France is positive in the analysed cases, but on a rather low level (see Figure 2.10b). Although domestic capacity recedes due to additional imports, gains in consumer surplus and congestion rent exceed slight losses in revenues of renewable energies.

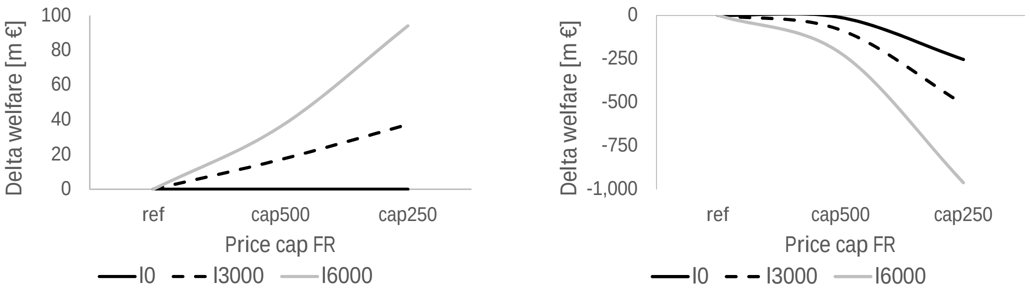


(A) Germany (B) France

FIGURE 2.10: Welfare effects with a price cap in Germany

The effect if only France has a price cap is illustrated in Figure 2.11. The shapes look similar to the previous case: The French price cap causes welfare losses in France that increase with increasing interconnector capacity and decreasing cap. For the neighbouring country that has no cap – in this case Germany – the effect is the other way around: The domestic welfare is increasing driven by increasing imports during peak demand as prices – in contrast to the capped market – signal scarcity.

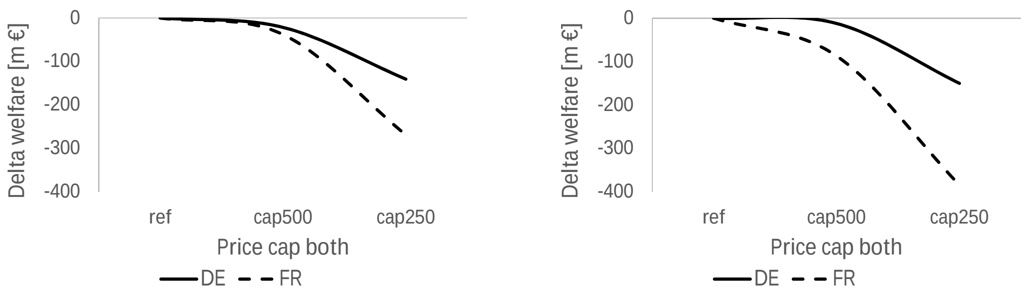
Negative effects of price caps increase in asymmetric (interconnected) markets – mainly affecting the market with the cap – but what is the effect in case of symmetric caps? Figure 2.12 illustrates the effect on welfare in Germany and France for both levels of



(A) Germany (B) France

Figure 2.11: Welfare effects with a price cap in France

interconnector capacity if both markets have the same price cap. The overall shape is similar for both regions: A price cap of 500 €/MWh has only a limited effect on welfare compared to a cap of 250 €/MWh. Again we can observe that a higher trade capacity is increasing the negative effects of price caps, even if they are synchronous. This is driven by the price dampening effect of trade which forces more capacity to recede: The price cap forces a decrease of generation capacity which otherwise could not recuperate investment costs. The higher the interconnector capacity the more of the missing capacity is replaced by imports (if capacity is available in the neighbouring market) which forces a stronger capacity cutback. Ultimately, less demand can be served which results in the higher deadweight losses.



(A) $\bar{l} = 3000$ (B) $\bar{l} = 6000$

Figure 2.12: Welfare in Germany and France with synchronous price caps

In summary our results suggest that in interconnected markets the negative effects of price caps are more severe (at least for the market having the cap) than in isolated markets. This is due to the supply flowing to neighbouring markets during scarcity of supply as the neighbours allow for higher prices. Additionally, imports hinder the recovering of prices with receding capacity which results in a sharper drop of generation capacity and ultimately welfare.

For the neighbouring markets the effects are much less pronounced. They even benefit from the additional imports during peak demand. But generation capacity is negatively affected in neighbouring markets, too.

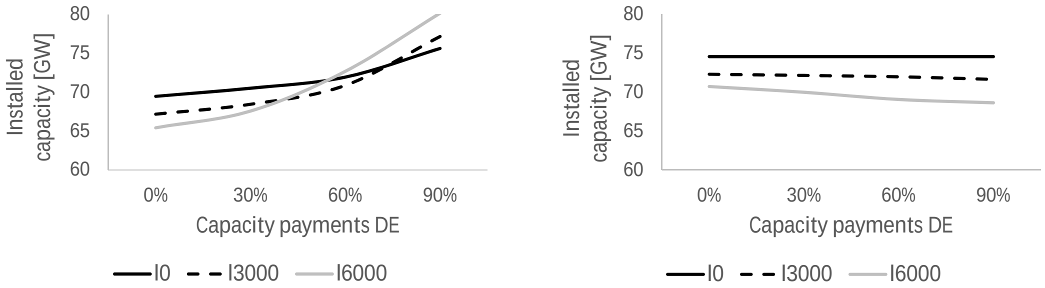
2.3.3.3 The effects of capacity payments with unilateral price cap

In the following we analyse the ability of capacity payments to compensate for losses triggered by price caps. For the case of isolated markets our results show that appropriate CP restore the optimal level of capacity and welfare. Now we analyse if the same holds true for connected markets. At first, like in the previous section, we start by assuming that only Germany has a price cap while prices in France are not restricted. Capacity payments to support domestic generation capacity are also assumed to be paid in Germany exclusively (the case of a price cap and capacity payments in France yields similar results, see Appendix C). Afterwards we introduce price caps in both markets to analyse if markets can rely on the efforts of their neighbours.

First we look at the effect of capacity payments on capacity. Like for the isolated case, capacity payments seem fit to increase the installed capacity and thereby restore the optimal level (see Figure 2.13a). Depending on the interconnector capacity the effect of CP is of varying magnitude: The higher the interconnector capacity the stronger is the effect on capacity. This is in line with the previous observation of a stronger decrease of generation capacity with higher trade capacities: The higher capacity decrease originated from the price dampening effect of imports which prevented prices to recover with decreasing generation capacity. As capacity payments directly support domestic investments this reinforcing negative influence of imports can be circumvented.

The impact of CP in Germany on capacity in France is small but increasing with higher transmission capacity (see Figure 2.13b). Capacity is decreasing as the additional capacity in Germany decreases prices in France. This indicates that in interconnected markets CP are not able to restore the welfare optimum like they did in isolated markets: Capacity in France already decreased due to the price cap and the resulting additional imports from Germany. As CP further decrease the capacity in France the equilibrium outcome of markets with no cap and no CP cannot be restored.

Increasing capacity in Germany as a result of capacity payments increases consumer surplus. At the same time capacity payments arise which have to be paid for. The net effect of increasing consumer surplus minus CP is illustrated in Figure 2.14a: With existing price caps the increase in consumer surplus exceeds the costs of capacity payments at first. For higher CP the net effect becomes negative. In contrast to the case of an isolated market, Germany cannot achieve the same welfare level as without a cap.



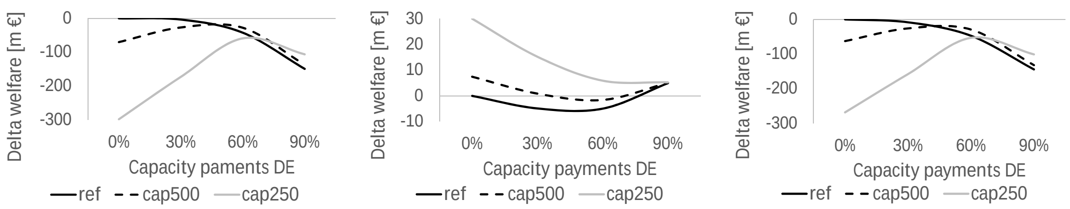
(A) Germany (B) France

FIGURE 2.13: Effects of capacity payments in Germany on capacity ($\bar{p} = 250$ €/MWh)

Although capacity payments increase domestic welfare up to a certain point the cost for the support scheme exceeds the benefits for the customers before the original level is reached. This is due to the distorted price signals still caused by the price cap which result in exports during scarce domestic supply.

For France the effects of CP in Germany are – compared to the effect on Germany – small (see Figure 2.14b). For lower levels of CP welfare slightly decreases driven by the receding generation capacity which is crowded out by the additional imports from Germany: Imports are less reliable than domestic capacity due to import restrictions. As a result, peak prices increase whereas medium prices decrease (the average price remains unchanged). As supply during high demand is more valuable than during medium demand, consumer surplus decreases. For higher levels of CP this negative impact on consumer surplus is overcompensated by an increase of congestion rent: The higher the CP are the more average prices diverge as CP lower the average market price in Germany. This increases the congestion rent.

In total, capacity payments cannot fully restore the optimal welfare level in Germany as well as overall (see Figure 2.14c): As the distortion of price signals persists, some inefficiencies in the allocation of capacity and supply remain.

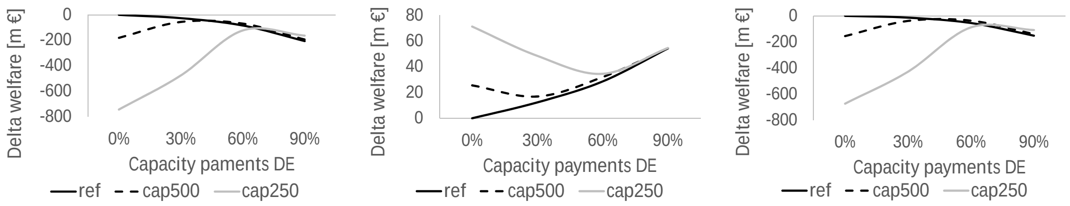


(A) Germany (B) France (C) Total

FIGURE 2.14: Welfare effect of capacity payments in Germany ($\bar{l} = 3{,}000$ MW)

Given an interconnector capacity of 6,000 MW the patterns for the welfare effect of capacity payments look similar for Germany and also the total effect. For France, the benefits of additional congestion rent exceed losses in consumer surplus at all times –

France is never worse off compared to the reference case with no price cap (see Figure 2.15).



(A) Germany (B) France (C) Total

FIGURE 2.15: Welfare effect of capacity payments in Germany ($\bar{l} = 6{,}000\,\mathrm{MW}$)

In summary, capacity payments are less effective in interconnected than in isolated markets. They can restore some of the welfare losses triggered by price caps, but cannot solve the inefficient price signals resulting in suboptimal allocation of capacity and supply. Thus some of the welfare losses remain.

The overall effects of CP on neighbouring markets are comparably small. They do not indicate a reason for concerns regarding the influence on price signals and the resulting effects on imports and exports.

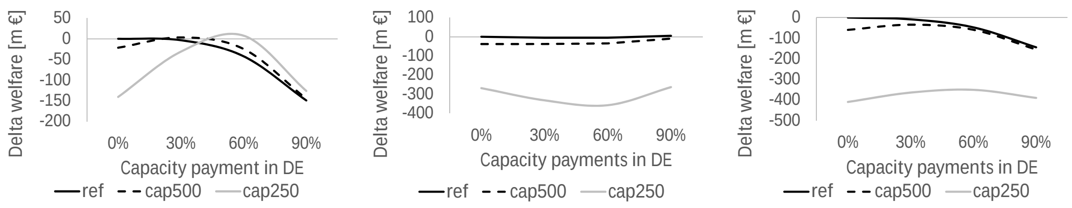
2.3.3.4 Spill-over effects of capacity payments with bilateral price caps

The previous analysis did not show large effects of price caps on neighbouring markets, they might even benefited from increasing congestion rent. Those analyses focused on neighbouring markets that do not have a price cap themselves. Now we look at the case where both markets have price restrictions. Building on the previous analysis we address the following questions: If both markets have a cap, do capacity payments in one market also help to overcome the negative effects in the other market? And if this is the case, does it reward free-riding behaviour?

We start our analysis by assuming that both markets face the same price cap and Germany introduces capacity payments. As seen earlier, capacity payments can have a positive effect on domestic welfare with unilateral caps. We can observe the same in this scenario for an interconnector capacity of 3,000 MW (see Figure 2.16 comparing welfare with the base case of no cap and no CP).[37] Welfare in Germany may even exceeds the level without caps and CP: Net benefits from increasing consumer surplus minus capacity payments recover most of the consumer surplus lost due to the price cap. Additional congestion rent results in extra welfare gains.

[37]The effects for an interconnector capacity of 6,000 MW are similar (see Appendix D).

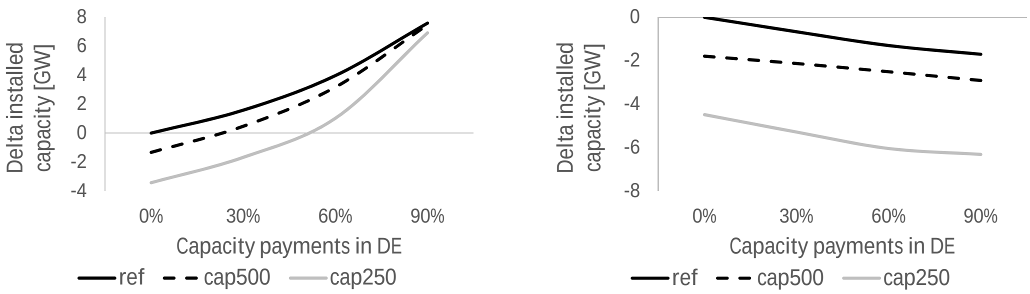
For France the effect is negligible for a common cap of 500 €/MWh – with a cap of 250 €/MWh capacity payments in Germany result in additional welfare losses for France (up to the optimal level of CP in Germany).



(A) Germany (B) France (C) Total

FIGURE 2.16: Effects of capacity payments in DE on welfare ($\bar{l} = 3{,}000$ MW)

This is because the additional capacity in Germany crowds out generation capacity in France to an even lower level than was already reached due to the price cap (see Figure 2.17). Assuming a price cap of 250 €/MWh in both markets, Germany reaches about the same level of capacity at CP of almost 60 % of investment costs of the peak technology – France looses about 1.5 GW of capacity additionally.



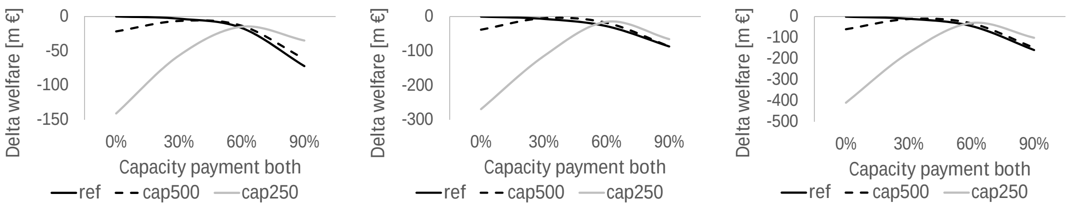
(A) Germany (B) France

FIGURE 2.17: Effects of capacity payments in DE on installed capacity ($\bar{l}$=3,000 MW)

Our results indicate that capacity payments are beneficial for the introducing country given that both markets have a price cap. For the neighbouring market the effects are rather small and, depending on the actual set up, may even be harmful. We complement our analysis by introducing capacity payments also in France. Thus, in the following, both markets feature the same price cap as well as the same level of CP.

Figure 2.18 illustrates the welfare effects for a trade capacity of 3,000 MW.[38] France can clearly benefit from having its own capacity payments instead of relying on its neighbour. Given an optimal choice of capacity payments, total welfare is only slightly below the optimal value achieved with no cap and no CP. The remaining losses are due to insufficient price signals – capacity and supply are not perfectly allocated.

[38] See Appendix D for the case of 6,000 MW.



(A) Germany (B) France (C) Total

FIGURE 2.18: Effects of capacity payments in both markets on welfare ($\bar{l} = 3{,}000\,\text{MW}$)

In summary, our results concerning possible spill-over effects do not support concerns regarding incentives for free-riding. First of all, adjacent markets with no own CRM do not appear to hinder the effectiveness of capacity payments. Furthermore, relying on capacity support in neighbouring markets does not help to overcome domestic deficits. If neighbouring markets both support capacity investments, most of the deadweight losses resulting from price caps can be prevented.

2.4 Conclusions

Existing literature on market imperfections and capacity mechanisms mainly focuses on a national perspective. Recent concerns that national CRMs might harm neighbouring markets by distorting market prices, for example raised by the EU Commission, are only rarely discussed.

We extend the literature on cross-border effects by using an equilibrium model with real world data. Thereby, we are able to cover the complex interaction of supply and demand in interconnected markets. We apply our model to the case of Germany and France to analyse the cross-border and interaction effects of price caps and capacity payments.

Our results indicate that the interaction with neighbouring markets not only influences the effect of market failure on domestic welfare but also the efficiency of capacity remuneration mechanisms. In our case, price caps result in more severe welfare losses and capacity payments are less efficient in countering the resulting market distortions. Given the frequently expected capacity shortages in the future and the numerous CRMs currently already implemented in the European electricity market, this raises further questions concerning the efficiency of the existing uncoordinated national mechanisms.

At least our results do not indicate serious negative effects of price caps on neighbouring markets. Similarly, this should be the case for a strategic reserve (which is already implemented or discussed in some markets) or excess capacity that may be triggered by overly high capacity payments or oversized capacity auctions as they have similar properties.

Likewise, our results do not indicate (relevant) positive effects of capacity payments on neighbouring markets that face insufficient price signals themselves. Thus, we do not identify incentives for free-riding on the capacity efforts of neighbours.

2.5 Appendix

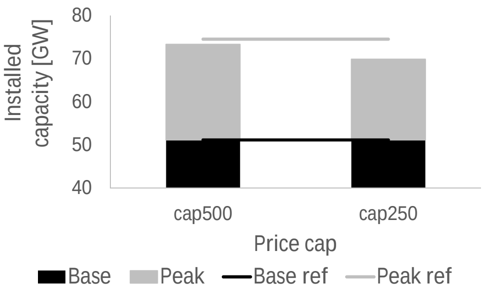
2.5.1 Appendix A: Model assumptions

Table 2.3: Assumptions concerning annual wind and photovoltaics generation

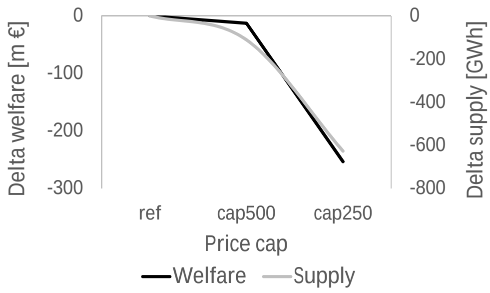
	Wind generation [TWh]	Photovoltaics generation [TWh]
DE	109	48
FR	55	32

2.5.2 Appendix B: Isolated France

Figure 2.19 illustrates the effects of price caps and capacity payments in France (with no connection to Germany). The general shape of the effects of price caps is similar to those for Germany – but the installed capacity, welfare and supply are affected to a larger extent. This is the result of a steeper residual demand curve in the relevant segment around the installed capacity: A decrease in installed capacity results in a smaller increase in average market prices. Therefore, capacity in France has to decrease to a larger extent – compared to Germany – in order to meet the equilibrium conditions of investment.



(A) Installed capacity



(B) Welfare and supply

FIGURE 2.19: Effects of price caps on capacity, supply and welfare in an isolated France

The stronger decrease of capacity results in less demand being served which again results in a higher deadweight loss. The negative effects of price caps can – equal to the German case – be countered by capacity payments that fully restore the optimal welfare level given an appropriate level of payments (see Figure 2.20).

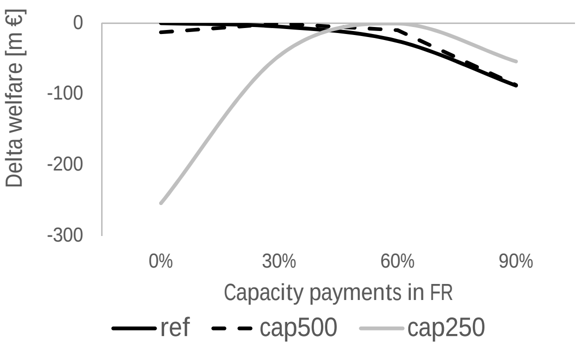
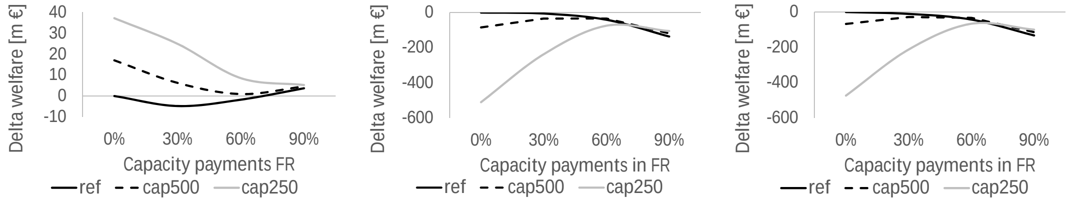


FIGURE 2.20: Effects of price caps and capacity payments in an isolated France

2.5.3 Appendix C: Interconnected France

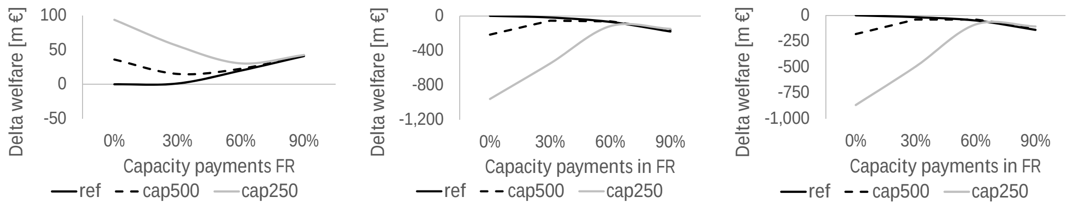
The effects of price caps and capacity payments in France are similar to the case for Germany (see Figure 2.21), the conclusions drawn earlier also apply here (see section 2.3.3.3). The effect in France is more pronounced than in the case for Germany which is due to the steeper residual supply curve as discussed in section 2.5.2.



(A) Germany (B) France (C) Total

FIGURE 2.21: Welfare effect of price caps and CP in France ($\bar{l} = 3{,}000\,\mathrm{MW}$)

Similar but amplified effects can be observed for a higher interconnector capacity (see Figure 2.22). This is in line with previous results.

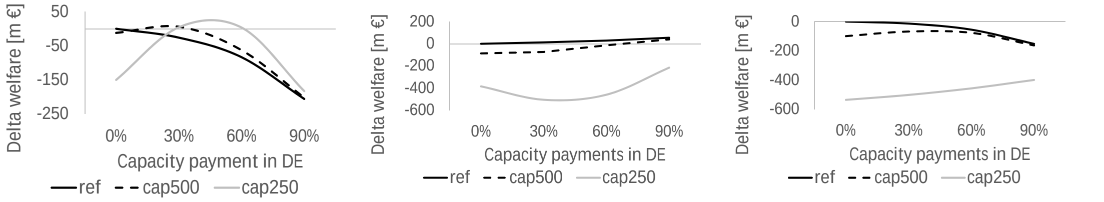


(A) Germany (B) France (C) Total

FIGURE 2.22: Welfare effect of price caps and CP in France ($\bar{l} = 6{,}000\,\mathrm{MW}$)

2.5.4 Appendix D: Bilateral price caps – Additional results

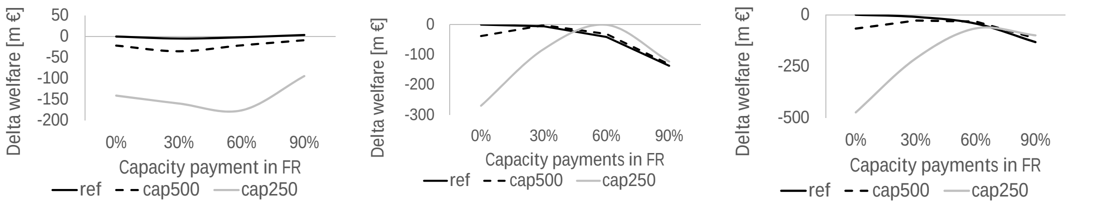
Here we complement the analysis from section 2.3.3.4. We start with the case of capacity payments in Germany and a interconnector capacity of 6,000 MW. The effects are similar to the case with less transmission capacity (see Figure 2.23). This time, the optimal level of CP is slightly lower than in the earlier example.



(A) Germany (B) France (C) Total

FIGURE 2.23: Effects of capacity payments in DE on welfare ($\bar{l} = 6{,}000$ MW)

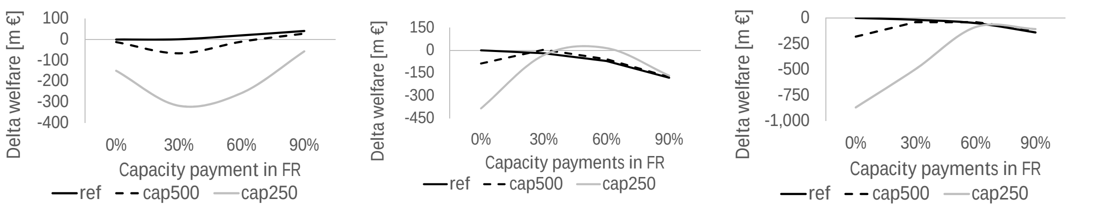
The following Figure 2.24 illustrates the effects of capacity payments in France on welfare if prices in both markets are limited by a cap.



(A) Germany (B) France (C) Total

FIGURE 2.24: Effects of capacity payments in FR on welfare ($\bar{l} = 3{,}000$ MW)

Figure 2.25 illustrates the same case but this time both markets are connected by a higher transmission capacity.

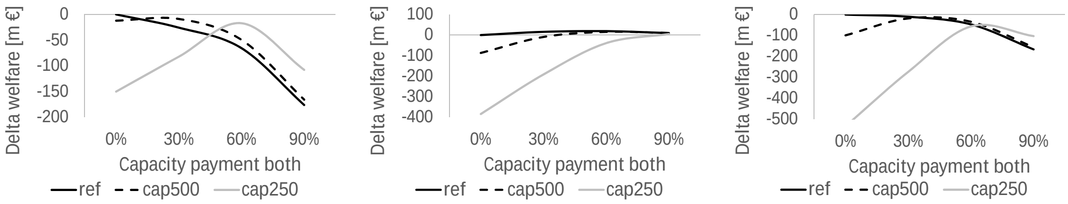


(A) Germany (B) France (C) Total

FIGURE 2.25: Effects of capacity payments in both markets on welfare ($\bar{l} = 6{,}000$ MW)

The welfare effects of equal price caps and capacity payments in both countries and interconnector capacity $\bar{l} = 6{,}000$ MW are illustrated in Figure 2.26.

For all three cases presented here the results in the main body of this text are also applicable.



(A) Germany (B) France (C) Total

FIGURE 2.26: Effects of capacity payments in both markets on welfare ($\bar{l} = 6{,}000$ MW)

Part II

Strategic behaviour in spacial natural resource markets

Assessing market structures in resource markets – An empirical analysis of the market for metallurgical coal using various equilibrium models

The prevalent market structures found in many resource markets consist of high concentration on the supply side and low demand elasticity. Market results are therefore frequently assumed to be an outcome of strategic interaction between producers. Common models to investigate the market outcomes and underlying market structures are games representing competitive markets, strategic Cournot competition and Stackelberg structures that take into account a dominant player acting first followed by one or more players. We add to the literature by expanding the application of mathematical models and applying an Equilibrium Problem with Equilibrium Constraints (EPEC), which is used to model multi-leader-follower games, to a spatial market. Using our model, we investigate the prevalent market setting in the international market for metallurgical coal between 2008 and 2010, whose market characteristics provide arguments for a wide variety of market structures. Using different statistical measures to compare model results with actual market outcomes, we find that two previously neglected settings perform best: First, a setting in which the four largest metallurgical coal exporting firms compete against each other as Stackelberg leaders, while the remainders act as Cournot followers. Second, a setting with BHPB acting as sole Stackelberg leader.

3.1 Introduction

Many resource markets suffer from high concentration on the supply side and low demand elasticity. Market results are therefore frequently assumed to be an outcome of strategic interaction between producers. The use of mathematical models to analyse market outcomes to gain insights into underlying market structures has a long tradition in the economic literature. Common models are one-stage games representing competitive markets or Cournot competition. More advanced two-stage models of the Stackelberg kind take into account a single leader followed by one or more players. We add to the literature by expanding the application of mathematical models and applying an Equilibrium Problem with Equilibrium Constraints (EPEC) to a spatial market, i. e., a setup with multiple, geographically disperse demand and supply nodes. This model class is used to simulate multi-leader-follower games. This enables us to investigate more complex market structures that have been neglected in previous studies on resource markets. Omitting these market structures may result in false conclusions about the prevalent state of competition.

The paper at hand investigates which market structure was prevalent in the international market for metallurgical coal during the time period 2008 to 2010.[39] The international metallurgical coal market is particularly suited for this kind of analysis since, first, the supply side is dominated by four large mining firms (hereafter referred to as the Big-Four), namely BHP Billiton (BHPB), Rio Tinto, Anglo American and Xstrata. Second, metallurgical coal is an essential input factor in producing pig iron and difficult to substitute, causing demand to be rather price inelastic. Third, in the period under scrutiny in this paper, yearly benchmark prices were negotiated between representatives of the Big-Four and representatives of the large Asian steel makers (Bowden, 2012). Fourth, one of the firms of the Big-Four, BHP Billiton, is by far the largest firm in the international market for metallurgical coal. Nonetheless, the other firms played a central role in the negotiations as well. Consequently, a wide variety of market structures may be a plausible approximation of the actual market setting.

Our research adds to that of Graham et al. (1999) and Trüby (2013) who were the first to analyse the market for metallurgical coal. The former investigates various market settings for the year 1996, in which firms or consumers simultaneously choose quantities. In contrast, the latter focusses on the time period from 2008 to 2010. Regarding the market structures, the author arrives at the conclusion that assuming the Big-Four

[39] The terms metallurgical and coking coal are often used interchangeably in the related literature as well as throughout this paper. Yet, this is not entirely correct since metallurgical coal includes coals (as it is the case in our data set) that technically are thermal coals but can be used for metallurgical purposes as well, such as pulverised coal injection (PCI).

jointly act as a Stackelberg leader provides the best fit to the actual market outcome. However, Trüby finds that it cannot be ruled out that firms in the market simply engaged in an oligopolistic Cournot competition. We add to the literature by extending the scope of possible market structures.[40]

More specifically, we simulate one scenario in which the Big-Four compete against each other at a first stage, i. e., choose output to maximise individual profits, while the remaining firms form a Cournot fringe and act as followers. This constitutes a multi-leader-follower game. In another scenario, BHP Billiton takes on the role as the sole Stackelberg leader, with the rest of the Big-Four choosing quantities simultaneously with the remaining players as followers. Thereby, we broaden the range of market structures analysed in the field of spatial resource markets as multi-leader games have thus far been omitted from existing studies. As investigating collusive behaviour in markets using simulation models crucially depends on an appropriate and comprehensive market representation, multi-leader games may help to expose previously overlooked market structures. Since it is a priori not clear which is the correct demand elasticity, we run the market simulations for a wide range of values. To assess whether one of the market structures is superior to the others, we compare simulated prices, trade flows and production volumes of the Big-Four to realised market outcomes. In order to compare trade flows, different statistical measures/tests are applied as suggested by, e. g., Bushnell et al. (2008), Paulus et al. (2011), and Hecking and Panke (2014).

This paper contributes to the literature on applied industrial organisation and, more specifically, the analysis of the international market for metallurgical coal. We expand previous studies by applying an Equilibrium Problem with Equilibrium Constraints (EPEC), a mathematical programme used to model multi-leader-follower settings, to a spatial market, i. e., a market with multiple, geographically disperse supply and demand nodes. In doing so, we find that the two additional market settings proposed in this paper provide a good fit with realised market outcomes for the time period 2008 to 2010. In addition, by analysing production volumes and profits of the Big-Four, we enhance the market structure analysis by providing an additional plausibility check. We are able to show that even if simulated prices and trade flows fit well with market outcomes, a scenario in which the Big-Four form a Cartel that acts as a Stackelberg leader is less likely since production volumes deviate from actual production. More importantly, additional revenues of the Big-Four from forming and coordinating a cartel are rather small compared to a scenario in which all four compete against each other at a first stage. Accounting for the transaction costs caused by the coordination of the cartel

[40] Graham et al. (1999) indicate that there could be market power on the demand side as well. However, given that two out of the three years under consideration in our paper are characterised by high prices, we focus on setups with market power on the supply side. An analysis of market power on the demand side could be a subject of further research.

would further decrease possible benefits. Concerning the demand elasticity, we detect that simulated prices for elasticities from -0.3 to -0.5 seem to be within a reasonable range for most of the market structures.

Summing up our findings, one of the main advantages of simulation models is that they allow us to assess different market structures. Yet, as shown in our paper, it may be difficult to decide on one setting that provides the best fit. Consequently, such analyses need to be accompanied by additional analyses similar to our comparison of production volumes of the Big-Four. To be able to further narrow down the number of potential market structures, additional data such as firm-by-firm export volumes, which were not available for all relevant firms in our example, would be helpful.

The remainder of this paper is structured as follows: Section 3.2 offers an overview of the relevant literature, while the methodology is described in Section 3.3. Section 3.4 briefly describes the numerical data used in this study. Section 3.5 is devoted to the analyses of the empirical results. Section 3.6 concludes.

3.2 Literature review

Commodity markets have often been subject to concerns of high concentration on the supply side, with several prominent examples being the markets for energy resources such as oil, natural gas or metallurgical coal. Consequently, there has been substantial academic research in an attempt to assess whether companies or countries exercised market power. In order to do so, one of two different methodological approaches – econometric methods or simulation models – is applied. While both approaches have their respective advantages and disadvantages[41], one of the most persuasive arguments in favour of using simulation models to assess the exercise of market power is that they are highly flexible with respect to the specific market structure. This, in principle, not only enables researchers to answer the question whether or not market power in a specific market has been exercised, but also provides hints as to which kind of market structure is prevalent, e. g., whether firms form a cartel or show no signs of explicit cooperation.

The use of mathematical programming models to analyse spatial markets has a long tradition in economics. Enke (1951) first described the problem of spatial markets, proposing a solution method using a simple electric circuit to determine equilibrium prices and quantities in competitive markets. Samuelson (1952) showed how the problem can be cast into a (welfare) maximisation problem and thereafter be solved using linear programming. Together with Takayama and Judge (1964, 1971) who extend the spatial

[41]For a brief overview of the various econometric approaches used in the literature and their respective advantages and drawbacks, see Germeshausen et al. (2014).

market representation (e. g., by including monopolistic competition), Samuelson's work is generally considered to have laid the groundwork for spatial market analysis using mathematical programming.

Advances in the representation of markets were made during the 1980s by modelling imperfect competition (e. g., by Harker, 1984, 1986, Nelson and McCarl, 1984). This has frequently been done since then, e. g., for steam coal markets (Haftendorn and Holz, 2010, Kolstad and Abbey, 1984, Trüby and Paulus, 2012), natural gas markets (Boots et al., 2004, Egging et al., 2010, Gabriel et al., 2005a, Growitsch et al., 2013, Holz et al., 2008, Zhuang and Gabriel, 2008), wheat markets (Kolstad and Burris, 1986), oil markets (Huppmann and Holz, 2012) or for the coking coal and iron ore markets (Hecking and Panke, 2014). A multi-fuel market model is presented in Huppmann and Egging (2014)

We focus our analysis on the metallurgical coal market. A recent analysis of short-term market outcomes by Trüby (2013) indicates that the market from 2008 to 2010 may has been characterised by firms exercising market power. This rejects the previous findings by Graham et al. (1999), although this study focuses on 1996.

Most of the aforementioned studies use models that assume players make decisions simultaneously. This model type can be extended to represent bi-level games, the classical example being Stackelberg games (Stackelberg, 1952). There are several applications for this type of problem, which can be modelled as a Mathematical Problem with Equilibrium Constraints (MPEC). MPECs are constrained optimisation problems, with constraints including equilibrium constraints (see Luo et al., 1996, for an overview of MPECs). MPECs have for instance been used to model power markets, e. g, by Gabriel and Leuthold (2010), Wogrin et al. (2011) and natural gas markets, e. g., by Siddiqui and Gabriel (2013). Bi-level games are, due to non-linearities, computationally more challenging to solve in comparison to one-level games.

The single-leader Stackelberg game can be extended to a multi-leader-follower game in which several players make decisions prior to one or more subsequent players. Any solution to this game must maximise leaders' profits while simultaneously taking into account the equilibrium outcome of the second stage. This results in an Equilibrium Problem with Equilibrium Constraints (EPEC). Due to the concatenation of several MPEC problems to one EPEC and the resulting high non-linearity, EPECs are even more difficult to solve than MPECs. Previous EPEC models have mostly been used to analyse electricity markets, e. g., by Barroso et al. (2006), Sauma and Oren (2007), Shanbhag et al. (2011), Yao et al. (2008) and Wogrin et al. (2013a). In addition, Lorenczik et al. (2017) analyse investment decisions in the metallurgical coal market.

3.3 Methodology

3.3.1 Market Structures

Due to its market structure (with few large producers and relatively low demand elasticity), the metallurgical coal market is often presumed to lack competition. This suspicion is confirmed by a recent study showing that market outcomes can be reproduced by assuming strategic rather than competitive behaviour. Trüby (2013) finds that over the years 2008 to 2010, assuming perfect competition, neither trade flows nor prices match well with actual market results. In contrast, the non-competitive market structures considered in the paper perform reasonably well with the exception of the Cournot Cartel case.[42] The paper's conclusion regarding the market structures is that assuming the Big-Four jointly act as a Stackelberg leader provides the best fit to the actual market outcome. However, it cannot be ruled out that firms in the market simply engaged in an oligopolistic Cournot competition. Therefore, two of the scenarios analysed in Trüby (2013), namely the case of Cournot competition (hereafter, referred to as MCP, which is the programming approach used to simulate the market setting) and a setting in which the Big-Four form a cartel that acts as the Stackelberg leader (MPEC Cartel) are taken into consideration in this paper as well to ease the comparison of results.

We expand the range of investigated market structures by analysing a multi-leader-follower game as well as one additional market setting involving one Stackelberg leader. In the multi-leader-follower game, the Big-Four compete against each other at the first stage and take into account the reaction of the other firms engaging in Cournot competition at the second stage (EPEC Big 4). We reason that this setting is relevant since, first, benefits in terms of additional revenues from forming a cartel are rather small when compared to the EPEC Big 4 scenario, even without accounting for the transaction costs that go along with coordinating a cartel. Thus, while still acting as leaders, it is reasonable to assume that the Big-Four compete against each other. Second, the simulated production volumes by the Big-Four fit historical production data better in the two additional settings proposed in this paper than in the MPEC Cartel case. Thus, they are worth a closer investigation. Both reasons will be discussed in depth in Section 3.5.3.

Finally, we simulate an additional single Stackelberg leader setting in which BHP Billiton sets quantities in a first stage with the remaining firms being followers (MPEC BHBP).

[42] In the Cournot Cartel case, the Big-Four are assumed to engage in a cartel and, thus, jointly optimise their total supply. Trüby (2013) finds that under this market setting, prices could only be reproduced when assuming very high elasticities. Concerning trade flows, the linear hypothesis tests suggest that simulated trade flows do not resemble actual market outcomes in 2009 for all elasticities, while in the other years the H_0-hypothesis could be rejected for elasticities up to -0.2 (2008) and -0.3 (2010).

The main reason that modelling such a market structure is intuitive is the fact that BHBP is by far the world's most important coking coal miner. Figure 3.1 provides an overview of the market structures investigated in this paper.

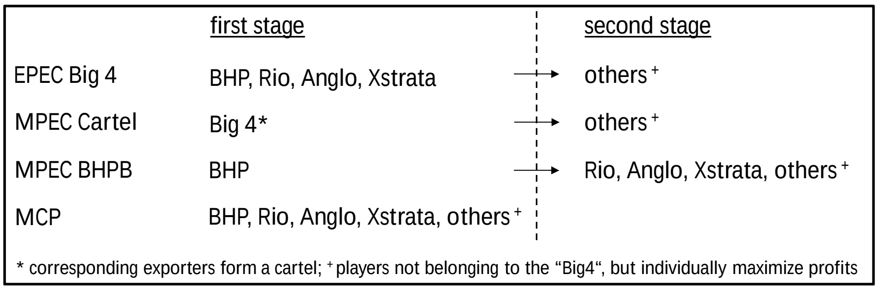


FIGURE 3.1: Overview of modelled market structures

To simulate the different aforementioned coking coal market settings, three different types of simulation models are used. The first calculates the expected market outcome in a Cournot oligopoly in which all players decide simultaneously about produced and shipped quantities. The two other models constitute bi-level games in which players act in consecutive order. In the Stackelberg game, one player (or a group of players forming a cartel) acts first followed by the remaining players. The last model type represents a market with multiple (Stackelberg) leaders and one or more followers. From a modelling perspective, the first model constitutes a Mixed Complementary Problem (MCP). The second and third models are implemented as a Mathematical Problem with Equilibrium Constraints (MPEC) and an Equilibrium Problem with Equilibrium Constraints (EPEC), respectively.

3.3.2 Model descriptions

Although we focus our analysis on the coking coal market, the model is suitable for a multitude of similar commodity markets such as the iron ore, copper ore, oil or gas market, which are characterised by a high concentration on the supply side and therefore may not be competitive. Thus, we use general terms for the model description as well as generic notation to emphasise the applicability of our approach to markets other than the coking coal market. Table 3.1 summarises the most relevant nomenclature used throughout this section, i. e., displays the abbreviations used for the various model sets, parameters and variables and provides a short description. Additional symbols are explained throughout the text where necessary.

TABLE 3.1: Model sets, parameters and variables

Abbreviation	Description
Model sets	
$i \in I$	Players
$j \in J$	Markets
$m \in M$	Production facilities
Model parameters	
a_j	Reservation price [per unit]
b_j	Linear slope of demand function
c_m	Variable production costs [per unit]
cap_m	Production capacity [units per year]
$tc_{i,j}$	Transportation costs [per unit]
Model variables	
P_j	Market price [per unit]
$s_{i,j}$	Supply [units]
x_m	Production [units]

3.3.2.1 The MCP model

The first model assumes a market in which all producers decide simultaneously about the use of production facilities and the delivery of goods. Each player $i \in I$ maximises profits according to:

$$\max_{x_m, s_{i,j}: m \in M_i} \sum_j P_j \cdot s_{i,j} - \sum_{j \in J} tc_{i,j} \cdot s_{i,j} - \sum_{m \in M_i} c_m \cdot x_m$$

subject to

$$\begin{aligned}
&cap_m - x_m \geq 0, \ \forall m \in M_i \ (\lambda_m) \\
&\sum_{m \in M_i} x_m - \sum_j s_{i,j} \ \geq 0 \ (\mu_i) \\
&P_j = a_j - b_j \cdot (s_{i,j} + S_{-i,j}), \ \forall j \\
&s_{i,j} \geq 0, \ \forall j \\
&x_m \geq 0, \ \forall m \in M_i \ .
\end{aligned}$$

Total supplied quantities $S_{-i,j}$ $(= \sum_{-i \neq i} s_{-i,j})$ to market j by other producers $(-i)$ are taken as given. Hence, each producer maximises revenues minus costs (production plus transportation) taking into account capacity restrictions (with λ_m being the dual variable for the capacity limit) and the restriction that total production has to be greater than total supply (with μ_i as the respective dual variable). As all production facilities of each player are located in the same area, transportation costs between production and specific demand nodes are assumed to be identical. Since different years are not interlinked, they can be optimised separately.

Maximising each players' profits is equivalent to finding a solution that satisfies the following Karush-Kuhn-Tucker (KKT) conditions simultaneously for all players:

$$
\begin{aligned}
&0 \leq tc_{i,j} - P_j + b_j \cdot s_{i,j} + \mu_i \perp s_{i,j} \geq 0, \ \forall \ i, j \\
&0 \leq c_m + \lambda_m - \mu_i \perp x_m \geq 0, \ \forall \ m \in M_i \\
&0 \leq cap_m - x_m \perp \lambda_m \geq 0, \ \forall \ m \\
&0 \leq \sum_{m \in M_i} x_m - \sum_j s_{i,j} \perp \mu_i \geq 0, \ \forall \ i \\
&P_j = a_j - b_j \cdot (s_{i,j} + S_{-i,j}), \ \forall \ j \\
&s_{i,j} \geq 0, \ \forall \ i, j \\
&x_m \geq 0, \ \forall \ m \ ,
\end{aligned}
$$

with the perp operator ($\perp$) meaning that the product of the expressions to the left and to the right has to equal zero. The first inequality reflects the first order condition for the optimal supply of player i to region j: Marginal revenues of additional supply (i. e., market price P minus transportation costs tc and the marginal costs of supply μ) have to equal supply times the slope of the linear demand function b, i. e., the reduction of revenue due to the negative price effect of additional supply. The second inequality, which represents the first order condition for production, reflects the marginal costs of supply μ as the sum of variable production costs c and the scarcity value of capacity λ. The third and fourth conditions represent the complementarity conditions forcing production to be within the capacity limit (with λ being the scarcity value of capacity) and production to meet supply (with marginal production costs μ). The equality condition constitutes the linear demand function followed by non-negativity constraints for supply and production.

Due to the strict quasi-concave objective function and the convexity of restrictions, the KKT conditions are necessary and sufficient for an optimal solution.

3.3.2.2 The MPEC model

In the MPEC model, we seek to represent a Stackelberg market structure with one leader (l) taking into account the equilibrium decisions of the follower(s). The model equations are as follows:

$$
\max_{x_m, s_{l,j}, \lambda_m, \mu_i} \sum_j P_j \cdot s_{l,j} - \sum_{j \in J} tc_{l,j} \cdot s_{l,j} - \sum_{m \in M_l} c_m \cdot x_m
$$

subject to

$$
\begin{aligned}
&0 \leq tc_{i,j} - P_j + b_j \cdot s_{i,j} + \mu_i \perp s_{i,j} \geq 0, \ \forall \ i \neq l, j \\
&0 \leq c_m + \lambda_m - \mu_i \perp x_m \geq 0, \ \forall \ m \in M_{i \neq l} \\
&0 \leq cap_m - x_m \perp \lambda_m \geq 0, \ \forall \ m \in M_{i \neq l} \\
&0 \leq \sum_{m \in M_i} x_m - \sum_j s_{i,j} \perp \mu_i \geq 0, \ \forall \ i \neq l \\
&P_j = a_j - b_j \cdot (S_{-i,j} + s_{l,j}), \ \forall \ j \\
&s_{i,j} \geq 0, \ \forall \ i, j \\
&x_m \geq 0, \ \forall \ m \ .
\end{aligned}
$$

Thus, the leader decides on supply taking the equilibrium outcome of the second stage (which influences the market price) into account. The followers ($-i$) take the other followers' as well as the leader's supply as given. The objective function is non-convex and thus solving the MPEC problem in the form previously described does usually not guarantee a globally optimal solution. Thus, we transform the model into a Mixed Integer Linear Problem (MILP) that can be solved to optimality with prevalent solvers.

There exist several approaches for linearising the existing non-linearities. Due to its simple implementation, we follow the approach presented by Fortuny-Amat and McCarl (1981) for the complementary constraints (for an alternative formulation see Siddiqui and Gabriel, 2013). For instance, the non-linear constraint

$$0 \leq c_m - P_j + b_j \cdot s_{i,j} + \lambda_m \perp s_{i,j} \geq 0$$

is replaced by the following linear constraints

$$
\begin{aligned}
&0 \leq c_m - P_j + b_j \cdot s_{i,j} + \lambda_m \leq M \cdot u_{i,j} \\
&0 \leq s_{i,j} \leq M(1 - u_{i,j})
\end{aligned}
$$

with M being a large enough constant (for hints on how to determine M, see Gabriel and Leuthold (2010)).

For the remaining non-linear term in the objective function ($P_j \cdot s_{i,j}$), we follow the approach presented by Pereira et al. (2005) using a binary expansion for the supply variable $s_{i,j}$. The continuous variable is replaced by discrete variables

$$s_{i,j} = \Delta_s \sum_k 2^k b^s_{k,i,j}$$

where Δ_s represents the step size, i. e., the precision of the linear approximation, and k the number of steps. Variables $b^s_{k,i,j}$ are binary. The term $P_j \cdot s_{i,j}$ in the objective function is replaced by $P_j \cdot \Delta_s \sum_k 2^k z^s_{k,i,j}$. In addition, the following constraints have to be included in the model

$$
\begin{aligned}
&0 \leq z^s_{k,i,j} \leq M^s b^s_{k,i,j} \\
&0 \leq P_j - z^s_{k,i,j} \leq M^s \left(1 - b^s_{k,i,j}\right) \ .
\end{aligned}
$$

The thereby formulated model constitutes a MILP that can be reliably solved to a globally optimal solution.[43]

3.3.2.3 The EPEC model

The EPEC model extends the Stackelberg game by enabling the representation of several leaders taking actions simultaneously under consideration of the reaction of one or more followers. The solution of an EPEC constitutes the simultaneous solution of several MPECs. Whereas MPECs are already difficult to solve due to their non-linear nature, it is even more difficult to solve EPECs. KKT conditions generally cannot be formulated for MPECs as regularity conditions are violated. Our model is solved using a diagonalisation approach. In doing so, we reduce the solution of the EPEC to the solution of a series of MPECs. The iterative solution steps are as follows:

1. Define starting values for the supply decisions $s^0_{l,j}$ of all leaders $l \in L$, a convergence criterion ϵ, a maximum number of iterations N and a learning rate R
2. $n = 1$
3. For all leaders,
   (a) Fix the supply decisions for all but the current leader
   (b) Solve current leader's MPEC problem to obtain optimal supplies $s^n_{l,j}, \ \forall j$
   (c) Set $s^n_{l,j}$ equal to $(1 - R) \cdot s^{n-1}_{l,j} + R \cdot s^n_{l,j}, \ \forall j$
4. If $|s^n_{l,j} - s^{n-1}_{l,j}| < \epsilon$ for all producers: equilibrium found, quit
5. If $n = N$: failed to converge, quit
6. $n = n + 1$: return to step 3.

EPECs may or may not have one or multiple (pure strategy) equilibrium solutions, and only one solution can be found per model run. In addition, if the iterations do not

[43] Within the range of the discretisation of the production variable.

converge to an equilibrium, this does not necessarily mean that no solution exists. This problem can partially be solved using multiple initial values for the iteration process, but it cannot be guaranteed that additional equilibria have not been missed. Despite these drawbacks, diagonalisation has been used widely and successfully in the corresponding literature (see Gabriel et al. (2012) and the literature cited therein).

For each EPEC setting, we run our model five times with varying start values and iteration orders to check for multiple equilibria. Each run converged to similar results with deviations of prices from the mean values of maximum 5%, single trade flows below 1.2 Mt and total production per mine below 0.6 Mt. Profits of the Big-Four and the cartel groups differed to a maximum of 1%. Whether theses deviations are due to a multiplicity of (similar) equilibra or to the (lack of) precision of the applied algorithm is not quite clear. In consideration of the almost equal results, we refrain from further analyses of the deviations.

3.4 Data

Modelling international commodity markets may be computationally challenging due to their spatial nature, i. e., multiple supply and demand nodes. In most empirical examples, each supply node is able to transport the commodity to each demand node giving rise to a large set of potential trade routes. The possible routes rapidly increase with additional demand or supply nodes. Whether a certain set of trade routes turns out to be computationally challenging depends on which market structure one would like to analyse. While solvers for Mixed Complementary Problems such as PATH (see Dirkse and Ferris, 1995) can handle quite large systems of equations and variables, the same setup may be intractable when formulated as a Mathematical Problem with Equilibrium Constraints (MPEC) or other more complex problems such as an Equilibrium Problems with Equilibrium Constraints (EPEC) due to their high non-linearity.

Since we are particularly interested in how well a multi-leader-follower game is able to model the coking coal market, we had to reduce the number of mines per player to one to keep the model feasible.[44] To ensure comparability, the same data setup was used for all market structures analysed in this paper irrespective of whether the respective solvers may have been able to handle larger sets of equations and variables (see Appendix A for production and shipping costs as well as capacities).

In total, the model used to conduct our empirical analysis consists of twelve supply nodes and six demand nodes. The supply side consists of individual firms as well as

[44] We would like to thank Johannes Trüby for allowing us to use his extensive mine-by-mine dataset on the international market for metallurgical coal.

countries. In addition to each of the four firms belonging to the Big-Four, i. e., BHP Billiton (BHPB), Rio Tinto, Anglo American and Xstrata, eight country supply nodes are included in the model of the international coking coal market (Table 3.2 shows which countries on the supply and demand side are represented in the model). When aggregating the data, production capacities of each mine belonging to the same firm or country are simply added up. Concerning production costs, we use the quantity-weighted average of the individual mines of a firm or country.

TABLE 3.2: Overview of firms and countries used in the model

Supply nodes	**Demand nodes**	**Countries/regions belonging to demand node**
BHP Billiton	JP_KR	Japan and Korea
Rio Tinto	CN_TW	China and Taiwan
Anglo American	IN	India
Xstrata Australia	LAM	Latin America (mainly Brazil and Chile)
Canada	EUR_MED	Europe and Mediterranean
China	Other	Africa and Middle East
Indonesia		
New Zealand		
Russia		
South Africa		
United States		

The demand side is represented by six nodes, most of which represent a demand cluster, with India being the only exception. The demand clusters were chosen based on geographical proximity and importance for international trade of metallurgical coal. Geographical proximity is important because shipment costs, which represent a large share in total import costs, largely depend on the shipping distance. Due to their minor importance in terms of the share of total import volumes, we included Africa and the Middle East in one demand node despite the large area this demand node covers. Inverse demand functions are assumed to be linear (see Table 3.4 in Appendix A for the used market data).[45] Since it is a priori not clear which is the correct elasticity, we run the market analyses for a range of values. More specifically, we consider elasticities from -0.1 to -0.6. This is in line with Bard and Loncar (1991), who estimated the elasticity of coking coal demand to lie in the range from -0.15 to -0.5, with Western European (Asian) demand elasticity lying in the lower (upper) part of this range. Graham et al.

[45]Choosing a linear functional form is a simplification of the real, unobservable demand function. It implies that the absolute price reaction to a specific absolute change in coking coal output is constant. The price elasticity, however, is not constant and depends on the price/demand combination. The stated elasticities refer to the elasticity at the reference price/demand combination. The choice of a linear demand function simplifies solving the model, particularly the two-stage ones.

(1999) finds that for 1996, a demand elasticity of -0.3 characterises best the actual market outcomes, whereas Trüby (2013) concludes that for the years 2008 to 2010, demand elasticity falls in the range from -0.3 to -0.5.

3.5 Results

In this section, the model results are presented and discussed. We start out by comparing the prices under the different market settings to the actual market prices. This allows us to narrow down the range of elasticities we need to focus on. In a second step, we use three statistical measures, namely a linear regression test as suggested by Bushnell et al. (2008), Spearman's rank correlation coefficient, and Theil's inequality coefficient, to assess whether trade flows simulated under different market structures match actual trade flows. Finally, revenues and production volumes of the Big-Four are analysed.

3.5.1 Prices

Figure 3.2 displays the actual FOB benchmark in 2008 (straight black line) as well as the simulated FOB prices for a range of elasticities (-0.1 to -0.6) and for the four market structure settings analysed in this paper. Four observations can be made: First, for very low elasticities, i. e., between -0.1 and -0.2, none of the market settings is able to reproduce actual market prices. Although only the results for 2008 are displayed in Figure 3.2, taking a look at the other years (see Figure 3.6 in Appendix C) confirms this conclusion.

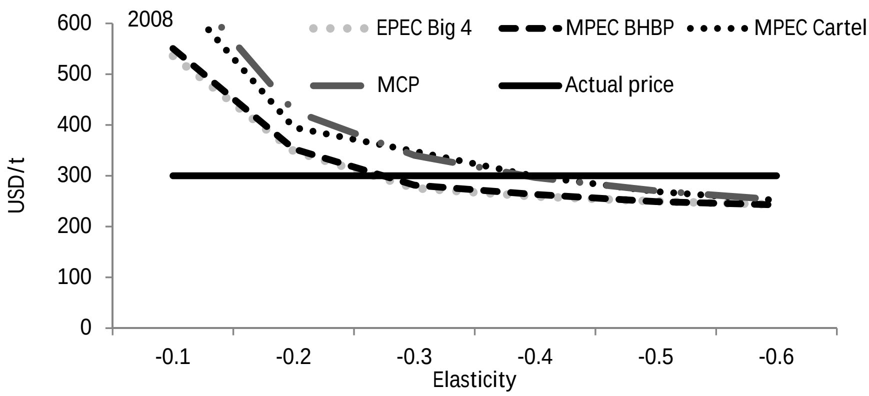


FIGURE 3.2: FOB Prices for a range of (abs.) elasticities – model results vs. actual benchmark price

Second, prices in the multi-leader-follower setting, EPEC Big 4, as well as in the setting in which BHP Billiton acts as a Stackelberg leader, MPEC BHPB, are more or less equivalent. This result is caused by the interaction of three effects (our argumentation follows Daughety (1990)): First, each following firm that becomes a Stackelberg leader has the incentive to increase its output since, now, it takes into account the optimal reaction of the remaining followers to a change in the output of the Stackelberg leaders. Second, increasing the number of leaders causes the output of each (incumbent) leader to drop. This may be interpreted as the result of the intensifying Cournot competition between the leaders. Third, the total output of the followers decreases with each firm becoming a Stackelberg leader. In our simulations, these effects seem to counterbalance each other, which is why the two market settings, EPEC Big 4 and MPEC BHPB, result in similar market outputs and prices.

Third, another interesting aspect is that (for low demand elasticities) prices for the case in which the Big-Four form a cartel that acts as a Stackelberg leader (labelled MPEC Cartel) are below the prices in the Cournot oligopoly (MCP).[46] In other words, the output-increasing effect of becoming a leader is stronger than the output-decreasing effect of collusion (forming the cartel). Building on Shaffer (1995), the intuition behind this finding can be explained as follows: For the case of N identical firms, zero marginal costs and a linear demand, the output of a cartel with k-members that acts as a Stackelberg leader is higher than in a Cournot oligopoly for k lower than $\frac{N+1}{2}$, but is decreasing in k. In other words, the bigger the cartel becomes, the more dominant the output-reducing collusion effect.[47] This is also in line with the results for the case in which BHPB acts as single leader (MPEC BHPB).

Finally, the higher the elasticity, the more the simulated prices converge. This can be explained by two effects: First, with increasing elasticity, total production increases as well (along with decreasing prices). As such, the capacity utilisation over all players increases from a minimum of 79 % (MCP, eta -0.1) to around 97 % (all scenarios with eta -0.6) for 2008. This narrows the ability to differentiate strategic behaviour as more players produce at their capacity limit. Second, increased price elasticity of demand itself narrows the potential for strategic choice of production as prices react more severely to changes in output.

Consequently, we conclude that the range of elasticities may be narrowed down to the range of -0.3 to -0.5, which is in line with previous analyses (see Section 3.4).

[46]For higher demand elasticities (i. e., larger than -0.3), prices of both cases are identical (given the tolerance of the applied linearisation method).

[47]In the case of $k = N$, i. e., the cartel consists of all firms N in the market, the price in the market would equal the monopoly price.

3.5.2 Trade flows

In a first step, we investigate whether simulated trade flows under the different market structures match the actual market outcomes by regressing the former on the latter. If the two were a perfect match, then the estimated linear equation would have a slope of one and an intercept of zero. Table 3.3 shows the p-values of the F-test that checks whether the coefficient of the slope and the intercept jointly equal one and zero, respectively, for six different elasticities and the four market structures.[48]

Taking a closer look at Table 3.3, we can conclude that all four market settings provide a reasonable fit with actual trade flows in the relevant range of elasticities (-0.3 to -0.5). This finding generally holds true for lower elasticities as well, with one exception. In the case of the MCP scenario, trade flows in 2008 and 2010 for an elasticity of -0.1 and in 2009 for an elasticity of -0.1 and -0.2 do not seem to provide a reasonable fit since the H_0-hypothesis is rejected. It should, however, be noted that 2009 was special in the sense that it was characterised by a significant drop in utilisation rates of the mines since steel demand and, thus, demand for coking coal plummeted compared to the previous year because of the financial crisis.

TABLE 3.3: P-values of the F-tests ($\beta_0 = 0$ and $\beta_1 = 1$) for a range of elasticities

Elasticity	EPEC Big 4			MPEC BHPB		
	2008	2009	2010	2008	2009	2010
e = -0.1	0.86	0.86	0.64	0.86	0.85	0.68
e = -0.2	1.00	0.80	0.90	1.00	0.81	0.92
e = -0.3	0.92	0.57	0.98	0.92	0.57	0.99
e = -0.4	0.85	0.44	0.95	0.84	0.46	0.97
e = -0.5	0.74	0.48	0.91	0.73	0.50	0.92
e = -0.6	0.59	0.52	0.84	0.59	0.52	0.85
Elasticity	MPEC Cartel			MCP		
	2008	2009	2010	2008	2009	2010
e = -0.1	0.79	0.76	0.70	0.08*	0.02**	0.06*
e = -0.2	1.00	0.66	0.12	0.22	0.09*	0.16
e = -0.3	0.43	0.45	0.37	0.43	0.25	0.34
e = -0.4	0.75	0.85	0.73	0.67	0.52	0.59
e = -0.5	0.78	0.49	0.92	0.77	0.73	0.81
e = -0.6	0.57	0.40	0.85	0.61	0.90	0.84

Significance levels: 1% '***' 5% '**' 10% '*'

In order to cross-check the results from the linear hypothesis test, two additional indicators are taken into consideration. Figure 3.3 depicts Spearman's rank correlation and Theil's inequality coefficient for the different market settings and the whole range

[48]See Appendix B for more details on the methodology used in this subsection.

of elasticities in 2008.[49] Both coefficients confirm the analysis of the linear hypothesis test since neither of the two indicators allows us to discard one of the market settings when looking at the relevant range of elasticities.

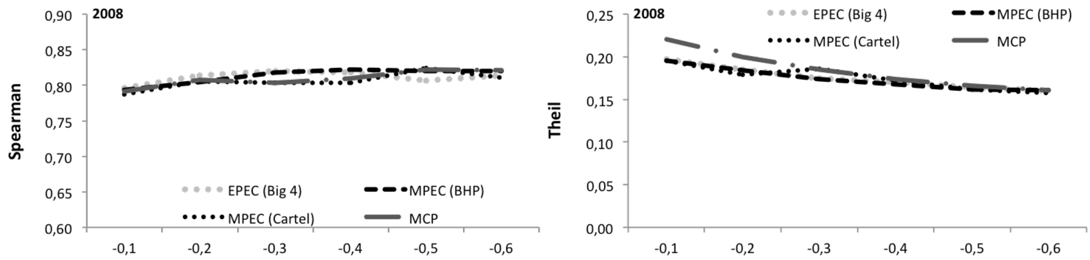


FIGURE 3.3: Spearman's correlation coefficients and Theil's inequality coefficients for a range of (abs.) elasticities

3.5.3 Production and revenues of the Big-Four

So far the conducted analyses have not provided significant evidence that one of the market structures investigated in this paper performs better or worse than another. Therefore, we take a closer look at two further components: revenues and production volumes of the Big-Four.

When analysing the differences in profits of the Big-Four between the various market structures simulated in this paper, we can observe that, as expected, the Big-Four make the largest profits in the MPEC Cartel setting. However, relative differences between the different market structures are negligible ($< 1\%$), which becomes obvious when comparing the bars in Figure 3.4.[50]

Thus, the conclusion that can be drawn from this comparison is that the gains of forming and coordinating a cartel are small even when neglecting transaction costs that go along with maintaining the cartel.

Turning now to production, we compare the absolute difference in simulated versus actual production volumes of the Big-Four cumulated over the time period investigated in this paper (2008 to 2010). This indicator was chosen because it captures differences in the total production volumes of the Big-Four as well as deviations in each firm's production volumes. In addition, we compare the sum of squared differences between actual and modelled production to assess the structure of deviations. The resulting differences are depicted in Figure 3.5 for a demand elasticity of -0.4, which is the mean value of the range of elasticities found to be relevant (see Subsection 3.5.1). As can

[49]Conclusions remain unchanged when focusing on the other two years, as can be seen in Figure 3.7 in Appendix C.

[50]The results for 2009 and 2010 are similar.

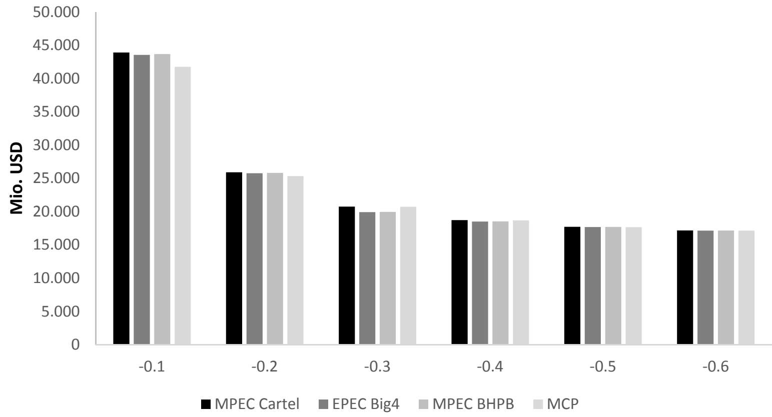


FIGURE 3.4: 2008's profits of the Big-Four in the three two-stage-games for the whole range of elasticities

be seen in the left diagram, cumulated absolute differences to historical data lie in the range of 8% to 17%, with the MPEC Cartel setting performing worst. On the other hand, the market structures in which BHP Billiton is the sole Stackelberg leader and the case of four non-colluding leaders perform best. Taking a closer look at the individual differences of the two settings with the largest differences, it becomes obvious that the MCP setting performs reasonably well in 2008 and 2010 but fails to reproduce the decline in production of the Big-Four in 2009. This is also the reason for the poor performance regarding squared deviations. In contrast, the MPEC Cartel setting constantly overestimates the production of BHP Billiton and underestimates the one of Rio Tinto, with the reason being that this minimizes the overall production costs of the cartel. In the two cases that perform best (MPEC BHPB and EPEG Big 4), we observe no significant patterns.

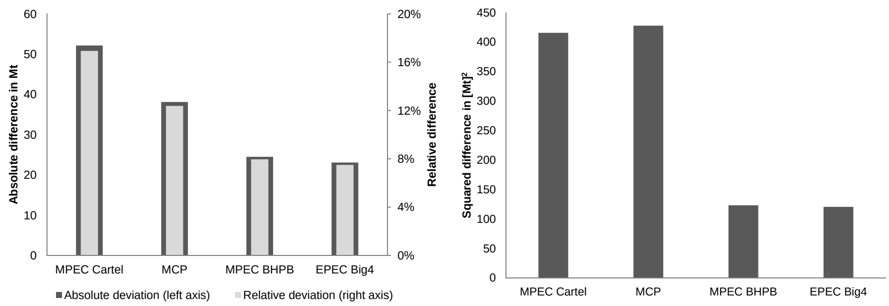


FIGURE 3.5: Cumulated absolute and squared difference in production volumes of the Big-Four to actual market outcomes at an elasticity of -0.4

In summary, three conclusions may be drawn from our analyses: i) We are able to support previous findings that the setting in which a cartel of the Big-Four acts as the Stackelberg leader, MPEC Cartel, as well as the Cournot oligopoly setting sufficiently reproduce actual trade flows and prices. ii) However, we also show that additional revenues from forming a cartel are rather small and individual production volumes of the Big-Four in the cartel setting do not match well with actual production numbers. Thus, we argue that a market structure with a cartel of the Big-Four that moves first is less likely than the other scenarios. iii) We find that the two settings with one or more leading firms reproduce actual trade flows and prices as good as the cartel and the Cournot settings. In addition, these two settings perform better than the former two settings with respect to the production volumes of the Big-Four. In particular, the methodology introduced in this paper to represent multi-leader-follower games scored among the best results in all tests used in our analysis.

3.6 Conclusions

Previous analyses of the prevailing market structure in spatial resource markets mainly focussed on the comparison of actual market outcomes to market results under perfect competition, Cournot competition and with a single (Stackelberg) leader. We add to these analyses by developing a model able to represent multi-leader market structures. We apply our model to the metallurgical coal market, which is especially suited as its market structure suggests a multitude of possible markets structures that have partly been neglected in previous analyses. Thereby, we are able to demonstrate the practicability and usefulness of our approach.

Trüby (2013) shows that market results of the metallurgical coal market indicate non-competitive behaviour. Actual prices and trade flows could rather be explained by Cournot competition or a game in which the Big-Four form a cartel that acts as a single Stackelberg leader. Our results confirm that a Cournot oligopoly as well as a cartel consisting of the Big-Four fit well with observed prices and trade flows of the metallurgical coal market from 2008 to 2010. Based on our results, however, the same is true for two additional settings: First, a market with BHPB acting as a Stackelberg leader and the remaining players competing afterwards in a Cournot fashion (MPEC BHBP). Second, a multi-leader market structure where the Big-Four independently act first followed by the remaining players (EPEC Big 4). By additionally analysing profits and comparing the actual production data with models results, we conclude that the two latter scenarios are even more likely than the previously suggested market structures.

For 2009, in which overall demand has been low, model outcomes, in particularly concerning prices, do not fit as well as for the other two years. Hence, taking into account market power on the demand side – as suggested by Graham et al. (1999) – might be more appropriate than the market settings analysed in this paper.

To improve the accuracy of current market structure analyses and to further narrow down the set of potential market structures, it could be useful to have more detailed firm and market data, also for smaller market participants. In order to be able to solve the computationally challenging non-linear bi-level games, we had to aggregate our dataset. Improving available solution methods for these problems to obtain mine-by-mine results may help to discriminate between the goodness of fit of different model results with actual market data. However, this would require detailed data availability. Unfortunately, neither mine-by-mine market results nor detailed profitability data on a firm level were available.

Our results demonstrate the multiplicity of possible market structures able to explain actual market outcomes concerning trade flows and market prices. By analysing the production data, we were able to identify two promising candidates for the underlying market structure. However, we are aware of the fact that the market structures analysed in this paper may not cover the whole range of potentially interesting settings, e. g., as indicated by Graham et al. (1999), the demand side could be exerting market power as well.

From this findings, two conclusions can be drawn: First, omitting potential scenarios can lead to false conclusions of the prevailing market structure. This is relevant especially when it comes to judging if market outcomes reveal collusive behaviour. Second, a market structure analysis solely based on market outcomes with respect to price and trade flows may not be sufficient to determine the actual market structure but should rather be completed using additional analyses.

These conclusions lead to the following subjects for future research: First, expanding the range of market settings under consideration, in particular including market power on the demand side, could give additional insights. This is especially relevant given the varying observable levels of demand in different years. Second, including more years could strengthen the explanatory power of the findings and eventually help to identify changing market structures over time. Third, expanding the model to include investment decisions could further strengthen the understanding of current markets and their development as research on this topic is rather thin.[51]

[51] See Lorenczik et al. (2017) for an analysis for the coking coal market that includes investment decisions.

3.7 Appendix

3.7.1 Appendix A: Input data

Table 3.4: Reference demand [Mt] and price [US$/t]

	2008		2009		2010	
	demand	price	demand	price	demand	price
JP_KR	80	300	71	129	87	227
CN_TW	10	300	26	129	42	227
IN	26	300	26	129	35	227
LAM	16	300	15	129	17	227
EUR_MED	63	300	43	129	58	227
Other	18	300	10	129	7	227

Table 3.5: Production costs [US$/t]

	2008	2009	2010
Australia	67	71	73
Canada	100	101	104
China	91	114	117
Indonesia	110	112	113
New Zealand	72	73	75
Russia	162	163	156
South Africa	51	52	53
USA	117	108	113
Anglo American	67	69	70
BHP Billiton	76	77	80
Rio Tinto	78	79	82
Xstrata	63	65	67

Table 3.6: Production capacities [Mtpa]

	2008	2009	2010
Australia	37.4	34.4	42.6
Canada	25.6	28.0	28.0
China	4.0	2.1	2.1
Indonesia	2.1	2.1	2.5
New Zealand	2.6	2.6	2.6
Russia	15.2	15.5	15.5
South Africa	0.8	0.8	0.8
USA	52.2	57.2	60.2
Anglo American	15.1	15.1	16.3
BHP Billiton	63.6	63.6	71.4
Rio Tinto	15.0	15.0	16.2
Xstrata	13.2	14.5	15.0

Table 3.7: Shipping costs [US$/t]

	CN_TW			EUR_MED			IN		
	2008	**2009**	**2010**	**2008**	**2009**	**2010**	**2008**	**2009**	**2010**
Australia	24.7	13.8	15.9	42.9	18.9	20.9	29.9	15.4	17.5
Canada	30.5	15.6	17.6	37.6	17.6	19.6	37.1	17.4	19.5
China	15.2	10.5	12.4	41.8	18.6	20.6	26.5	14.4	16.4
Indonesia	17.9	11.5	13.5	39.9	18.2	20.2	23.5	13.4	15.5
New Zealand	29.6	15.3	17.4	42.5	18.8	20.8	32.3	16.1	18.2
Russia	16.7	11.1	13.1	16.5	11.0	13.0	27.4	14.7	16.7
South Africa	31.6	15.9	18.0	32.7	16.2	18.3	25.1	14.0	16.0
USA	41.7	18.6	20.6	23.7	13.5	15.6	37.8	17.6	19.6

	JP_KR			LAM			Other		
	2008	**2009**	**2010**	**2008**	**2009**	**2010**	**2008**	**2009**	**2010**
Australia	24.8	13.9	15.9	36.2	17.2	19.2	33.7	16.5	18.5
Canada	26.4	14.4	16.4	36.4	17.2	19.3	41.2	18.5	20.5
China	15.1	10.4	12.4	42.5	18.8	20.8	32.1	16.0	18.1
Indonesia	22.2	13.0	15.0	37.7	17.6	19.6	26.9	14.5	16.6
New Zealand	29.2	15.2	17.3	32.3	16.1	18.1	36.2	17.2	19.2
Russia	12.4	9.3	11.2	33.0	16.3	18.4	27.2	14.6	16.7
South Africa	34.9	16.8	18.9	26.0	14.2	16.3	26.2	14.3	16.4
USA	39.2	18.0	20.0	27.9	14.8	16.9	36.5	17.3	19.3

3.7.2 Appendix B: Statistical measures[52]

In order to assess the accuracy of our model, we compare market outcomes such as production, prices and trade flows to our model results. In comparing trade flows, we follow, for example, Kolstad and Abbey (1984), Bushnell et al. (2008) and more recently Trüby (2013) as well as Hecking and Panke (2014) by applying three different statistical measures: a linear hypothesis test, the Spearman rank correlation coefficient and Theil's inequality coefficient. In the following, we briefly discuss the setup as well as some of the potential weakness of each of the three tests.

Starting with the linear hypothesis test, if the actual and model trade flows had a perfect fit, the dots in a scatter plot of the two data sets would align along a line starting at zero and have a slope equal to one. Therefore, we test model accuracy by regressing actual trade flows A_t on the trade flows of our model M_t, with t representing the trade flow between exporting country $e \in E$ and importing region $d \in D$, as data on trade flows is available only on a country level. Using ordinary least squares (OLS), we estimate the following linear equation:

$$A_t = \beta_0 + \beta_1 * M_t + \epsilon_t.$$

Modelled trade flows have a bad fit with actual data if the joint null hypothesis of $\beta_0 = 0$ and $\beta_1 = 1$ can be rejected at typical significance levels. One of the reasons why this test is applied in various studies is that it allows hypothesis testing, while the other two tests used in this paper are distribution-free and thus do not allow such testing. However, there is a drawback to this test as well, since the results of the test are very sensitive to how good the model is able to simulate outliers. To improve the evaluation of the model accuracy regarding the trade flows, we apply two more tests.

The second test we employ is the Spearman's rank correlation coefficient, which, as already indicated by its name, can be used to compare the rank by volume of the trade flow t in reality to the rank in modelled trade flows. Spearman's rank correlation coefficient, also referred to as Spearman's rho, is defined as follows:

$$rho = 1 - \sum_{t}^{T} d_t^2/(n^3 - n)$$

with $d_{i,j}$ being the difference in the ranks of the modelled and the actual trade flows and T being the total number of trade flows. Since Spearman's rho is not based on a distribution, hypothesis testing is not applicable. Instead, one looks for a large value

[52] This section has already been published in Hecking and Panke (2014), which is co-authored by one of the authors of this paper.

of *rho*. However, Spearman's rank correlation coefficient does not tell you anything about how well the predicted trade flows compare volumewise to the actual trade flow volumes. For example, *rho* could be equal to one despite total trade volume being ten times higher in reality as long as the market shares of the trade flows match.

Finally, we apply the normed-version of Theil's inequality coefficient U, which lies between 0 and 1, to analyse the differences between actual and modelled trade flows. A U of 0 indicates that modelled trade flows perfectly match actual trade flow, while a large U hints at a large difference between the two data sets. Theil's inequality coefficient is defined as:

$$U = \frac{\sqrt{\sum_t^T (M_t - A_t)}}{\sqrt{\sum_t^T M_t^2} + \sqrt{\sum_t^T A_t^2}} .$$

3.7.3 Appendix C: Prices and statistical measures for trade flows

Prices

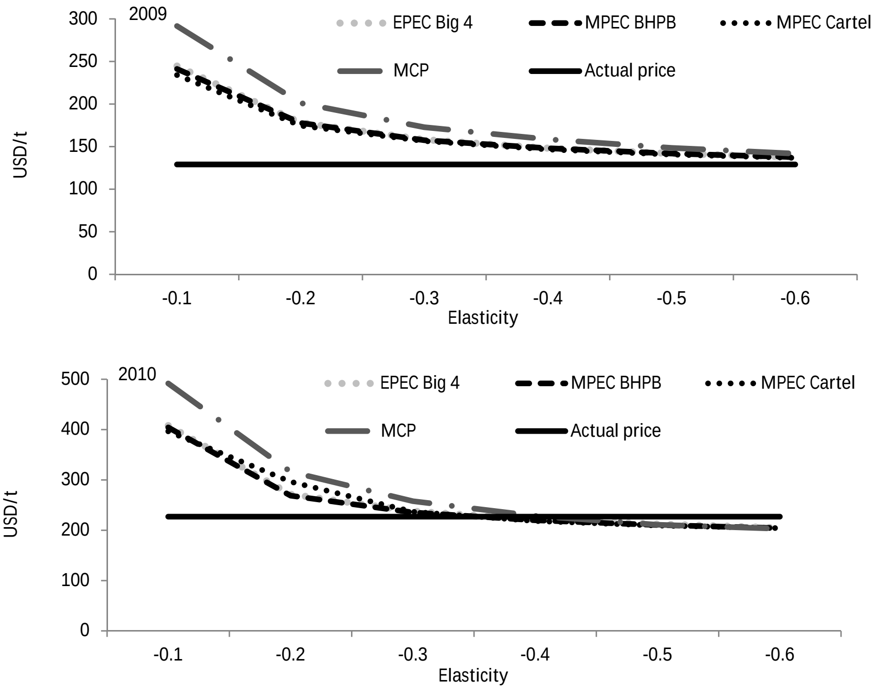


FIGURE 3.6: FOB Prices for a range of (abs.) elasticities – model results vs. actual benchmark price

Statistical measures for trade flows

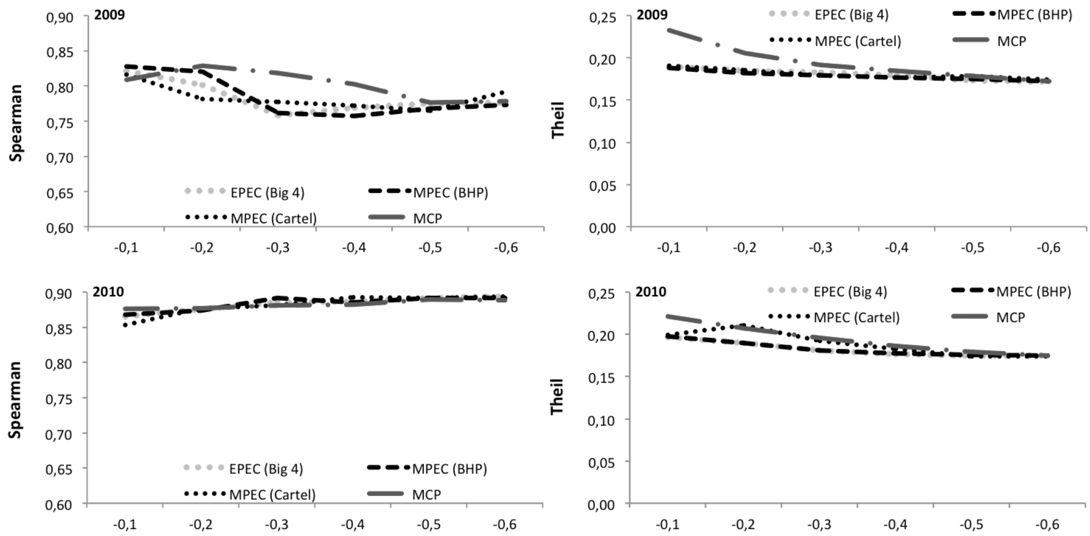


FIGURE 3.7: Spearman's correlation coefficients and Theil's inequality coefficients for a range of (abs.) elasticities

Modeling strategic investment decisions in spatial markets

Markets for natural resources and commodities are often oligopolistic. In these markets, production capacities are key for strategic interaction between the oligopolists. We analyze how different market structures influence oligopolistic capacity investments and thereby affect supply, prices and rents in spatial natural resource markets using mathematical programing models. The models comprise an investment stage and a supply stage in which players compete in quantities. We compare three models, a perfect competition and two Cournot models, in which the product is either traded through long-term contracts or on spot markets in the supply stage. Tractability and practicality of the approach are demonstrated in an application to the international metallurgical coal market. Results may vary substantially between the different models. The metallurgical coal market has recently made progress in moving away from long-term contracts and more towards spot market-based trade. Based on our results, we conclude that this regime switch is likely to raise consumer rents but lower producer rents, while the effect on total welfare is negligible.

4.1 Introduction

Markets for natural resources and commodities such as iron ore, copper ore, coal, oil or gas are often highly concentrated and do not appear to be competitively organized at first glance. In such markets, large companies run mines, rigs or gas wells and trade their product globally. In the short term, marginal production costs and capacities are given and determine the companies' competitive position in the oligopolistic market. However, in the longer term, companies can choose their capacity and consequently alter their competitive position.

Investing in production capacity is a key managerial challenge and determining the right amount of capacity is rarely trivial in oligopolistic markets. Suppliers have to take competitors' reactions into account not only when deciding on the best supply level but also when choosing the best amount of capacity.

In this paper, we introduce three different models to address this capacity expansion problem in oligopolistic natural resource markets under varying assumptions of market structure and conduct. Moreover, we pursue the question as to how different market structures influence capacity investments, supply, prices and rents. The models comprise two stages: an investment stage and a supply stage in which players compete in quantities. We explicitly account for the spatial structure of natural resource markets, i. e., demand and supply regions are geographically separated and market participants incur distance-dependent transportation costs.

The first model assumes markets to be contestable; hence investment follows competitive logic. Solving this model yields the same result as would be given by a perfectly competitive market. The second model assumes the product to be sold through long-term contracts under imperfect competition. Even though supply takes place in stage two, the supply and investment decisions are made simultaneously in stage one. The long-term contract that is fulfilled in stage two determines the level of capacity investment in stage one. Any production capacity that is different from the one needed to produce the quantity of the best-supply equilibrium in stage two reduces the respective players profits and is not a Nash equilibrium. The outcome is termed 'open-loop Cournot equilibrium' and corresponds to the result of a static one-stage Cournot game (accounting for investment costs). The third model assumes that investment and supply decisions are made consecutively: In stage one, when investment takes place, none of the oligopolists can commit to their future output decision in stage two (unlike in the open-loop case). In stage two, when the market clears, the investment cost spent in the first stage is sunk and the players base their output decision solely on production cost. The resulting equilibrium is termed 'closed-loop Cournot equilibrium' and may differ from the open-loop outcome.

Intuitively, the lack of commitment in the closed-loop game and therefore the repeated interaction of the oligopolists would suggest a higher degree of competition and thus lower prices and higher market volumes than in the open-loop equilibrium. However, the players anticipate this strategic effect and make their investment decisions accordingly. How prices and volumes rank compared to the open-loop game is parameter-dependent and requires a numerical analysis. As discussed for instance in Fudenberg and Tirole (1991) in a more general context, each player in the closed-loop model has a strategic incentive to deviate from his first stage open-loop action as he can thereby influence the other players' second stage action. Applying this general economic framework to the capacity expansion problem examined in this paper, indeed tends to lead to higher investment and supply levels in the closed-loop model and hence to lower prices.

Computing open-loop games is relatively well understood, and existence and uniqueness of the equilibrium can be guaranteed under certain conditions (see, e. g., Harker, 1984, 1986, Takayama and Judge, 1964, 1971). The open-loop Cournot model can be solved via the Karush-Kuhn-Tucker conditions as a mixed complementarity problem (MCP). Oligopolistic spatial equilibrium models have been widely deployed in analyzing resource markets, without taking investments decisions into account, e. g., for steam coal markets (Haftendorn and Holz, 2010, Kolstad and Abbey, 1984, Trüby and Paulus, 2012), metallurgical coal markets (Graham et al., 1999, Trüby, 2013), natural gas markets (Gabriel et al., 2005b, Growitsch et al., 2013, Holz et al., 2008, Zhuang and Gabriel, 2008), wheat markets (Kolstad and Burris, 1986), oil markets (Huppmann and Holz, 2012) or for iron ore markets (Hecking and Panke, 2014). Investments in additional production capacity have been analyzed for example in Huppmann (2013) with investment and production decisions being made simultaneously and therefore implicitly assuming a market structure with long-term contracts.

Closed-loop models are computationally challenging due to their non-linear nature. Depending on the problem this can be resolved. Gabriel and Leuthold (2010) for instance model an electricity market with a Stackelberg leader using linearization to guarantee a globally optimal solution. Closed-loop models in energy market analysis have primarily been used to study restructured electricity markets (e. g., Daxhelet and Smeers, 2007, Shanbhag et al., 2011, Yao et al., 2008, 2007). Murphy and Smeers (2005) and Wogrin et al. (2013a,b) have analyzed the implications of closed- and open-loop modeling on market output and social welfare as well as characterized conditions under which closed- and open-loop model results coincide.

Our two-stage model consists of multiple players on both, the first and second stage (investment in stage one and supply in stage two), and therefore existence and uniqueness of (pure strategy) equilibria cannot be guaranteed. The closed-loop model, which is formulated as an Equilibrium Problem with Equilibrium Constraints (EPEC), is implemented using a diagonalization approach (see, e. g., Gabriel et al., 2012). In doing so, we reduce the solution of the EPEC to the solution of a series of Mathematical Programs with Equilibrium Constraints (MPEC). Concerning the solution of the MPECs we implement two algorithms, grid search along the investment decisions of the individual players and a Mixed Integer Linear Program reformulation following Wogrin et al. (2013a).

We demonstrate the tractability and practicality of our investment models in an application to the international metallurgical (or coking) coal trade. Metallurgical coal is, due to its special chemical properties, a key input in the process of steel-making. The market for this rare coal variety is characterized by a spatial oligopoly with producers mainly located in Australia, the United States and Canada competing against each other and providing the bulk of the traded coal (Bowden, 2012, Trüby, 2013). The players hold existing mining capacity and can invest into new capacity. Investment and mining costs differ regionally. Key uncertainties in this market are demand evolution and price responsiveness of demand. We therefore compute sensitivities for these parameters to demonstrate the robustness of our results.

Our findings are generally in line with previous results found in the literature on two-stage games with players choosing capacity and output, i. e., we find that prices and supply levels in the closed-loop game fall between those in the perfect competition and the open-loop game (see, e. g., Murphy and Smeers, 2005). If investment costs are low compared to variable costs of supply, the strategic effect of the two-stage optimization in the closed-loop game diminishes. With investment costs approaching zero, the closed-loop result converges to the open-loop result. Hence, the closed-loop model is particularly useful for capital-intensive natural resource industries in which the product is traded on spot markets.

The numerical results for supply levels, prices and rents in the metallurgical coal market analysis differ markedly between the three models. Consistent with actual industry investment pipelines, our model suggests that the bulk of the future capacity investment comes from companies operating in Australia followed by Canadian and US firms. Starting in 2010, the metallurgical coal market has undergone a paradigm shift, moving away from long-term contracts and more towards a spot market-based trade – with similar tendencies being observed in other commodity markets such as the iron ore trade. In light of our findings, this effect is detrimental to the companies' profits but beneficial to consumer rents. The effect on welfare is negligible: Gains in consumer rents and losses in producers' profits are of almost equal magnitude.

The contribution of this paper is threefold: First, by extending the multi-stage investment approach to the case of spatial markets, we introduce a novel feature to the literature on Cournot capacity expansion games. Second, we outline how our modeling approach can be implemented and solved to analyze capacity investments in natural resource markets. We thereby extend previous research on natural resource markets, which has typically assumed capacities to be given. Finally, we illustrate and discuss the model properties on the basis of a real-world application to the international metallurgical coal trade and draw conclusions for this market. In doing so, we also take into account existing capacities of the players and hence incorporate a feature which to our knowledge has been ignored in previous work on multi-stage Cournot capacity expansion games. By comparing open- and closed-loop model results, we illustrate possible consequences of the ongoing regime switch from long-term contracts to a more spot market-based trade in the international metallurgical coal market. Our analysis in particular allows for the first quantification of the magnitude of the divergence between open- and closed-loop model results in a real-world application.

The remainder of the paper is structured as follows: Section 4.2 describes the models developed in this paper and Section 4.3 provides details about their implementation. The data is outlined in Section 4.4, results are presented in Section 4.5. Section 4.6 discusses computational issues and Section 4.7 concludes.

4.2 The Model

We introduce three different approaches to the capacity expansion problem – two open-loop models and a closed-loop model. In the open-loop models, all players decide simultaneously on their investment *and* production levels, whereas in the closed-loop model all players *first* decide on their *investment* levels simultaneously and *then*, based on observed investment levels, they simultaneously decide on their *production* levels. The two open-loop models vary in their underlying market structure: one model assumes perfect competition, the other model assumes Cournot competition with a competitive fringe. The closed-loop model also assumes Cournot competition with a competitive fringe.

While similar open-loop models have previously been studied, the introduced closed-loop model varies from existing closed-loop models by taking into account also the spatial structure of the market as well as considering existing capacities of the players.

4.2.1 General Setting and Notations

Table 4.1 summarizes the most relevant nomenclature used throughout this section. Additional symbols are explained where necessary. We assume a spatial, homogeneous good market consisting of producers $i \in I$, production facilities $m \in M$ and demand regions $j \in J$. Each producer i owns production facilities $m \in M_i \subset M$. Furthermore, we assume that $M_i \cap M_j = \emptyset$ for $i \neq j$, i. e., production facilities are exclusively owned by one producer. Producers decide on both their investment in production facilities as well as on their supply levels.

As in equilibrium added capacities are fully utilized, no stock constraint for new capacities is modeled. Therefore we implicitly assume that mines will be exhausted after their depreciation period (see Section 4.4).

The supply from production facility m to market j is given by $x_{m,j}$. Total production of production facility m is hence given by $\sum_j x_{m,j}$. It is limited by the facilities' capacity $cap_m^0 + y_m$, where cap_m^0 is the initial production capacity and y_m denotes the capacity investment. Capacity investments y_m are non-negative and limited by y_m^{max}. The upper bound on capacity expansion is chosen sufficiently high not to impose restrictions on economically favorable investments but is rather used to ease the solution algorithm (the upper limit restricts the solution space of the non-linear MPEC and enables the equidistant separation of investments in the case of the line search, see Section 4.3).

Capacity investments in an existing production facility (i. e., $cap_m^0 \neq 0$) can be interpreted as capacity expansions, and investments in the case of $cap_m^0 = 0$ as newly built production facilities.

Investment expenditures for facility m are given by C_m^{inv}. We assume that C_m^{inv} is a linear function in the investment level y_m, with k_m denoting marginal investment costs, i. e.,

$$C_m^{inv}(y_m) = k_m \cdot y_m.$$

Variable costs C_m^{var} are specific to the production facility m. They are composed of transportation costs $\tau_{m,j}$ per unit delivered from m to market j as well as the variable production costs v_m. We assume that v_m is a linear function in the total production of the facility. Total variable costs of facility m therefore amount to

$$C_m^{var}(\mathbf{x}_m) = \sum_j (x_{m,j} \cdot \tau_{m,j}) + v_m(\sum_j x_{m,j}),$$

with $\mathbf{x}_m = (x_{m,j})_j$ denoting the production vector of facility m.

Market prices P_j in market j are given by a linear inverse demand function, i. e.,

$$P_j = a_j - b_j \cdot \sum_m x_{m,j}.$$

TABLE 4.1: Model sets, parameters and variables

Abbreviation	Description
Model sets	
$m \in M$	Production facilities
$j \in J$	Markets
$i \in I$	Players
Model parameters	
k_m	Marginal investment costs [US$ per unit per year]
v_m	Variable production costs [US$ per unit]
$\tau_{m,j}$	Transportation costs [US$ per unit]
a_j	Reservation price [US$ per unit]
b_j	Linear slope of demand function
cap_m^0	Initial production capacity [units per year]
y_m^{max}	Maximum capacity expansion [units per year]
Model variables	
C_m^{var}	Total variable production costs [US$]
C_m^{inv}	Investment expenditures [US$]
$x_{m,j}$	Supply [units]
P_j	Market price [US$ per unit]
y_m	Capacity investments [units per year]

4.2.2 Model 1: The Open-Loop Perfect Competition Model

In the open-loop perfect competition model (in the following simply termed 'perfect competition model'), each producer $i \in I$ solves the optimization problem

$$\max_{\mathbf{x}_m, y_m : m \in M_i} \sum_{m \in M_i} \Big(\sum_{j \in J} P_j \cdot x_{m,j} - C_m^{var}(\mathbf{x}_m) \Big) - \sum_{m \in M_i} C_m^{inv}(y_m)$$

subject to

$$\begin{aligned}
&P_j = a_j - b_j \cdot (X_{i,j} + X_{-i,j}), \ \forall j \\
&cap_m^0 + y_m - \sum_j x_{m,j} \geq 0, \ \forall m \in M_i \ (\lambda_m) \\
&y_m^{max} - y_m \geq 0, \ \forall m \in M_i \ (\theta_m) \\
&x_{m,j} \geq 0, \ \forall m \in M_i, j \\
&y_m \geq 0, \ \forall m \in M_i
\end{aligned}$$

while taking the supplies $X_{-i,j}$ of the other producers $(-i)$ as given. Here and in the following, we use the abbreviation $X_{I_1,j} = \sum_{i \in I_1} \sum_{m \in M_i} x_{m,j}$ for some $I_1 \subset I$.

Hence, in the perfect competition model, each producer *simultaneously* makes his ("long-term") investment and ("short-term") production decisions in order to maximize profits. In doing so, each producer takes capacity restrictions into account. However, players do not take into account their influence on price.

Any solution to the above optimization problem has to satisfy the short-term Karush-Kuhn-Tucker (KKT) conditions

$$
\begin{aligned}
&0 \leq \frac{\partial C_m^{var}(\mathbf{x}_m)}{\partial x_{m,j}} - [a_j - b_j \cdot (X_{i,j} + X_{-i,j})] + \lambda_m \perp x_{m,j} \geq 0, \ \forall i, m \in M_i, j \\
&0 \leq cap_m^0 + y_m - \sum_j x_{m,j} \perp \lambda_m \geq 0, \ \forall i, m \in M_i
\end{aligned}
$$

as well as the long-term KKT conditions

$$
\begin{aligned}
&0 \leq k_m - \lambda_m + \theta_m \perp y_m \geq 0, \ \forall i, m \in M_i \\
&0 \leq y_m^{max} - y_m \perp \theta_m \geq 0, \ \forall i, m \in M_i.
\end{aligned}
$$

In equilibrium, all KKT conditions have to hold simultaneously. Uniqueness of the solution is guaranteed due to the quasi-concave objective function and the convexity of the restrictions. The derived KKT conditions are thus necessary and sufficient for obtaining the solution.

4.2.3 Model 2: The Open-Loop Cournot Model with Competitive Fringe

In the open-loop Cournot model with competitive fringe (in the following simply termed 'open-loop model'), each producer $i \in I$ solves an optimization problem identical to the one for the perfect competition model described above. However, each producer may take additionally into account his influence on price which is represented by the conjectural variation parameter ψ_i, where $\frac{\partial P_j}{\partial x_{m,j}} = \psi_i \cdot b_j$ for all $m \in M_i$. Cournot behavior with a competitive fringe can then be represented as $\psi_i = 1$ for the Cournot players and $\psi_i = 0$ for the competitive fringe.[53]

Any solution to the open-loop Cournot model with competitive fringe then satisfies the short-term Karush-Kuhn-Tucker (KKT) conditions

$$
\begin{aligned}
&0 \leq \frac{\partial C_m^{var}(\mathbf{x}_m)}{\partial x_{m,j}} - [a_j - b_j \cdot (X_{i,j} + X_{-i,j})] + \psi_i \cdot b_j \cdot X_{i,j} + \lambda_m \perp x_{m,j} \geq 0, \ \forall i, m \in M_i, j \\
&0 \leq cap_m^0 + y_m - \sum_j x_{m,j} \perp \lambda_m \geq 0, \ \forall i, m \in M_i
\end{aligned}
$$

as well as the long-term KKT conditions

$$
\begin{aligned}
&0 \leq k_m - \lambda_m + \theta_m \perp y_m \geq 0, \ \forall i, m \in M_i \\
&0 \leq y_m^{max} - y_m \perp \theta_m \geq 0, \ \forall i, m \in M_i.
\end{aligned}
$$

[53] The perfect competition model also follows from this specification by setting $\psi_i = 0$ for all i.

In equilibrium, the KKT conditions of both the Cournot players and the competitive fringe have to hold simultaneously. As in the perfect competition case, uniqueness of the solution is guaranteed due to the quasi-concave objective function and the convexity of the restrictions. The derived KKT conditions are therefore again necessary and sufficient for obtaining the solution.

4.2.4 Model 3: The Closed-Loop Model

In the closed-loop model, producers play a two-stage game: In the first stage, oligopolistic producers l ($l \in L \subset I$) decide on their investment levels. In the second stage, they choose, based on *observed* investment decisions of the other oligopolistic producers, their production and supply levels. In addition, in the second stage, a further player, the competitive fringe (F), makes his supply decisions. The competitive fringe is not allowed to invest in either stage.[54] As opposed to the oligopolistic producers, the competitive fringe is a price taker.

4.2.4.1 The Second Stage Problem

For a given investment vector (y_l, y_{-l}) of the oligopolistic producers, let the second stage problem of producer i be given by

$$\max_{x_{m,j}:m\in M_i} \sum_{m\in M_i} \big(\sum_{j\in J} P_j \cdot x_{m,j} - C_m^{var}(\mathbf{x}_m)\big)$$

subject to

$$\begin{aligned}
&P_j = a_j - b_j \cdot (X_{i,j} + X_{-i,j}),\ \forall j \\
&cap_m^0 + y_m - \sum_j x_{m,j} \geq 0,\ \forall m \in M_i\ (\lambda_m) \\
&x_{m,j} \geq 0,\ \forall m \in M_i, j.
\end{aligned}$$

As in the open-loop model, producer i decides on his supplies while taking the supplies of the other producers ($-i$) as given. A producer's influence on price is again assumed to be represented by a conjectural variation parameter ψ_i, which is equal to one for the oligopolistic producers and zero for the competitive fringe. Note that the competitive fringe may not invest and therefore $y_m = 0$ for the fringe.

[54] In our application to the metallurgical coal market, this restriction also holds true for the player in the perfect competition model corresponding to the competitive fringe in the closed-loop model as well as for the competitive fringe in the Cournot open-loop model. For better readability, the model descriptions in the preceding two subsections are slightly more general, i. e., allowing potentially all players to invest.

The corresponding KKT conditions to this problem are then given by

$$0 \leq \frac{\partial C_m^{var}(\mathbf{x}_m)}{\partial x_{m,j}} - [a_j - b_j \cdot (X_{i,j} + X_{-i,j})] + \psi_i \cdot b_j \cdot X_{i,j} + \lambda_m \perp x_{m,j} \geq 0, \ \forall m \in M_i, j$$
$$0 \leq cap_m^0 + y_m - \sum_j x_{m,j} \perp \lambda_m \geq 0, \ \forall m \in M_i.$$

In the second stage equilibrium, the KKT conditions of all producers have to hold simultaneously. In the following, let $\tilde{x}_{m,j}(y_l, y_{-l})$ denote the second stage production equilibrium for a given investment vector (y_l, y_{-l}).

4.2.4.2 The First Stage Problem

The first stage problem for oligopolistic producer $l \in L$ is given by

$$\max_{y_m : m \in M_l} \sum_{m \in M_l} \big(\sum_{j \in J} \tilde{P}_j \cdot \tilde{x}_{m,j}(y_l, y_{-l}) - C_m^{var}(\tilde{\mathbf{x}}_m(y_l, y_{-l})) \big) - \sum_{m \in M_l} C_m^{inv}(y_m)$$

subject to

$$\tilde{P}_j = a_j - b_j \cdot (\tilde{X}_{l,j}(y_l, y_{-l}) + \tilde{X}_{-l,j}(y_l, y_{-l}) + \tilde{X}_{F,j}(y_l, y_{-l})), \ \forall j$$
$$y_m^{max} - y_m \geq 0, \ \forall m \in M_l$$
$$y_m \geq 0, \ \forall m \in M_l,$$

i. e., producer l chooses his investment levels in order to maximize profits for a given investment strategy of the other oligopolistic producers (y_{-l}) under consideration of the resulting second stage equilibrium outcome.

Combining the second stage and the first stage problem, we obtain the following MPEC for producer l, hereafter referred to as MPEC$_l$:

$$\max_{\Omega_l} \sum_{m \in M_l} \big(\sum_{j \in J} (a_j - b_j \cdot (X_{l,j} + X_{-l,j} + X_{F,j})) \cdot x_{m,j} - C_m^{var}(\mathbf{x}_m) \big) - \sum_{m \in M_l} C_m^{inv}(y_m)$$

subject to

$$y_m^{max} - y_m \geq 0, \ \forall m \in M_l$$
$$y_m \geq 0, \ \forall m \in M_l$$
$$0 \leq \frac{\partial C_m^{var}(\mathbf{x}_m)}{\partial x_{m,j}} - [a_j - b_j \cdot (X_{i,j} + X_{-i,j})] + \psi_i \cdot b_j \cdot X_{i,j} + \lambda_m \perp x_{m,j} \geq 0, \ \forall i, m \in M_i, j$$
$$0 \leq cap_m^0 + y_m - \sum_j x_{m,j} \perp \lambda_m \geq 0, \ \forall i, m \in M_i$$

given the investment vector (y_{-l}) of the other oligopolistic producers. Here, Ω_l is given by:[55]

$$\Omega_l = \{(y_m)_{m \in M_l}; \quad (x_{m,j}, \lambda_m)_{m \in M, j \in J}\}.$$

An investment strategy $(\tilde{y}_l, \tilde{y}_{-l})$ is a closed-loop equilibrium if for all $l \in L$, $\tilde{y}_l$ solves l's MPEC problem MPEC$_l$ given $\tilde{y}_{-l}$. The problem of finding a closed-loop equilibrium is hence of EPEC type (Gabriel et al., 2012), and therefore existence and uniqueness of equilibria typically is non-trivial and parameter dependent.

4.2.5 Discussion of the Models and Equilibrium Concepts

Closed-loop strategies allow players to condition their actions on actions taken in previous stages; in open-loop strategies, this is not possible. Thus, equilibria in the closed-loop model are by definition subgame perfect, whereas open-loop equilibria are typically merely dynamically (time) consistent. The latter is a weaker equilibrium concept than subgame perfection. It requires only that no player has an incentive at any time to deviate from the strategy he announced at the beginning of the game, "given that no player has deviated in the past and no agent expects a future deviation" (Karp and Newbery, 1992). Therefore, with subgame perfect equilibria requiring actions to be optimal in every subgame of the game, i. e., requiring that no player has an incentive to deviate from his strategy regardless of any deviation in the past, an equilibrium of the open-loop model may fail to be an equilibrium in the closed-loop game.[56]

Fudenberg and Tirole (1991) and the literature cited therein generally address the issue of diverging results of open-loop models in comparison to closed-loop models and provide intuition for the divergence: In the closed-loop model, in contrast to the open-loop model, a player's influence via its own actions in the first stage on the other players' actions in the second stage is taken into account. Applying this intuition to the special case of the capacity expansion problem, Murphy and Smeers (2005) show that in the closed-loop equilibrium, marginal investment costs may be higher than the sum of the short-term marginal value implied by the KKT conditions. In particular, they note that "the difference between the two characterizes the value for the player of being able to manipulate the short-term market by its first stage investments." This may lead to higher investments and supplies and hence lower prices in the closed-loop model compared to the open-loop model.

[55] Note that the first stage decision variable is separated from the second stage decision variables by a semicolon. The latter are indirectly determined by the choice of the first stage decision variable.

[56] See Selten (1965) for the first formalization of the concept of subgame perfect equilibria and, e. g., Karp and Newbery (1989) for a general account on dynamic consistency.

The existing literature on the subject, in particular the above mentioned Murphy and Smeers (2005) as well as Wogrin et al. (2013b), provides general properties of closed-loop and open-loop models and conditions for diverging and non-diverging results between the two models, assuming simplified settings (e. g., ignoring existing capacities). We conjecture that in a spatial application with non-generic data and existing capacities available to the players, equilibria are likely to deviate between the two modeling approaches, which is confirmed by our application to the metallurgical coal market (see Sections 4.4 and 4.5). Analytical analysis is no longer available in this setting due to increased complexity and thus makes a numerical analysis necessary. The numerical approach is also suitable to address an issue which to our knowledge has not yet been comprehensively touched upon in previous literature: a quantification of the magnitude of the divergence between closed-loop and open-loop model results.

4.3 Implementation

4.3.1 Model 1: The Open-Loop Model

Both open-loop models introduced in Section 2, i. e., the open-loop perfect competition model and the open-loop Cournot competition model with competitive fringe, are implemented as mixed complementarity problems (MCP).

4.3.2 Model 2: The Closed-Loop Model

We solve the closed-loop model using diagonalization (see for instance Gabriel et al., 2012):

1. Set starting values for the investment decisions y_l^0 of all oligopolistic producers $l \in L$, a convergence criterion ϵ, a maximum number of iterations N and a learning rate R
2. $n = 1$
3. Set $y_l^n = y_l^{n-1}$
4. Do for all $l \in L$
   (a) Fix the investment decisions y_{-l}^n of $-l$
   (b) Solve player l's MPEC problem MPEC$_l$ to obtain an optimal investment level y_l

(c) Set y_l^n equal to $R \cdot y_l + (1 - R) \cdot y_l^n$

5. If $|y_l^n - y_l^{n-1}| < \epsilon$ for all producers $l \in L$: quit
6. If $n = N$: quit
7. $n = n + 1$ and go back to step 3

Diagonalization thus reduces the closed-loop problem to a series of MPEC problems. Concerning the solution of the MPECs, we implement two procedures: grid search along the investment decision y_l and a reformulation of the MPEC as a Mixed Integer Linear Program (MILP).

Both approaches differ with respect to the simplification of the decision variables: With grid search we discretize the investment decision which is reasonable for many investment choices in real life. Thus, solving the MPEC problem reduces to solving a series of MCP problems with the choice of production volumes remaining continuous. On the contrary, in the MILP approach we discretize the production decisions but retain a continuous choice of investments in new capacity. The discretization may result in missing the global optimal solution.[57] As both approaches result in very similar outcomes (see Section 4.6) we are confident that our obtained results are valid.

Implementing both the grid search and MILP reformulation allows for the comparison of the computer run-times of the two models, with grid search typically being faster for reasonable grid sizes (see Section 4.6 for details on this issue).

4.3.2.1 Grid Search

When applying grid search along the investment decision y_l, MPEC$_l$ simplifies to a sequence of complementarity problems. In our implementation, the grid width in the grid search is the same for all producers; the number of steps for a producer is thus dependent on his capacity expansion limit.

4.3.2.2 MILP Reformulation

In addition to grid search, we implement a MILP reformulation of the MPEC. Non-linearities arise in the MPEC due to the complementarity constraints and the non-linear term in the objective function.

[57] A third way of approaching the non-linearities in the model might be using the strong duality theorem to linearize the original MPEC as described in Ruiz and Conejo (2009).

The former are replaced by their corresponding disjunctive constraints (see Fortuny-Amat and McCarl, 1981), e. g., we replace

$$0 \leq cap_m^0 + y_m - \sum_j x_{m,j} \perp \lambda_m \geq 0$$

by

$$\begin{aligned}
&M^\lambda b_m^\lambda \geq \lambda_m \\
&M^\lambda (1 - b_m^\lambda) \geq cap_m^0 + y_m - \sum_j x_{m,j}
\end{aligned}$$

for some suitably large constant M^λ and binary variables b_m^λ.

For the discretization of the non-linear term in the objective function, we proceed following Pereira et al. (2005) using a binary expansion of the supply variable. The binary expansion of $x_{m,j}$ is given by

$$x_{m,j} = \underline{x} + \Delta_x \sum_k 2^k b_{k,m,j}^x,$$

where $\underline{x}$ is the lower bound, Δ_x the stepsize, k the number of discretization intervals and $b_{k,m,j}^x$ binary variables. Substituting $P_j \cdot \underline{x} + \Delta_x \sum_k 2^k z_{k,m,j}^x$ for $P_j \cdot x_{m,j}$, we have to impose the additional constraints

$$\begin{aligned}
&0 \leq z_{k,m,j}^x \leq M^x b_{k,m,j}^x \\
&0 \leq P_j - z_{k,m,j}^x \leq M^x (1 - b_{k,m,j}^x)
\end{aligned}$$

for some suitably large constant M^x.

4.4 Data Set

The models are parametrized with data for the international metallurgical coal market (see Table 4.1 and Appendix A). Yet, as the structure of the international metallurgical coal trade is (from a modeling perspective) similar to that of other commodities, the model could easily be calibrated with data for other markets.

Metallurgical coal is used in steel-making to produce the coke needed for steel production in blast furnaces and as a source of energy in the process of steel-making. Metallurgical coal is distinct from thermal coal, which is typically used to generate electricity or heat.

Currently around 70 % of the global steel production crucially relies on metallurgical coal as an input.[58]

International trade of metallurgical coal amounted to 250 million tonnes (Mt) in 2012.[59] International trade is predominantly seaborne, using dry bulk vessels. Up until 2010, metallurgical coal was almost exclusively traded through long-term contracts. Since then, the market has begun to move away from this system towards more spot market-based trading. While the share of spot market activity has increased rapidly, a substantial amount of metallurgical coal is still traded through long-term contracts.

Key players in this market are large mining companies such as BHP-Billiton, Anglo-American, Glencore and Rio Tinto. These companies produce mainly in Australia and, together with Peabody Energy's Australian operations, control more than 50 % of the global export capacity. In addition, adding to this the market share of the Canadian Teck consortium and the two key metallurgical coal exporters from the United States, Walter Energy and Xcoal, results in almost three quarters of the global export capacity, marketed by an oligopoly of eight companies. For the sake of simplicity and computational tractability, we aggregate these players' existing mines into one mining operation per player. Smaller exporters from Australia, the United States, Russia, New Zealand, Indonesia and South Africa are aggregated into three players: one Cournot player from Australia (AUS6), one Cournot player from the United States (USA1) and one competitive fringe player that comprises all other regions (Fringe). This results in eleven asymmetric players who differ with respect to their existing production capacity and the associated production and transport costs (see Table 4.2).[60]

Table 4.2: Existing Capacity, Variable and Investment Costs

Players	Existing Capacity [Mtpa]	Variable Costs [US$/t]	Investment Costs [US$/tpa]	Max. Investment [Mtpa]
USA1	38	122.0	-	-
USA2	9	122.1	98.2	50
USA3	11	141.0	98.0	50
AUS1	54	118.3	218.1	50
AUS2	11	118.4	218.0	50
AUS3	17	118.5	217.9	50
AUS4	10	118.6	217.8	50
AUS5	12	118.0	218.2	50
AUS6	18	118.1	-	-
CAN	26	105.0	161.0	20
Fringe	26	78.0	-	-

[58]See WCA (2011).
[59]See IEA (2013).
[60]Data on capacities and costs are taken from Trüby (2013).

We assume that the three players representing the smaller exporters, i. e., USA1, AUS6 and Fringe, cannot invest in additional capacity. Hence, only the largest eight companies can endogenously expand their supply capacity. The investment decision, made in stage one, is based on the players' capacities and costs in 2011. We consider one investment cycle with capacities becoming available after six years serving one demand period. Investment costs per tonne of annual production capacity (tons per annum) are broken down into equal annual payments based on an annuity calculation using an interest rate of 10 % and a depreciation time of 10 years. The profitability of investments is evaluated based on the comparison of annuity and profits in the considered production stage. We therefore assume that returns are constant over the years of production. Note that production cost of new mines correspond to the production cost of the respective player's existing mine.

The two largest importers of metallurgical coal are Europe and Japan, followed by India, China and Korea. These key importers account for more than 80 % of the trade. We aggregate these and the remaining smaller countries into two demand regions: Europe-Atlantic and Asia-Pacific.[61] The former also includes the Mediterranean's neighboring countries and importers from the Atlantic shores of the Americas. The latter includes importers with coastlines on the Pacific or the Indian Ocean. Exporters from the United States have a transport cost advantage in the Europe-Atlantic region, while Canadian and Australian exporters are located closer to the consumers in the Asia-Pacific region (see Table 4.5 in the Appendix). We assume the inverse import demand function for metallurgical coal to be linear. The function can be specified using a reference price and a corresponding reference quantity in combination with a point-elasticity *eta*.

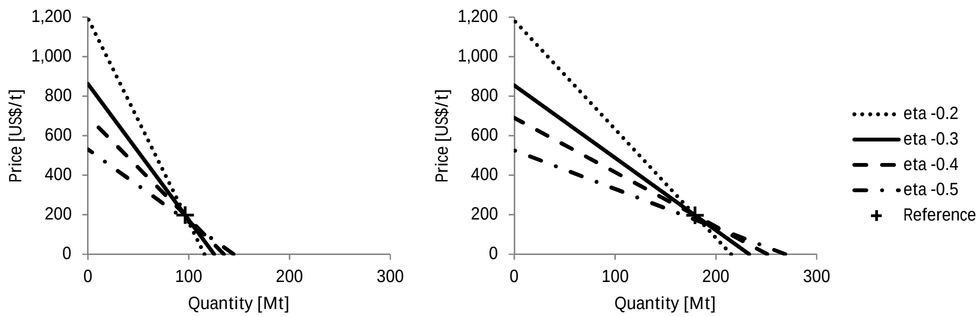


FIGURE 4.1: Demand functions for Europe-Atlantic (left) and Asia-Pacific regions (right) with varying elasticity

[61]Our approach covers 100 % of the global seaborne metallurgical coal imports and exports (based on data from 2011).

4.5 Results

In practice, investors in production capacity face demand evolution as a key uncertainty. Accounting for this uncertainty, we run sensitivities in which we vary the point-elasticity parameter *eta* across the range -0.2 to -0.5 (see Figure 4.1).[62] This bandwidth is generally considered reasonable in the metallurgical coal market (see Trüby, 2013, and the literature cited therein). Furthermore, we vary the reference demand quantity (see Table 4.6 in Appendix A) from 60 % to 140 % to account for different demand evolution trajectories. The presentation of the results is structured around the variation of these demand parameters followed by a general discussion of the findings.

4.5.1 Variation of Demand Elasticity

Decreasing the point elasticity parameter *eta* results in a flatter gradient of the linear demand function (see Figure 4.1). A decreasing *eta* (i. e., a more negative *eta*) expresses an increasing price responsiveness of consumers which, *ceteris paribus*, limits the extent to which the oligopolists can exploit their market power. Consequently, with decreasing *eta*, average prices achieved in the imperfect competition cases (open-loop and closed-loop) are decreasing while total production is increasing (Figure 4.2). Note that in the perfect competition case, the aggregate supply and the aggregate demand curves intersect below the reference point resulting in an increase in production with decreasing *eta* and, correspondingly, with increasing marginal costs, an increase in production results in an increase in price. In the two Cournot models with imperfect competition the oligopolistic mark-up on marginal costs leads to market prices exceeding the reference price.

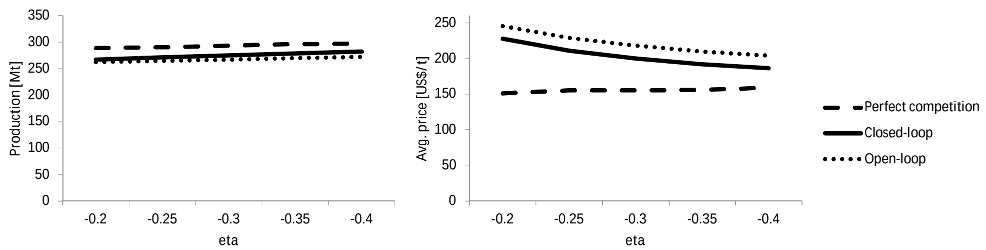


FIGURE 4.2: Total production (left) and average market price (right) for varying demand elasticity

[62] For *eta* smaller than -0.4, closed-loop model runs did not converge. Therefore, the results presented in this section only comprise the range -0.2 to -0.4. For a discussion on computational issues, see Section 4.6.

A variation of *eta* impacts the investment trends differently in the three cases (Figure 4.3). However, the capacity expansion investments need to be interpreted in concert with the corresponding utilization of the existing capacity. Intuitively, one would expect investment into additional capacity to be highest in the perfect competition case. Yet, in our setup, the investment level in the perfectly competitive case falls between the two cases with imperfect competition. This effect stems from the significant amount of existing capacities which – with the exception of some very high-cost capacities – are utilized before additional production capacity is built. Murphy and Smeers (2005) show that in their model which does not account for existing capacities, investment levels are indeed highest under perfect competition.

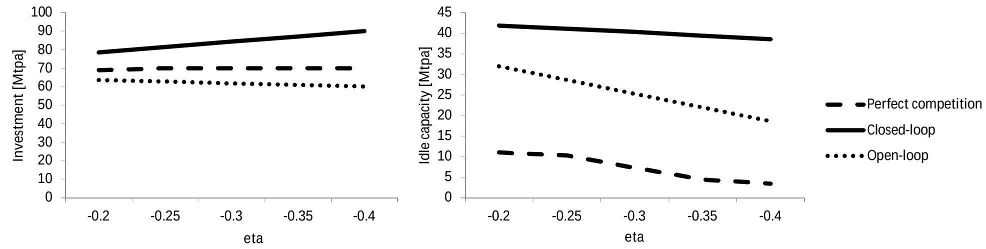


FIGURE 4.3: Capacity investments (left) and idle capacity (right) for varying demand elasticity

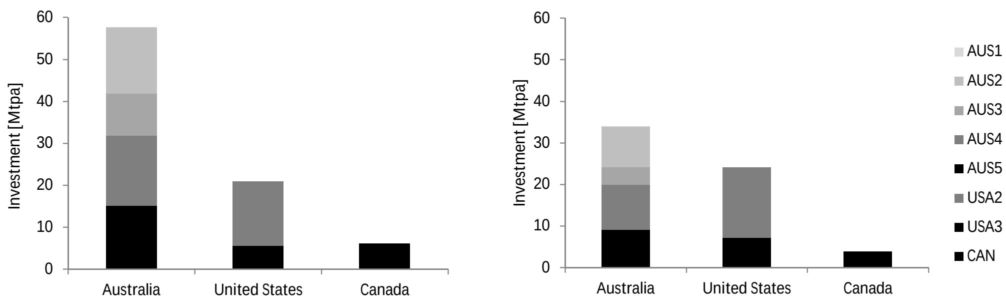


FIGURE 4.4: Capacity investments for the closed-loop (left) and open-loop model (right) ($eta = -0.3$)

Of particular interest is the ranking of the closed-loop and open-loop case in terms of capacity expansion and capacity withholding. Note that withholding (or idle capacity), here and in the following, concerns only exiting capacities. Each player exhausts existing capacities before investing in additional capacities. Newly built capacities are always fully utilized in equilibrium as otherwise players could increase their profit by reducing investments. Investments in the open-loop case are strictly lower than in the closed-loop case independent of the elasticity while less capacity is withheld in the open-loop case. However, the investment behavior of individual players may differ from the aggregate industry behavior; as can be seen in Figure 4.4 two players from the United States invest more in the open-loop model than in the closed-loop model contrary to what the rest of the industry does.

The investment level is higher in the closed-loop case compared to the open-loop case as the capacity expansion in the first stage can be used strategically to influence the supply decisions of the other players in the second stage. In the closed-loop case, the two-stage structure introduces, informally speaking, an additional element of competition that does not exist in the open-loop case. The open-loop result is a stable Nash-equilibrium of the one-stage game in which players decide simultaneously on capacity investments and production. In the closed-loop case, the sequential form of the game leads to an interaction between the capacity choice of the first stage and the production decision of the second stage which may alter players' decisions compared to the open-loop case. However, asymmetries play a critical role for the divergence of open-loop and closed-loop equilibria: Wogrin et al. (2013b) demonstrate that for symmetric players (under certain mild conditions) open-loop and closed-loop equilibria are identical.

To get an intuition for the difference between the open-loop and the closed-loop game, suppose for a moment that there are two players, an incumbent with infinite existing capacity and an entrant without any capacity. Both players face the same production costs while the entrant faces additional (non-zero) investment costs. It is important to note that in both games the production decision of each player depends on his own costs and the costs of the other player: higher own costs result in lower own production and higher output of rivals – higher costs of rivals result in higher own output as the production of the rivals is lower.

In the *open-loop* case, the incumbent decides on his output, knowing about the total costs of the entrant (investment and production costs), which are featuring in his first-order condition. In equilibrium the entrant produces and invests so that marginal revenue equals the sum of marginal production and investment costs. Since the entrant's total costs are higher, he produces less than the incumbent in equilibrium. The incumbent's profits are higher than the entrant's as he produces more and does not have to pay for capacity.

Contrarily, in the *closed-loop* case, the choice of capacity and production is sequential. Solving the problem by backward-induction, assume that the entrant has built more capacity than in the open-loop case. With the production decision now being based on the incumbent's and the entrant's production costs, the entrant will now produce more and the incumbent less than in the open-loop case as in the latter both, investment and production costs, appear in the entrant's (and with reversed sign the incumbent's) first-order condition. But when would the entrant build more capacity in the closed-loop case than in the open-loop case? Only if the potential gains in the second stage (the spot market) exceed the increased investment cost from building more capacity in the first stage. The entrant anticipates in the first stage that by increasing his capacity he can reduce the incumbent's production in the second stage. Knowing this, the entrant decides on his optimal capacity investment while anticipating his influence on the incumbent's level of production. The interaction between the decisions taken in the different stages is not present in the open-loop model. In the asymmetric entrant/incumbent case considered in this intuition, it is indeed profitable for the entrant to deviate from his open-loop outcome and to invest more. As a result, the incumbent's profits are still higher than the entrant's (as he does not have to pay for capacity) but are lower than in the open-loop case while the entrant can raise his profits compared to the open-loop case.[63]

An analytical solution of this game becomes non-trivial when more than one player makes subsequent investment and supply decisions as these decisions mutually influence each other. Yet, this little example is useful to provide a better understanding of why the closed-loop case features higher investment levels but also higher withholding of existing capacities than the open-loop case.

In our application to the metallurgical coal market, capacity is exclusively withheld by the two largest players (one producing in Australia and the other in the United States), in both models of imperfect competition. Capacity expansion and withholding are following opposing trends in our models of imperfect competition, i. e., the open-loop model exhibits a lower level of investment but also a lower level of unused capacity while the higher investment levels in the closed-loop model come with a higher level of idle capacity. Thus, it is a-priori unclear how the two models would rank in terms of total supply and market prices. A numerical solution of our models yields that supply is higher in the closed-loop case than in the open-loop case. Consequently, market prices are lower in the closed-loop case. This result is in line with the findings of Murphy and Smeers (2005).

Industry profits, consumer rent and social welfare are depicted in Figure 4.5. Industry profits decrease with decreasing *eta* and so does consumer rent (a higher price-responsiveness of consumers limits market power exploitation but also potential consumer rent). The existence of profits in the perfect competition model is due to capacity restrictions of existing mines and limited expansion potential for new mines. Social welfare is similar in all three models: in a perfectly competitive market welfare is slightly higher than in the Cournot models. Welfare is lowest in the open-loop case (Figure 4.6). Thus, the different underlying assumptions concerning the prevailing market structure in the international metallurgical coal trade (long-term contracts versus spot market)

[63]See calculations in the Appendix for a more formal statement of this analysis.

primarily influences the surplus distribution rather than its sum: in the open-loop case in which the product is traded through long-term contracts, companies can earn higher profits, while consumer surplus is higher in markets with spot market-based trade.

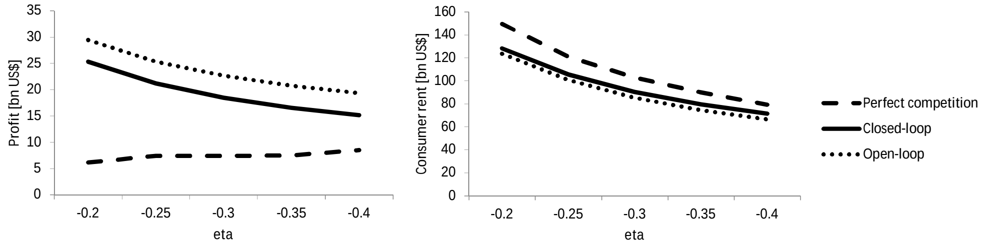


FIGURE 4.5: Accumulated profits (left) and consumer rent (right) with varying demand elasticity

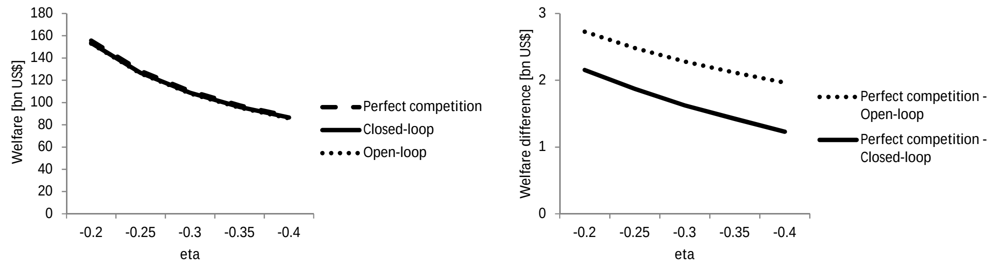


FIGURE 4.6: Overall welfare (left) and welfare differences (right)

4.5.2 Variation of Reference Demand

For the variation of reference demand, the point elasticity *eta* has been fixed to a value of -0.3; thus the case of 100 % reference demand corresponds to the depicted results of the previous subsection with the same demand elasticity. Variations of the reference demand results in a shift of the demand curve to the right for values larger than 100 % and a shift to the left for values lower than 100 %.

As in the previous subsection, supply is highest under perfect competition and lowest in the open-loop case for any demand variation (Figure 4.7). Accordingly, prices are highest in the open-loop case followed by the closed-loop and the perfect competition cases. As one would expect, supply and average prices increase with increasing demand.

For low reference demand levels, the existing capacities of small players are almost sufficiently high to produce the quantities needed for their best-supply response in stage two. Therefore, the results in the open-loop and closed-loop cases almost coincide at 60 % reference demand as investment activity is low. Investments in additional production capacity are increasing monotonously with growing reference demand (Figure 4.8). As

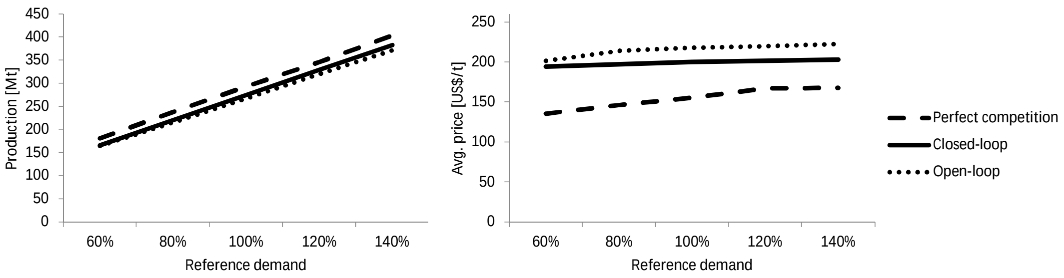


FIGURE 4.7: Total production (left) and average market price (right) for varying reference demand

with the variation of the demand elasticity, investments are consistently lower in the open-loop case than in the closed-loop case. For low demand levels, investments in the competitive model are below those in the models with imperfect competition as existing capacities are sufficient to serve demand rendering investments unprofitable. In the Cournot models, investment into additional production capacity is still profitable for small players as they can count on players with large existing capacities to withhold some output.

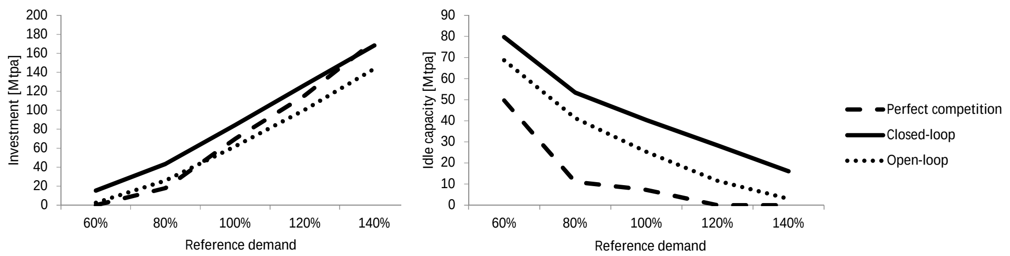


FIGURE 4.8: Capacity investments (left) and idle capacity (right) for varying reference demand

For high demand levels, investments under perfectly competitive conduct exceed those even in the closed-loop model. The order of idle capacity is similar to the case of varying demand elasticity: idle capacity is highest in the closed-loop model followed by the open-loop case (both due to strategic considerations) and the perfect competition model (due to market prices below the marginal costs of costlier capacities).

With increasing demand, profits as well as consumer rents increase (Figure 4.9). Again, results for the open-loop and closed-loop cases almost coincide if reference demand is very low as investments play a minor role. In the case of high reference demand, profits in the open-loop model exceed those in the closed-loop model. Results for consumer rents are vice versa. Total welfare turns out to be quite similar for all three models with the highest welfare occurring in the perfectly competitive model followed by the closed-loop and open-loop models (Figure 4.10).

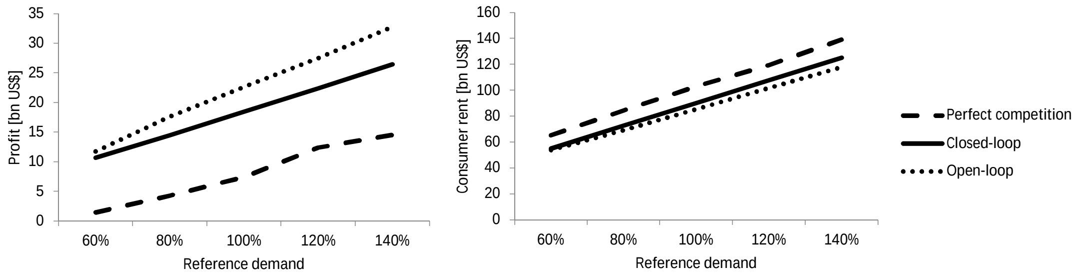


FIGURE 4.9: Accumulated profits (left) and consumer rent (right) with varying reference demand

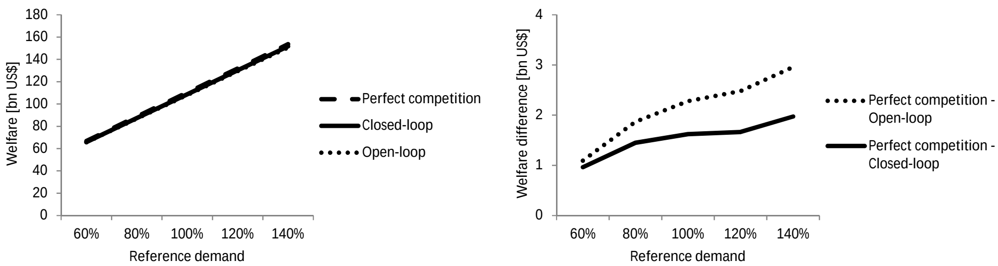


FIGURE 4.10: Overall welfare (left) and welfare difference (right, open-loop minus closed-loop)

4.5.3 Summary

Asymmetric existing capacities are an important driver of our results. While welfare is highest in the perfect competition case, investment levels in this case fall between the two Cournot models as existing capacities are sufficient to absorb additional demand. Profits are highest in the open-loop case followed by the closed-loop and perfect competition models. Moving away from long-term contracts towards a spot market-based trade reduces profits of all players, however, companies with large existing capacities are affected to a larger degree: the two large firms (one from Australia and one from the United States) who are responsible for the withholding of capacity in the Cournot models together receive 23 % of the industry profits in the open-loop case but see their share of profits diminished to 17 % in the closed-loop case.

In our modeling setup the competitive fringe has no strategic relevance. Fringe players neither invest nor withhold, i. e., they always produce to capacity. In essence, the fringe determines the residual demand that the oligopolists optimize against but it does not introduce any sort of first-mover vs. follower relationship.

The magnitude of result deviations between the different models, and thus the implications for market participants are quite significant. The models of imperfect competition differ, for instance, in capacity expansions between 19 % and up to 33 % (low and high demand elasticity, respectively).

Even though social welfare differs only slightly between the open-loop and closed-loop models in our calculations for the metallurgical coal market, the difference may be higher for other markets with different model parameters. In addition, the surplus distribution between consumer rent and profits differs significantly and has policy implications since – in natural resource markets – production and consumption take place in different countries.

4.6 Computational Issues

Equilibria in a closed-loop model, if any exist, do not necessarily have to be unique. Therefore, we perform a robustness check for our closed-loop results by using different starting values for capacity investments. Starting values are randomly drawn from a reasonable range of possible investments, with the maximum investment of each player as given in Table 4.2. Limiting the range of possible investments drastically reduces computer run-times and increases the probability of finding equilibria. In addition, calculations are made with starting values set to zero and to the open-loop results. The algorithm terminates if overall adjustments of investments δ are less than $\epsilon = 0.1$ million tons per annum compared to the previous iteration. We use a learning rate parameter R for the adoption rate of new investments in order to avoid cycling behavior. The learning rate parameter is randomly set between 0.6 and 1.0 (see Gabriel et al., 2012). Calculations have been done on a 16 core server with 96 GB RAM and 2.67 GHz using CPLEX 12.2.

Table 4.3 shows calculation statistics when using the MILP version of our model (see Subsection 4.3.2.2). We perform six runs per parameter setting using random starting values. Most runs converged to an equilibrium before the maximum number of iterations was reached. With increasing demand elasticity, the algorithm had difficulties to converge. In the case of $eta = -0.4$, only every third run converged to an equilibrium; for $eta < -0.4$, no equilibrium could be found at all. Using either zero investments or open-loop results as starting values, a closed-loop equilibrium was found, except for $eta < -0.4$.

Table 4.3: Computation time and convergence to equilibrium - MILP version (random, zero, open-loop starting values)

Scenario	Convergence (max. 10 iterations)	Iterations until convergence (only converged runs, max. 10)	Calculation time (only converged runs) [h]
reference case (eta -0.3, dem 1.0)	6/6, yes, yes	6-7 (avg. 6.8), 7, 6	10.7-13.7 (avg. 12.4), 7.1, 5.2
eta -0.2	6/6, yes, yes	7-8 (avg. 7.3), 6, 6	9.2-14.1 (avg. 11.0), 5.7, 4.1
eta -0.25	6/6, yes, yes	7-10 (avg. 8.2), 7, 6	11.4-14.9 (avg. 12.8), 6.9, 5.5
eta -0.35	6/6, yes, yes	6-8 (avg. 7.2), 6, 6	11.3-15.8 (avg. 12.7), 5.1, 5.6
eta -0.4	2/6, yes, yes	7-8 (avg. 7.5), 9, 7	12.2-12.7 (avg. 12.4), 5.6, 7.8
eta -0.45	0/6, no, no	-, -, -	-, -, -.
eta -0.5	0/6, no, no	-, -, -	-, -, -
dem 0.6	5/6, yes, yes	7-9 (avg. 7.4), 7, 7	1.9-3.5 (avg. 2.2), 0.1, 0.2
dem 0.8	6/6, yes, yes	7-8 (avg. 7.5), 6, 5	3.6-8.8 (avg. 7.1), 2.0, 2.3
dem 1.2	6/6, yes, yes	6-9 (avg. 7.8), 7, 6	9.8-13.9 (avg. 11.9), 8.3, 6.0
dem 1.4	6/6, yes, yes	6-10 (avg. 8.3), 7, 6	7.7-11.2 (avg. 8.7), 9.7, 5.7

Figure 4.11 illustrates the iterative solution process for a single model run for $eta = -0.5$ using random starting values. The model run did not converge to an equilibrium.[64] After initial adjustments of investments in the first iterations, investments start to cycle in a rather small range. Total investments from iteration 5 to 10 vary between 89 million tons per annum and 97 million tons per annum. This range is typical for all runs regardless of the starting values. The maximum range for a single player's investment deviations is 3 Mtpa. Thus, even if no equilibrium is reached, analyzing the solution process may hint to possible market developments.

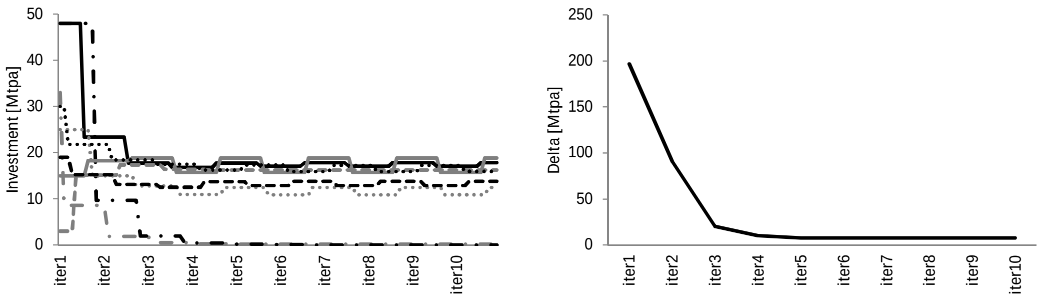


Figure 4.11: Course of investments of single players during solution process ($eta = -0.5$)

Using zero investments or open-loop equilibrium results as starting values led to a significant reduction of computer run-times compared to random starting values. This is probably due to the rather large range of random starting values and the (comparably) rather small equilibrium investments. Thus, starting from zero investments in most cases is closer to the equilibrium values than starting with random values. In summary, using reasonable starting values can support the solution process significantly.

[64]In our iterative approach, convergence depends on the choice of (an arbitrarily small) ϵ.

If the algorithm converged, model results were identical for all runs with the same parameters concerning demand level and demand elasticity. Thus, even if the existence of multiple equilibria cannot be excluded, equilibria appear to be stable.

Calculations using the MILP version of our model usually took several hours to converge to an equilibrium. Applying the grid search approach (see Section 4.3.2.1) reduced computer run-times significantly. The conceptional difference between both approaches lies in the simplification of the decision variables: With grid search we discretize the investment decision. On the contrary, in the MILP approach we discretize the production decisions but retain a continuous choice of investments in new capacity.

The same calculations as in the MILP version have been done using grid search with investment steps of 0.1 million tons per annum and the same convergence criterion as in the MILP version ($\epsilon = 0.1$ million tons per annum). The model was implemented in GAMS using GUSS (see Bussieck et al., 2012).

Table 4.4: Computation time and convergence to equilibrium - Grid Search (random, zero, open-loop starting values)

Scenario	Convergence (max. 10 iterations)	Iterations until convergence (only converged runs, max. 10)	Calculation time (only converged runs) [min]	Accumulated absolute difference between investments in MILP and grid version [%]
reference case (eta -0.3, dem 1.0)	6/6, yes, yes	6-7 (avg. 6.3), 7, 6	2.8-15.7 (avg. 9.3), 2.2, 2.4	0.7-0.9, 0.8, 0.8
eta -0.2	6/6, yes, yes	5-7 (avg. 6.3), 7, 5	3.5-16.7 (avg. 10.2), 2.3, 2.0	1.0, 1.0, 1.0
eta -0.25	6/6, yes, yes	6-7 (avg. 6.7), 7, 6	2.4-16.5 (avg. 9.4), 2.2, 2.4	0.8, 0.7, 0.8
eta -0.35	6/6, yes, yes	6-8 (avg. 7.0), 7, 6	2.8-17.5 (avg. 10.4), 2.2, 2.4	0.8-1.2, 1.2, 0.8
eta -0.4	6/6, yes, yes	6-7 (avg. 6.5), 7, 6	2.5-16.2 (avg. 9.3), 2.2, 2.4	1.2-1.5, 1.5, 1.5
eta -0.45	0/6, no, no	-, -, -	-, -, -.	-, -, -
eta -0.5	0/6, no, no	-, -, -	-, -, -	-, -, -
dem 0.6	6/6, yes, yes	6-8 (avg. 7.0), 5, 5	3.2-16.9 (avg. 9.7), 2.0, 2.0	2.5-3.7, 3.1, 3.3
dem 0.8	6/6, yes, yes	6-7 (avg. 6.7), 6, 6	2.8-16.2 (avg. 9.6), 2.6, 2.4	0.9, 1.0, 0.9
dem 1.2	6/6, yes, yes	5-7 (avg. 6.8), 7, 6	2.9-17.2 (avg. 10.1), 2.3, 2.5	0.3, 0.3, 0.3
dem 1.4	6/6, yes, yes	5-6 (avg. 6.3), 7, 6	2.9-16.1 (avg. 9.7), 2.3, 2.5	0.3-0.4, 0.4, 0.4

Applying grid search, the solution process took only several minutes to converge. Thus, reducing the optimization process from a series of computationally challenging MPECs to comparably easy-to-solve complementarity problems reduced overall computer run-times significantly. As for the MILP version, all model runs converged to the same equilibrium (for $eta \geq -0.4$) or did not converge at all (for $eta < -0.4$). Aggregated absolute deviations of investments between the MILP and the grid search version of our model vary between 0.3 % and 3.7 %. Thus, in our parameter setting, only minor differences in the results occurred.

4.7 Conclusions

We presented three investment models for oligopolistic spatial markets. Our approach accounts for different degrees of competition and as to whether the product is sold through long-term contracts or on spot markets. The models are particularly suited for the analysis of investments in markets for natural resources and minerals. We applied the models to the international metallurgical coal trade, which features characteristics similar to those of other commodity markets.

Results may differ substantially between the different models. The closed-loop model, which is computationally challenging, is particularly well suited for when the product is traded on a spot market and the investment expenditure is large compared to production costs. The open-loop model is appropriate for markets with perfect competition or imperfectly competitive markets on which the product is traded through long-term contracts. Moreover, the open-loop model approximates the closed-loop outcome when investment costs are minor.

Over the last several years, progress has been made in the metallurgical coal and iron ore markets to move away from long-term contracts and introduce spot markets in commodity trade. Similarly, efforts are being made to introduce spot market-based pricing between European natural gas importers and the Russian gas exporting giant Gazprom. Our results suggest that moving away from long-term contracts in oligopolistic markets is likely to stimulate additional investment and consequently reduce profits and increase consumer rents. The overall effect on welfare is negligible. However, in natural resource markets, export revenues and consumer rents from imports are typically accrued in different legislations. Hence, policy makers from exporting and importing countries are likely to have differing views on how commodity trade should be organized.

Further research is needed to improve methods for solving complex two-stage problems. In addition, further research could apply the models presented here to other oligopolistic mining industries such as the copper or iron ore trade. Given that static pricing models tend to give unsatisfactory results for the oil market, in which variable costs are low but capital expenditure is very high, the closed-loop approach may provide interesting insights into oligopolistic pricing when accounting for investments in capacity.

4.8 Appendix

4.8.1 Appendix A: Input data

Table 4.5: Distance

from	to	distance [Nautical miles]
United States	Europe-Atlantic	3,387
	Asia-Pacific	10,978
Australia	Europe-Atlantic	11,626
	Asia-Pacific	3,731
Canada	Europe-Atlantic	8,840
	Asia-Pacific	4,227
Fringe	Europe-Atlantic	5,018
	Asia-Pacific	3,037

Table 4.6: Reference Demand and Reference Price

Market	Reference Demand [Mt]	Reference Price [US$/t]
Europe-Atlantic	96	180
Asia-Pacific	179	180

4.8.2 Appendix B: Open- vs. closed-loop investments

We provide an intuition for the results presented in this paper by solving analytically a simplified model consisting of one market and two players. Player I is the incumbent in the market and has infinite existing capacity. Player E is the entrant to the market owning no existing capacity. The entrant can invest at cost k per unit, whereas the incumbent may not invest. Both players produce at variable production costs c and there are no transportation costs to the market. There is only one time period and the inverse residual demand curve for this period is given by $P = a - (x_I + x_E)$.

We solve both the open-loop and the closed-loop model for this simplified setting and show that there is an incentive for the players in the closed-loop model to deviate from their open-loop equilibrium quantities.[65]

The open-loop model

In the open-loop model, the entrant's optimization problem is given by

$$\max_{x_E, y_E} P \cdot x_E - c \cdot x_E - k \cdot y_E$$

subject to

$$\begin{aligned} &y_E - x_E \geq 0 \ (\lambda_E), \\ &x_E \geq 0, \\ &y_E \geq 0. \end{aligned}$$

From the corresponding KKT conditions, it is easy to see that in an open-loop equilibrium the capacity of the entrant is fully utilized, i. e., $y_E = x_E$. Therefore the optimization problem may be simplified to

$$\max_{x_E \geq 0} P \cdot x_E - (c + k) \cdot x_E.$$

Taking the derivative with respect to x_E, we obtain the first order condition

$$a - 2x_E - x_I - (c + k) = 0$$

from which we obtain the entrant's reaction curve

$$x_E = \frac{a - (c + k) - x_I}{2}.$$

[65] We restrict our attention to parameter settings in which both players produce. This restriction is adequate for the objective at hand, namely to provide intuition for the main results in the paper.

The incumbent faces a different optimization problem, as he may not invest but has infinite existing capacity. The incumbent's optimization problem is hence given by

$$\max_{x_I \geq 0} P \cdot x_I - c \cdot x_I$$

which yields, when taking the derivative with respect to x_I, the first order condition

$$a - 2x_I - x_E - c = 0.$$

The first order condition can be solved for x_I to obtain the incumbent's reaction curve

$$x_I = \frac{a - c - x_E}{2}.$$

Solving the system of equations consisting of the two reaction curves for the players' supply quantities, we obtain

$$x_I = \frac{a - c + k}{3}$$

and

$$x_E = \frac{a - c - 2k}{3},$$

which is the solution to the open-loop model if the non-negativity conditions for x_I and x_E are fulfilled.

The closed-loop model

In order to solve the closed-loop model we use backward induction. For this let y_E denote the first stage investment volume of the entrant. The entrant's second stage optimization problem is then given by

$$\max_{x_E \geq 0} P \cdot x_E - c \cdot x_E$$

subject to

$$y_E - x_E \geq 0 \ (\lambda_E).$$

The Lagrangian to this optimization problem is given by

$$\mathcal{L} = P \cdot x_E - c \cdot x_E + \lambda_E \cdot (y_E - x_E)$$

from which the KKT conditions follow:

$$x_E = \frac{a - (c + \lambda_E) - x_I}{2},$$
$$0 \leq \lambda_E \perp y_E - x_E \geq 0.$$

The incumbent faces the optimization problem

$$\max_{x_I \geq 0} P \cdot x_I - c \cdot x_I$$

which yields, as in the open-loop model, the reaction curve

$$x_I = \frac{a - c - x_E}{2}.$$

By inserting this in the above KKT condition we obtain the expression

$$x_E = \frac{a - c - 2\lambda_E}{3}.$$

The first stage optimization problem of the entrant is then given by

$$\max_{y_E \geq 0} P \cdot x_E - c \cdot x_E - k \cdot y_E$$

subject to

$$\begin{aligned}
& x_E = \frac{a - c - 2\lambda_E}{3} \\
& 0 \leq \lambda_E \perp y_E - x_E \geq 0, \\
& x_I = \frac{a - c - x_E}{2}.
\end{aligned}$$

Consider the case in which the capacity of the entrant is fully utilized, i. e., $x_E = y_E$. In this case the optimization problem may be simplified to

$$\max_{y_E \geq 0} P \cdot y_E - (c + k) \cdot y_E$$

subject to

$$\begin{aligned}
& \frac{a - c}{3} - y_E \geq 0 \ (\mu_E), \\
& x_I = \frac{a - c - y_E}{2}.
\end{aligned}$$

The Lagrangian to this optimization problem is given by

$$\mathcal{L} = \left[a - \left(y_E + \frac{a - c - y_E}{2} \right) \right] \cdot y_E - (c + k) \cdot y_E + \mu_E \cdot \left(\frac{a - c}{3} - y_E \right)$$

from which the KKT conditions follow:

$$\begin{aligned}
& y_E = \frac{a - c - 2k - 2\mu_E}{2}, \\
& 0 \leq \mu_E \perp \frac{a - c}{3} - y_E \geq 0.
\end{aligned}$$

In case of $\mu_E > 0$, we obtain

$$y_E = x_E = x_I = \frac{a-c}{3}$$

and $\lambda_E = 0$.[66]

This is indeed a solution in case

$$\mu_E = \frac{a-c-2k-2y_E}{2} = \frac{a-c}{6} - k > 0.$$

In case of $\mu_E = 0$, we obtain

$$y_E = x_E = \frac{a-c-2k}{2}$$

and

$$x_I = \frac{a-c+2k}{4}.$$

which is a solution if

$$\lambda_E = \frac{a-c-3y_E}{2} \geq 0$$

and

$$y_E \leq \frac{a-c}{3}.$$

Conclusion

The above calculations show that the investment in the closed-loop model is higher than in the open-loop model for the entrant as

$$\frac{a-c-2k}{2} \geq \frac{a-c-2k}{3}$$

and

$$\frac{a-c}{3} \geq \frac{a-c-2k}{3}.$$

We conclude that there is an incentive for the players in the closed-loop game to deviate from the open-loop equilibrium. Further investigation of the above considered cases also shows that the total supply is higher and prices are lower in the closed-loop model than in the open-loop model.

[66] This is also the limiting case for $x_E < y_E$.

Bibliography

50Hertz, 2016. Kennzahlen. Downloaded 02.04.2016. `http://www.50hertz.com/de/Kennzahlen`.

ACER, 2013. Capacity Remuneration Mechanisms and the Internal Market for Electricity. Pursuant to Article 11 of Regulation (EC) No 713/2009.

Amprion, 2016. Netzkennzahlen. Downloaded 02.04.2016. `http://www.amprion.net/netzkennzahlen`.

Bard, J., Loncar, T., 1991. Factors influencing the Demand for Australian Coal. ABARE - Australian Government Publishing Service.

Barroso, L., Carneiro, R., Granville, S., Pereira, M., Fampa, M., 2006. Nash equilibrium in strategic bidding: a binary expansion approach. IEEE Transactions on Power Systems 21 (2), 629–638.

Bertsch, J., Growitsch, C., Lorenczik, S., Nagl, S., 2016. Flexibility in Europe's power sector – An additional requirement or an automatic complement? Energy Economics 53, 118–131.

Boots, M., Rijkers, F., Hobbs, B., 2004. Trading in the downstream European gas market: A successive oligopoly approach. The Energy Journal 25 (3), 73–102.

Borenstein, S., Holland, S., 2005. On the Efficiency of Competitive Electricity Markets with Time-Invariant Retail Prices. The RAND Journal of Economics 36 (3), 469–493.

Bowden, B., 2012. A history of the pan-pacific coal trade from the 1950s to 2011: Exploring the long-term effects of a buying cartel. Australian Economic History Review 52 (1), 1–24.

Bushnell, J. B., Mansur, E. T., Saravia, C., 2008. Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets. American Economic Review 98 (1), 237–266.

Bussieck, M. R., Ferris, M. C., Lohmann, T., 2012. GUSS: Solving Collections of Data Related Models Within GAMS. In: Kallrath, J. (Ed.), Algebraic Modeling Systems. Vol. 104 of Applied Optimization. Springer Berlin Heidelberg, 35-56.

Capros, P., Mantzos, L., Tasios., N., DeVita, A., Kouvaritakis, N., 2010. Energy Trends to 2030 – Update 2009. Institute of Communication and Computer Systems of the National Technical University of Athens.

Cepeda, M., Finon, D., 2011. Generation capacity adequacy in interdependent electricity markets. Energy Policy 39 (6), 3128–3143.

Cramton, P., Ockenfels, A., 2012. Economics and design of capacity markets for the power sector. Zeitschrift für Energiewirtschaft 36 (2), 113–134.

Cramton, P., Ockenfels, A., Stoft, S., 2013. Capacity market fundamentals. Economics of Energy & Environmental Policy 2 (2).

Cramton, P., Stoft, S., 2005. A capacity market that makes sense. The Electricity Journal 18 (7), 43–54.

Cramton, P., Stoft, S., 2006. The Convergence of Market Designs for Adequate Generating Capacity with Special Attention to the CAISO's Resource Adequacy Problem – A White Paper for the Electricity Oversight Board.

Cramton, P., Stoft, S., 2008. Forward reliability markets: Less risk, less market power, more efficiency. Utilities Policy 16, 194–201.

Daughety, A. F., 1990. Beneficial concentration. The American Economic Review, 1231–1237.

Davison, J., 2009. The need for flexibility in Power Plants with CCS. Workshop on operation flexibility of power plants with CCS, Imperial College, London, 11th-12th November 2009.

Daxhelet, O., Smeers, Y., 2007. The EU regulation on cross-border trade of electricity: A two-stage equilibrium model. European Journal of Operational Research 181 (3), 1396–1412.

De Jonghe, C., Delarue, E., Belmans, R., D'haeseleer, W., 2011. Determining optimal electricity technology mix with high level of wind power penetration. Applied Energy 88, 2231–2238.

Denholm, P., Hand, M., 2011. Grid flexibility and storage required to achieve very high penetration of variable renewable electricity. Energy Policy 39, 1817–1830.

Dirkse, S. P., Ferris, M. C., 1995. The path solver: a nommonotone stabilization scheme for mixed complementarity problems. Optimization Methods and Software 5 (2), 123–156.

Egging, R., Holz, F., Gabriel, S. A., 2010. The World Gas Model: A multi-period mixed complementarity model for the global natural gas market. Energy 35 (10), 4016–4029.

Elberg, C., 2014. Cross-border effects of capacity mechanisms in electricity markets. EWI Working Paper (14/11).

Enke, S., 1951. Equilibrium among spatially separated markets: Solution by electric analogue. Econometrica 19 (1), 40–47.

ENTSO-E, 2011. Yearly electricity consumption data for Europe. Downloaded 02.11.2011. https://www.entsoe.eu/index.php?id=92.

ENTSO-E, 2016. European Network of Transmission System Operators for Electricity. Downloaded 02.04.2016. www.entsoe.net.

European Commission, 2013. Communication from the Commission: Delivering the internal electricity market and making the most of public intervention. C(2013) 7243 final.

European Commission, 2016a. Commission Staff Working Document – Accompanying the document 'Report from the Commission – Final Report of the Sector Inquiry on Capacity Mechanisms'. SWD(2016) 385 final.

European Commission, 2016b. EU Reference Scenario 2016 – Energy, Transport and GHG Emissiones Trends to 2050.

Finkenrath, M., 2011. Cost and performance of carbon dioxide capture from power generation. IEA Working Paper.

Fortuny-Amat, J., McCarl, B., 1981. A representation and economic interpretation of a two-level programming problem. The Journal of the Operational Research Society 32 (9), 783–792.

Fürsch, M., Hagspiel, S., Jägemann, C., Nagl, S., Lindenberger, D., Glotzbach, L., Tröster, E., Ackermann, T., 2011. Roadmap 2050 – A closer look. Cost-efficient RES-E penetration and the role of grid extensions.

Fürsch, M., Hagspiel, S., Jägemann, C., Nagl, S., Lindenberger, D., Tröster, E., 2012. The role of grid extensions in a cost-efficient transformation of the European electricity system until 2050. EWI working paper, no 12/04.

Fudenberg, D., Tirole, J., 1991. Game Theory. The MIT Press.

Gabriel, S., Zhuang, J., Kiet, S., 2005a. A large-scale linear complementarity model of the North American natural gas market. Energy Economics 27 (4), 639–665.

Gabriel, S. A., Conejo, A. J., Fuller, J. D., Hobbs, B. F., Ruiz, C., 2012. Complementarity Modeling in Energy Markets. International Series in Operations Research & Management Science. Springer New York.

Gabriel, S. A., Kiet, S., Zhuang, J., 2005b. A mixed complementarity-based equilibrium model of natural gas markets. Operations Research 53 (5), 799–818.

Gabriel, S. A., Leuthold, F. U., 2010. Solving discretely-constrained MPEC problems with applications in electric power markets. Energy Economics 32 (1), 3–14.

Germeshausen, R., Panke, T., Wetzel, H., 2014. Investigating the Influence of Firm Characteristics on the Ability to Exercise Market Power – A Stochastic Frontier Analysis Approach with an Application to the Iron Ore Market. ZEW Discussion Paper No. 14-105.

Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M., Draxl, C., 2011. The State-Of-The-Art in Short-Term Prediction of Wind Power – A Literature Overview.

Graham, P., Thorpe, S., Hogan, L., 1999. Non-competitive market behaviour in the international coking coal market. Energy Economics 21 (3), 195–212.

Growitsch, C., Hecking, H., Panke, T., 2013. Supply disruptions and regional price effects in a spatial oligopoly – an application to the global gas market. EWI working paper, no 13/08.

Haftendorn, C., Holz, F., 2010. Modeling and analysis of the international steam coal trade. The Energy Journal 31 (4), 205–230.

Harker, P. T., 1984. A variational inequality approach for the determination of oligopolistic market equilibrium. Mathematical Programming 30 (1), 105–111.

Harker, P. T., 1986. Alternative models of spatial competition. Operations Research 34 (3), 410–425.

Hecking, H., Panke, T., 2014. Quantity-setting Oligopolies in Complementary Input Markets – the Case of Iron Ore and Coking Coal. EWI working paper, no 14/06.

Hogan, W. W., 2005. On an 'energy only' electricity market design for resource adequacy. Center for Business and Government, John F. Kennedy School of Government, Harvard University.

Holttinen, H., 2005. Impact of hourly wind power variations on the system operation in the nordic countries. Wind energy 8 (2), 197–218.

Holttinen, H., Hirvonen, R., 2005. Wind Power in Power Systems. John Wiley & Sons, Ltd, Ch. 8: Power System Requirements for Wind Power, 143-167.

Holz, F., von Hirschhausen, C., Kemfert, C., 2008. A strategic model of European gas supply (GASMOD). Energy Economics 30 (3), 766–788.

Huppmann, D., 2013. Endogenous production capacity investment in natural gas market equilibrium models. European Journal of Operational Research 231 (2), 503–506.

Huppmann, D., Egging, R., 2014. Market power, fuel substitution and infrastructure – A large-scale equilibrium model of global energy markets. Energy 75, 483–500.

Huppmann, D., Holz, F., 2012. Crude Oil Market Power – A Shift in Recent Years? The Energy Journal 33 (4), 1–22.

IEA, 2011. World energy outlook 2011.

IEA, 2013. Coal Information 2013. IEA Publications Paris.

Jägemann, C., Fürsch, M., Hagspiel, S., Nagl, S., 2012. Decarbonizing Europe's power sector by 2050 – Analyzing the implications of alternative decarbonization pathways. EWI working paper, no 12/13.

Joskow, P., Tirole, J., 2007. Reliability and competitive electricity markets. The RAND Journal of Economics 38 (1), 60–84.

Joskow, P. L., 2006. Competitive Electricity Markets And Investment In New Generating Capacity. Working Papers 0609, Massachusetts Institute of Technology, Center for Energy and Environmental Policy Research.

Joskow, P. L., 2008. Capacity payments in imperfect electricity markets: Need and design. Utilities Policy 16 (3), 159–170.

Karp, L., Newbery, D. M., 1989. Time-consistent oil import tariffs. Discussion Paper 344, Center for Economic Policy Research, London.

Karp, L., Newbery, D. M., 1992. Dynamically consistent oil import tariffs. The Canadian Journal of Economics / Revue canadienne d'Economique 25 (1), 1–21.

Knaut, A., Paulus, S., 2016. When are consumers responding to electricity prices? An hourly pattern of demand elasticity. EWI working paper, no 16/07.

Kolstad, C. D., Abbey, D. S., 1984. The effect of market conduct on international steam coal trade. European Economic Review 24 (1), 39–59.

Kolstad, C. D., Burris, A. E., 1986. Imperfectly competitive equilibria in international commodity markets. American Journal of Agricultural Economics 68 (1), 27–36.

Lamadrid, A., Mount, T., Thomas, R., 2011. Geographical averaging and ancillary services for stochasticpower generation. International Journal of Innovations in Energy Systems and Power 6 (1), 13–21.

Lijesen, M. G., 2007. The real-time price elasticity of electricity. Energy Economics 29 (2), 249–258.

Lorenczik, S., Malischek, R., Trüby, J., 2017. Modeling strategic investment decisions in spatial markets. European Journal of Operational Research 256 (2), 605–618.

Lorenczik, S., Panke, T., 2016. Assessing market structures in resource markets – An empirical analysis of the market for metallurgical coal using various equilibrium models. Energy Economics 59, 179–187.

Luo, Z.-Q., Pang, J.-S., Ralph, D., 1996. Mathematical programs with equilibrium constraints. Cambridge Univ. Press.

Meyer, R., Gore, O., 2014. Cross-border effects of capacity mechanisms: Do uncoordinated market design policies countervail the goals of european market integration? Bremen Energy Working Papers 17.

Möst, D., Fichtner, W., 2010. Renewable energy sources in european energy supply and interactions with emission trading. Energy Policy 38, 2898–2910.

Murphy, F. H., Smeers, Y., 2005. Generation capacity expansion in imperfectly competitive restructured electricity markets. Operations Research 53 (4), 646–661.

Nagl, S., Fürsch, M., Jägemann, C., Bettzüge, M., 2011. The economic value of storage in renewable power systems – The case of thermal energy storage in concentrating solar plants. EWI working paper, no 11/08.

Nelson, C. H., McCarl, B. A., 1984. Including imperfect competition in spatial equilibrium models. Canadian Journal of Agricultural Economics 32 (1), 55–70.

Nicolosi, M., 2010. Wind power integration and power system flexibility - An empirical analysis of extreme events in Germany under the new negative price regime. Energy Policy 38, 7257–7268.

Nicolosi, M., 2012. The economics of electricity market integration – An empirical and model-based analysis of regulatory frameworks and their impacts on the power market. Ph.D. thesis, University of Cologne.

Paulus, M., Trüby, J., Growitsch, C., 2011. Nations as strategic players in global commodity markets: evidence from world coal trade. EWI working paper, no 11/04.

Pereira, M. V., Granville, S., Fampa, M. H. C., Dix, R., Barroso, L. A., 2005. Strategic bidding under uncertainty: A binary expansion approach. IEEE Transactions on Power Systems 20 (1), 180–188.

Prognos/EWI/GWS, 2010. Energieszenarien für ein Energiekonzept der Bundesregierung.

Richter, J., 2011. DIMENSION – A Dispatch and Investment Model for European Electricity Markets. EWI working paper, no 11/03.

RTE, 2016. Réseau de Transport d'Electricité – Prévisions de production éolienne en France. Downloaded 02.04.2016. `http://clients.rte-france.com/lang/an/visiteurs/vie/prod/prevision_production.jsp`.

Ruiz, C., Conejo, A., 2009. Pool strategy of a producer with endogenous formation of locational marginal prices. Power Systems, IEEE Transactions on 24 (4), 1855–1866.

Samuelson, P. A., 1952. Spatial price equilibrium and linear programming. The American Economic Review 42 (3), 283–303.

Sauma, E., Oren, S., 2007. Economic criteria for planning transmission investment in restructured electricity markets. IEEE Transactions on Power Systems 22 (4), 1394–1405.

Selten, R., 1965. Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. Zeitschrift für die gesamte Staatswissenschaft / Journal of Institutional and Theoretical Economics 121, 301–324.

Shaffer, S., 1995. Stable cartels with a cournot fringe. Southern Economic Journal, 744–754.

Shanbhag, U. V., Infanger, G., Glynn, P. W., 2011. A complementarity framework for forward contracting under uncertainty. Operations Research 59 (4), 810–834.

Siddiqui, S., Gabriel, S. A., 2013. An SOS1-Based Approach for Solving MPECs with a Natural Gas Market Application. Networks and Spatial Economics 13 (2), 205–227.

Stackelberg, H. V., 1952. The Theory of Market Economy. Oxford University Press, Oxford.

Takayama, T., Judge, G. G., 1964. Equilibrium among spatially separated markets: A reformulation. Econometrica: Journal of the Econometric Society 32 (4), 510–524.

Takayama, T., Judge, G. G., 1971. Spatial and Temporal Price and Allocation Models. North-Holland Publishing Company, Amsterdam.

TenneT, 2016. Netzkennzahlen. Downloaded 02.04.2016. `http://www.tennettso.de/site/de/Transparenz/veroeffentlichungen/netzkennzahlen`.

TransnetBW, 2016. Kennzahlen. Downloaded 02.04.2016. `https://www.transnetbw.de/de/kennzahlen`.

Trüby, J., 2013. Strategic behaviour in international metallurgical coal markets. Energy Economics 36, 147–157.

Trüby, J., Paulus, M., 2012. Market Structure Scenarios in International Steam Coal Trade. The Energy Journal 33 (3), 91–123.

Ummels, B., Gibescu, M., Pelgrum, E., Kling, W., 2006. System Integration of Large-Scale Wind Power in the Netherlands. IEEE Power Engineering Society General Meeting, 2006.

WCA, 2011. World Coal Association – Coal and Steel Facts 2011. World Coal Association.

Wogrin, S., Barquín, J., Centeno, E., 2013a. Capacity expansion equilibria in liberalized electricity markets: an EPEC approach. IEEE Transactions on Power Systems 28 (2), 1531–1539.

Wogrin, S., Centeno, E., Barquin, J., 2011. Generation capacity expansion in liberalized electricity markets: A stochastic mpec approach. IEEE Transactions on Power Systems 26 (4), 2526–2532.

Wogrin, S., Hobbs, B. F., Ralph, D., Centeno, E., Barquín, J., 2013b. Open versus closed loop capacity equilibria in electricity markets under perfect and oligopolistic competition. Mathematical Programming 140 (2), 295–322.

Xu, L., Threteway, D., 2012. Flexible Ramping Products – Second Revised Draft Final Proposal. California ISO (CAISO).

Yao, J., Adler, H., Oren, S. S., 2008. Modeling and computing two-settlement oligopolistic equilibrium in a congested electricity network. Operations Research 56 (1), 34–47.

Yao, J., Oren, S. S., Adler, I., 2007. Two-settlement electricity markets with price caps and Cournot generation firms. European Journal of Operational Research 181 (3), 1279–1296.

Zhuang, J., Gabriel, S. A., 2008. A complementary model for solving stochastic natural gas market equilibria. Energy Economics 30 (1), 113–147.