Essays on Market Design and Regulation in Energy Economics

Inauguraldissertation

zur

Erlangung des Doktorgrades

der

Wirtschafts- und Sozialwissenschaftlichen Fakultät

der

Universität zu Köln

2023

vorgelegt

von

M. Sc. Dominic Lencz

aus

Bad Soden am Taunus

Referent:	Prof. Dr. Marc Oliver Bettzüge
Korreferent:	Prof. Dr. Christian Tode
Tag der Promotion:	27.06.2023

Acknowledgements

First, I want to thank Prof. Dr. Marc Oliver Bettzüge and Prof. Dr. Christian Tode for supervising my thesis. Their support as well as their challenging and inspiring feedback shaped and strengthened my research over the years. Furthermore, my gratitude goes to Prof. Michael Krause, Ph.D for chairing the examination committee.

I am also grateful to the Institute of Energy Economics (EWI). On the one hand, for providing an encouraging working environment and financial support. On the other hand for the colleagues and the great conversations, the collaboration during challenging projects, the fun events and parties, as well as the exciting trips.

This thesis would not have been possible without the inspiring, motivating, and joyful teamwork with my former colleagues and co-authors Eren Çam, Samir Jeddi and Theresa Wildgrube. Further I would like to thank Fabian Arnold, Berit Hanna Czock, and Johannes Wagner for their the fruitful feedback and discussions which contributed to this thesis. My gratitude also extends to the administration and IT of the EWI.

The financial support from the Erdgas-BRidGE project, funded by the German Federal Ministry for Economic Affairs and Energy (BMWi) through research grant FKZ: 03ET4055B is gratefully acknowledged. Further my research benefited from the framework of the Hans-Ertel-Centre for Weather Research funded by the German Federal Ministry for Transportation and Digital Infrastructure through research grant BMVI/DWD 4818DWDP5A.

Finally, special thanks go to my friends and family, especially my parents who provided me with constant support and encouragement. Foremost, my thanks go to Alessia and Thilo for giving me the joy, the confidence to complete this journey, and for showing me what really matters in life.

Dominic Lencz Cologne, March 2024

Contents

List of Figures **v**

List of Tables **ix**

1. Introduction **1**

1.1. Motivation . . . 1

1.2. Outline . . . 4

1.2.1. How curtailment affects the spatial allocation of variable renewable electricity - What are the drivers and welfare effects 4

1.2.2. Complementing carbon prices with Carbon Contracts for Difference in the presence of risk - When is it beneficial and when not? . . . 4

1.2.3. Pricing short-term gas transmission capacity: A theoretical approach to understand the diverse effects of the multiplier system . . . 5

1.2.4. Internal and external effects of pricing short-term gas transmission capacity via multipliers . . . 5

1.3. Methodology, Assumptions, and Limitations . . . 6

2. How curtailment affects the spatial allocation of variable renewable electricity **9**

2.1. Introduction . . . 9

2.2. Model . . . 13

2.3. Spatial allocation under nodal pricing . . . 18

2.3.1. Capacity allocation ranges . . . 18

2.3.2. Effect of changes in the transmission capacity . . . 23

2.3.3. Effect of changes in the demand distribution . . . 25

2.3.4. Effect of changes in the availability profiles . . . 26

2.4. Spatial allocation under uniform pricing . . . 32

2.4.1. Capacity allocation ranges . . . 32

2.4.2. Effects of changes in the transmission capacity and demand distribution . . . 37
2.4.3. Effect of changes in the availability profiles . . . 39
2.5. Discussion . . . 42
2.6. Conclusion . . . 47

3. Complementing carbon prices with Carbon Contracts for Difference in the presence of risk 49
3.1. Introduction . . . 49
3.2. Carbon pricing regimes in the absence of risk . . . 53
3.2.1. Model framework in the absence of risk . . . 54
3.2.2. Policy ranking in the absence of risk . . . 58
3.3. Carbon pricing regimes in the presence of risk . . . 60
3.3.1. Model framework in the presence of risk and socially optimal production . . . 60
3.3.2. Policy ranking with damage risk . . . 62
3.3.3. Policy ranking with variable cost risk . . . 66
3.4. Carbon pricing regimes with potentially socially not optimal production . . . 71
3.4.1. Model framework in the presence of risk and socially not optimal production . . . 71
3.4.2. Policy ranking with damage risk . . . 73
3.4.3. Numerical application with risk aversion . . . 77
3.5. Discussion . . . 79
3.6. Conclusion . . . 82

4. Pricing short-term gas transmission capacity: A theoretical approach 85
4.1. Introduction . . . 85
4.2. The Model . . . 89
4.3. Results . . . 93
4.3.1. Deriving the Effects on Infrastructure Utilisation . . . 93
4.3.2. Deriving the Effects on Prices and Price Spreads . . . 98
4.3.3. Deriving the Effects on Surpluses and Welfare . . . 101
4.3.4. The regulated TSO: Transmission Tariff Adjustment . . . 104
4.4. Discussion . . . 106
4.4.1. Effects on Infrastructure Utilisation . . . 106
4.4.2. Effects on Hub Prices . . . 108

4.4.3. Effects on Surpluses and Welfare 109
4.5. Conclusion 111

5. Internal and external effects of pricing short-term gas transmission capacity via multipliers 113
5.1. Introduction 113
5.2. Identifying the main drivers 117
5.2.1. Internal effects of multipliers 117
5.2.2. External effects of multipliers 119
5.3. Methodology 120
5.3.1. Model 120
5.3.2. Assumptions and data 123
5.4. Results 125
5.4.1. Internal effects 126
5.4.2. External effects 133
5.4.3. Overall distributional effects 136
5.4.4. Comparing different optimal multiplier levels 138
5.5. Discussion 141
5.5.1. Overall effects 141
5.5.2. Regional effects 141
5.5.3. External effects and the EU optimum 142
5.6. Conclusion 144

A. Supplementary Material for Chapter 2 147
A.1. Properties of curtailment due to limited transmission capacity . . 147
A.1.1. Functional form of curtailment 147
A.1.2. Marginal curtailment given all capacity is allocated to node h 147
A.1.3. Effect of various parameters on marginal curtailment . . . 148
A.2. Historical availabilities for wind and solar in Germany and corresponding Beta distribution 151
A.3. Applied and historical densities for wind power 151
A.4. Description and explanation of formulas used to compute Figures 2.2-2.12 152
A.5. Effect of changing μ_i with the means of α_i on σ_i 153
A.6. Effects of the variance on the spatial allocation ranges when transmission capacity is high 154

B. Supplementary Material for Chapter 3 **155**
B.1. Proof of Proposition 1 155
B.2. Proof of Proposition 2 157
B.3. Proof of Proposition 3 160
B.4. Proof of Proposition 4 164
B.5. Regulatory solutions with variable cost risk and potentially socially not optimal production 168
B.6. Welfare difference compared to the social optimum in the presence of damage risk, and (ex post) potentially socially not optimal abatement due to an increase in σ_D 170

C. Supplementary Material for Chapter 4 **173**
C.1. Formal representation of the theoretical model 173
C.2. KKT points 174
C.3. Prices in region A 183
C.4. Surpluses and deadweight loss when no feasible $\underline{m}$ and $\overline{m}$ exist . 185

D. Supplementary Material for Chapter 5 **187**
D.1. Theoretical analysis 187
D.2. Reference case and model validation 188
D.3. Overview of regional price spreads 190
D.4. Overview of external effects on consumer surplus 191

Bibliography **193**

List of Figures

2.1. Schematic representation of the model setup. 17
2.2. Spatial allocation, marginal usable supply, and VRE share at different VRE penetration levels under nodal pricing. 23
2.3. Effect of changes in the transmission capacity (t) on the spatial allocation ranges under nodal pricing. 25
2.4. Effect of changes in the demand distribution on the spatial allocation ranges under nodal pricing. 26
2.5. Effect of the correlation among availability profiles on the spatial allocation ranges under nodal pricing. 28
2.6. Effect of the average in the availability profile on the spatial allocation ranges under nodal pricing. 30
2.7. Effect of the variance in the availability profile on the spatial allocation ranges under nodal pricing. 32
2.8. Spatial allocation, marginal usable and saleable supply, VRE shares and welfare losses at different VRE penetration levels under uniform pricing. 37
2.9. Effects of changes in the transmission capacity and demand distribution on the spatial allocation ranges and welfare losses under uniform pricing. 38
2.10. Effect of the correlation among availability profiles on the spatial allocation ranges and welfare losses under uniform pricing. 40
2.11. Effect of the average availability on the spatial allocation ranges and welfare losses under uniform pricing. 41
2.12. Effect of the average availability on the spatial allocation ranges and welfare losses under uniform pricing. 42

3.1. Sequence of actions in the different carbon pricing regimes. . . . 56
3.2. Market clearing. 57
3.3. Density of D and C_v following a truncated normal distribution with $P(C_v > D) = 0$. 61

3.4. Difference in welfare compared to social optimum in the presence of damage risk. 66
3.5. Difference in welfare compared to social optimum in the presence of cost risk. 71
3.6. Density of normally distributed D and C_V with $P(C_V > D) > 0$. 72
3.7. Difference in welfare compared to social optimum in the presence of damage risk and potentially welfare-reducing production. . . . 77
3.8. Difference in welfare compared to social optimum in the presence of damage risk, potentially welfare-reducing production and risk aversion. 79

4.1. Schematic representation of the model structure and the main assumptions 90
4.2. Development of the volumes for storage, ST capacity and LT capacity with respect to the multiplier (a); and development of transported volumes at time periods t_1 and t_2 with respect to the multiplier (b) 98
4.3. Development of the hub prices in region B (a) and the regional price spread between regions A and B (b), at time period t_1 and t_2 with respect to the multiplier 99
4.4. Producer, trader, consumer and TSO surpluses, and deadweight loss with respect to m 102
4.5. Volumes and prices when τ_c is adjusted such that the TSO does not earn a surplus 105
4.6. Producer, trader, consumer and TSO surpluses, and deadweight loss with respect to m when τ_c is adjusted such that the TSO does not earn a surplus 106

5.1. Schematic representation of the types of regions 120
5.2. Schematic representation of the spatial model structure 123
5.3. Capacity bookings by run-time and wasted capacity in each region when adjusting their multipliers 128
5.4. Relative change in import volumes in the peak-demand month and yearly storage volumes in each region when adjusting their multipliers 129
5.5. Absolute change in the average price (i.e. delta LT tariff) with respect to m1 level and the absolute change in the standard deviation in each region when adjusting their multipliers individually . . . 130

5.6. Consumer and storage operator surplus in each region when adjusting their multipliers individually 133
5.7. Changes in the consumer and storage operator surplus in the regions which lie downstream of Central when Central adjusts its multipliers: (a) Italy, (b) British, (c) Iberia 134
5.8. Changes in the consumer and storage operator surplus in the regions which are not directly connected to Central when Central adjusts its multipliers: (a) South East, (b) Baltic, and (c) the corresponding development of the standard deviation of Russian prices 135
5.9. Changes in the consumer and storage operator surplus in Central when (a) Italy adjusts its multipliers, (b) when South East adjusts its multipliers 136
5.10. Changes in the consumer, producer, trader, and storage surplus and welfare with respect to multipliers in the EU 138
5.11. Changes in regional consumer surplus with respect to how the multipliers are specified 140

A.1. Comparison of historical availabilities for wind and solar in Germany with the corresponding Beta distribution. 151
A.2. Historical availability densities for the years 2015-2022 and in this analysis applied densities. 151
A.3. Effect on σ when changing μ with the means of α. 154
A.4. Effect of the variance in the availability profile on the spatial allocation ranges under nodal pricing. 154

B.1. Difference in welfare compared to social optimum due to change in $P(c_v > D)$ by altering μ_D in the presence of damage risk and potentially welfare-reducing production. 171

C.1. Development of prices in region A at time periods t_1 and t_2 with respect to the multiplier 184
C.2. Surpluses and deadweight loss when no feasible $\underline{m}$ and $\overline{m}$ exist . 185
C.3. Surpluses and deadweight loss when no feasible $\underline{m}$ and $\overline{m}$ exist in the case where τ_c is adjusted 186

D.1. Simulated and the historical monthly storage levels in the EU . . 189
D.2. Simulated and the historical monthly import volumes from Russia into the EU 189

D.3. Simulated regional price levels for the gas year 2018 and the historical TTF price in the corresponding period 190
D.4. Change in the average inter-regional price spread and its standard deviation with respect to import region when each region adjusts their multipliers individually 190
D.5. The changes in consumer surplus in the regions and the total impact in the EU when multipliers are adjusted individually in the regions: (a) Central, (b) South East, (c) Baltic, (d) Italy, (e) British, (f) Iberia. 191

List of Tables

5.1. Notation used in the TIGER model extension 121
5.2. The chosen multiplier levels for the analysis 126
5.3. Multiplier levels maximising consumer surplus 139

1. Introduction

1.1. Motivation

Human activity is the cause of the climate problem,
so human action must be the solution.
António Guterres (2022)

Human activity is the cause of the climate problem. In the last century burning fossil fuels to produce goods, transport vehicles, and heat homes led to wealth and prosperity. An unabated burning of fossil fuels in the next century will lead us living in a climate hell, according to António Guterres (2022), secretary-general of the United Nations. Humanity can prevent such a climate hell by substantially reducing greenhouse gas emissions.

So far, the international community of states has failed to make a credible commitment to reduce carbon emissions. The main reason for this failure is the common nature of the climate. A state that reduces its carbon emissions receives only a fraction of the benefits, but bears all the costs of mitigation. To achieve cooperation in the fight against the climate crisis, economists propose international cooperation like climate clubs (Nordhaus, 2015) or reciprocal agreements on carbon prices (Cramton et al., 2017). While it is important to negotiate such binding agreements, a decade-long history of stalled negotiations shows that such agreements should not be expected soon.

Another approach to reduce carbon emissions is to unilaterally promote research and development of low-carbon technologies. This approach is less efficient, but it does not require cooperation between sovereign states. A well-known example is the German promotion of solar and wind power in the Renewable Energy Sources Act, introduced in the year 2000 (German Bundestag, 2000). Among other unilateral support programmes, it stimulated investment in the research and development. Between 2000 and 2020 the cost of photovoltaic panels were reduced by a factor of 20 (IEA, 2022a). The cost of wind turbines also fell. As a result of these and expected further cost reductions the IEA (2022c) expects variable renewable electricity (VRE) from solar and wind to dominate the global electricity production by 2050.

To integrate such levels of VRE, an appropriate spatial allocation is crucial. This is because, concentrating VRE capacity in the regions with the highest average availability maximises the potential supply, but is also likely to lead to high curtailments due to system constraints, such as limited transmission capacity.

The optimal spatial allocation is influenced by system characteristics, such as transmission capacity, demand distribution, VRE penetration, or VRE availability characteristics. Market designs that do not take into account nodal constraints, such as uniform pricing, lead to inefficiencies. A fundamental understanding of the effect of each system characteristic on the optimal spatial allocation and the distorting effect of uniform pricing can help to improve the spatial allocation of VRE capacity. This can reduce the costs of decarbonisation and thereby strengthen the fight against the climate crises.

In addition to generating electricity from renewable sources, consumers must be able to use the electricity, either directly or indirectly, i.e., as hydrogen produced from renewable electricity. However, most industrial processes, vehicles and heating systems run on fossil fuels. Replacing them require investments as well as research and development of new processes. Many low-carbon processes, like direct reduction steelmaking and electric arc furnaces, are not economically viable today, both in terms of investment and variable costs (Vogl et al., 2018). If the international community makes a credible commitment to a net-zero emissions world in the future, some of these processes likely become economic. Otherwise, these processes will probably remain uneconomical, and the sunk costs of the investment may threaten the viability of the investing company. As a result, risk-averse investors avoid such changes, which slows the development of low-carbon technologies and hampers efforts to tackle the climate crisis.

In a world without credible international agreements to reduce carbon emissions, national governments or a union of states, such as the European Union (EU), can promote carbon abatement technologies to combat the climate crisis. One promising approach is to offer Carbon Contracts for Difference (CCfDs) to companies that develop and invest in carbon abatement technologies. CCfDs hedge the risk of carbon price uncertainty, thereby reducing the risk of investment decisions for companies. This should encourage companies to invest in abatement technologies. However, if the abatement costs of other technologies fall faster than expected, the government is locked into supporting an inefficient abatement technology. Money that could otherwise be spend to achieve greater decarbonisation. Policymakers should therefore have a sound understanding of the implications from offering CCfDs. This should help to design CCfDs in such a way that the welfare gains from additional investment incentives exceed the welfare losses from the potential promotion of inefficient technologies.

On the path to decarbonisation, many studies assume natural gas to play a crucial role in the short and medium term due to its relatively low carbon intensity (Scharf et al., 2021). Carbon pricing can facilitate fuel switching from carbon intensive fuels such as lignite to less carbon-intensive gas. In addition, the EU may reduce costs for EU consumers, allowing for increased spending on decarbonisation, by pricing gas transmission capacity appropriately. This is because the pricing of transmission capacity can affect wholesale gas prices and thus the distribution of welfare between EU consumers and the group of international producers and traders (Hecking, 2015). One element in the pricing is the ratio between the price of short-term and long-term capacity which is called multiplier in the EU regulation (European Commission, 2017).

Understanding the diverse effects of multipliers fundamentally, may help policymakers to set them appropriately. Realised savings by EU consumers may then be used to foster investments into the abatement of carbon emissions. Consumer rent maximising multipliers may differ among EU states and induce distributional effects within the union. Understanding these effects will help to improve the setting of multipliers and thereby result in increased consumer benefits. Additionally, this would allow to arrange cross-national payments which mitigate unintended distributional effects and thereby increase the union-wide acceptance of the multiplier setting.

The individual chapters of this thesis cover different elements discussed in this motivation. Chapter 2 analyses the spatial allocation of VRE capacity, the underlying drivers, and the resulting welfare effects. Thereby the optimal allocation, and the allocation arising under uniform pricing are assessed. Chapter 3 examines how different sources of risk affect the efficiency of CCfDs and when these contracts are preferable to other policies. Chapter 4 and 5 analyses the pricing of short-term gas transmission capacity in EU including the resulting welfare and distributional effects. Thereby Chapter 4 aims to achieve a fundamental understanding by investigating the effects in two-node setting. Chapter 5 analyses a spatially more realistic setting of the EU considering upstream, downstream and transit regions. This allows to derive spatially differentiated effects and to decompose them into internal and external effects. Each chapter is based on an article to which all authors contributed equally:

1. How curtailment affects the spatial allocation of variable renewable electricity – What are the drivers and welfare effects. *EWI Working Paper 23/02* (Lencz, 2023).
2. Complementing carbon prices with Carbon Contracts for Difference in the presence of risk – When is it beneficial and when not? Joint work with Samir Jeddi and Theresa Wildgrube, *EWI Working Paper 21/09* (Jeddi et al., 2021).
3. Pricing short-term gas transmission capacity: A theoretical approach to understand the diverse effects of the multiplier system. Joint work with Eren Çam, *Energy Economics*, Vol. 95, 2021, Çam and Lencz (2021b).
4. Internal and external effects of pricing short-term gas transmission capacity via multipliers. Joint work with Eren Çam, *EWI Working Paper 21/04* (Çam and Lencz, 2021a).

The research questions and key findings of each chapter are discussed in more detail in the following. Afterwards, the methodology and the assumptions underlying the applied models are introduced.

1.2. Outline

1.2.1. How curtailment affects the spatial allocation of variable renewable electricity - What are the drivers and welfare effects

Variable renewable electricity (VRE), generated for instance by wind or solar power plants, is characterised by negligible variable costs and an availability that varies over time and space. Locating VRE capacity at sites with the highest average availability maximises the potential supply. However, potential supply must be curtailed, if system constraints prevent a local use or export. Such system constraints arise from the features defining the system, which I denote as system topology. Therefore, site choices that are unfavourable from a potential supply perspective may still be optimal from a total system cost perspective. Previous research has shown that first-best investments require nodal prices that take account of the system constraints. Market designs that do not reflect nodal prices, such as uniform pricing, typically fail to achieve optimal site choices. However, a profound theoretical understanding of the economic trade-offs involved in the optimal spatial allocation of VRE capacity is lacking. My paper contributes to filling this research gap. To do so, I develop a highly stylised model in which producers, taking into account the system topology, allocate VRE capacity in a one-shot game. Using the model, I analytically show that the optimal spatial allocation can be grouped into three spatial allocation ranges. Which of these ranges applies, I find to be highly dependent on the system topology parameters. In the first range, valid for relatively low VRE penetration levels, it is optimal to allocate all capacity to the node with the higher average availability. In the second and third range, it is optimal to allocate marginal capacity either fully or partially to the node with the lower average availability, i.e., the less favourable site from a potential supply perspective. For uniform pricing, I show that producers allocate capacity inefficiently when VRE penetration exceeds a certain threshold. The resulting welfare losses I find to be especially high when transmission capacity is low, the difference in average VRE availability is large, and demand is concentrated at the node with the lower availability.

1.2.2. Complementing carbon prices with Carbon Contracts for Difference in the presence of risk - When is it beneficial and when not?

Deep decarbonisation requires large-scale irreversible investments throughout the next decade. Policymakers propose CCfDs to incentivise such investments in the industry sector. CCfDs are contracts between a regulator and a firm that pay out the difference between a guaranteed strike price and the actual carbon price per abated emissions by an investment. Chapter 3 develops an analytical model to assess the welfare effects of CCfDs and compare it to other carbon pricing regimes.

In the model, a regulator can offer CCfDs to risk-averse firms that decide upon irreversible investments into an emission-free technology in the presence of risk. Risk can originate from the environmental damage or the variable costs of the emission-free technology. The chapter finds that CCfDs can be beneficial policy instruments, as they hedge firms' risk, encouraging investments when firms' risk aversion would otherwise inhibit them. In contrast to mitigating firms' risk by an early carbon price commitment, CCfDs maintain the regulator's flexibility to adjust the carbon price if new information reveals. However, as CCfDs hedge the firms' revenues, they might safeguard production with the emission-free technology, even if it is ex-post socially not optimal. In this case, regulatory flexibility can be welfare superior to offering a CCfD.

1.2.3. Pricing short-term gas transmission capacity: A theoretical approach to understand the diverse effects of the multiplier system

In the European Union's (EU) gas transmission system, transporting gas requires the booking of transmission capacities. For this purpose, long-term and short-term capacity products are offered. Short-term capacities are priced by multiplying long-term capacity tariffs with factors called multipliers, making them comparably more expensive. As such, the level of multipliers directly affects how capacity is booked and may significantly impact infrastructure utilisation and welfare—an issue that has not received attention in the literature so far. Using a theoretical approach, Chapter 4 shows that multipliers equal to 1 minimise costs and maximise welfare. In contrast, higher multipliers are associated with decreasing welfare. Yet, policymakers may favour higher multipliers, as multipliers greater than 1 but sufficiently low are found to maximise consumer surplus by leading to reduced hub prices and lower regional price spreads on average. These findings are expected to hold for the large majority of the EU countries. Nevertheless, situations are identified in which capacity demand can become inelastic depending on the proportion of multipliers with respect to the relative cost of transmission versus storage. In such cases, varying multipliers are found to have no effect on infrastructure utilisation, prices and welfare.

1.2.4. Internal and external effects of pricing short-term gas transmission capacity via multipliers

Chapter 4 indicates that, depending on the region, there exist multiplier levels that allow transport tariffs to be reduced and consumer surplus to be maximised. However, since multiplier levels in a region can cause externalities in other regions, it is not clear if individually optimal multipliers in regions would also lead to a joint optimum. In Chapter 5 insight into optimal multiplier levels in different regions in the EU is provided by using a numerical optimisation model to simulate the European gas dispatch. The chapter analyses the effects of multipliers in regional clusters; identify and differentiate between internal and external effects. It is shown that those effects and the individually optimal multiplier levels vary among regions depending on factors such as demand structure and storage availability. Our analysis confirms that individually adjusting multipliers in a region can cause external effects in other regions, depending largely on the location along the gas transport chain. With 92 million EUR per year, the potential EU consumer surplus gains with individually optimal multipliers is found to be 9% lower than the maximum achievable EU consumer surplus gains via multipliers. Hence, it is shown that because of the external effects of multipliers, individually optimal multipliers do not result in the EU optimum.

1.3. Methodology, Assumptions, and Limitations

The chapters of this thesis analyse different topics in Energy Economics and therefore apply different methodologies. Chapter 2 analyses the spatial allocation of VRE capacity for different market designs and system characteristics. Chapter 3 investigates the interaction between the regulator and firms investing in the decarbonisation of industrial processes in the presence of risk. Chapter 4 and 5 analyse the behaviour of traders who procure, store and transmit natural gas in the context of transmission pricing.

The methodology of the different chapters feature some commonalities. First, each chapter applies fundamental models to analyse the respective research questions. Second, the models are formulated as partial equilibrium models, isolating single markets and assuming an exogenous development in other markets. Third, the models assume fully rational players and the idealised market structure of perfect competition to assure an analytical or computational tractability. To interpret and transfer the results to the real world, it is crucial to understand the methodological approaches and particularly the underlying assumptions as well as their limitations.

The highly stylised model in Chapter 2 depicts the spatial allocation of VRE capacity by homogeneous producers under the nodal and uniform market design as well as for multiple system parameter configurations. Under nodal pricing, electricity prices reflect the nodal supply and demand, such that the producers' profit maximisation is modelled by minimising the total cost. In contrast, under uniform pricing, prices for electricity do not reflect grid limitations, such that the producers profit maximisation is modelled by minimising the costs which would occur in the absence of grid limitations. VRE availabilities are assumed to be beta distributed to account for the volatility of VRE. The resulting optimisation problem is solved analytically for the most system parameter configurations. For the remaining system parameter configurations the model is solved numerically using Python. While the model's simplicity allows to fundamentally understand the impact of various drivers on the spatial allocation of VRE capacity, additional effects likely occur when considering a more realistic setting. For example, the spatial allocation decision is modelled as a one-shot game in which producers can observe a fixed system. However, in reality VRE capacity is allocated continuously and system characteristics like transmission capacity or demand distribution evolve over time. Hence, the validity ranges identified in the model likely do not translate into consecutive phases of VRE penetration. In addition the model only considers one VRE technology, two nodes and neglects the existence of storage or demand flexibility. While the effects from storage or demand flexibility should be similar to those from reduced volatility which are analysed in the chapter, the presence of multiple VRE technologies and multiple nodes likely will introduce additional interaction effects.

Chapter 3 analyses the investment decision of heterogeneous firms in decarbonising a production process. Thereby two sources of risk and different carbon pricing regimes are analysed, including CCfDs. While the investment decision is met before the risk resolves, carbon prices are either set before or after the risk resolves, depending on the carbon pricing regime. The model is solved using backward induction to derive the sub-game perfect Nash equilibrium. Afterwards, the pricing regimes are ranked and compared to the social optimum. When analysing each risk separately, an analytical solution is gained. When applying both types of risk simultaneously the model is solved numerically. The risk is modelled as random variable which is assumed to be truncated normally distributed. Furthermore, the model assumes firms to face an utility that is exponential in profits, while the regulator is assumed to be risk neutral. Further, the absence of shadow costs of public funding is assumed. These assumptions allow to approximate the decision of the investors by a mean-standard deviation decision rule defined by Norgaard and Killeen (1980). As taxation has distortionary effects, public spending is likely to induce shadow costs of public funds. Furthermore, the degree of risk aversion of regulators and firms is hotly debated in the literature. Assuming a risk-averse regulator or less risk-averse firms would both dampen the positive welfare effects identified for CCfDs. The effect of accounting for the shadow cost of public funds would be ambiguous as it affects the price of carbon and CCfDs.

In Chapter 4 a stylised analytical model depicting the gas procurement, storage, and transmission capacity booking in the EU gas market is developed to analyse the effects of multipliers. The model considers two points in time and a network consisting of a demand and supply node which are connected by a transmission pipeline. Five groups of players, i.e., traders, producers, storage operators, the transmission system operator (TSO), and consumers interacting with each other are represented. Major assumptions are perfect competition with perfect foresight and a gas demand which is inelastic in the short run. The resulting problem coincides with the planner's problem, which is given by the maximisation of welfare. The maximisation of welfare is achieved by the minimisation of total costs arising from procurement, transport and storage. The linear cost minimisation problem is solved analytically using Karush-Kuhn-Tucker (KKT) conditions. In this analysis, capacities of production, pipeline and storage are assumed to be non-binding, representing a situation with sufficient capacities. In reality, there are some chronically congested pipelines; however, this assumption is generally representative of the overall situation in the European gas system. The assumption of perfectly inelastic demand is a common assumption for short-run gas market models; however, gas demand can nevertheless have a certain short-run elasticity, in particular in the power sector due to fuel switching between gas and coal-fired plants. Furthermore, demand reductions during the Ukrainian war showed the elasticity in the demand for heating. While assuming elastic demand would not change the main findings of the analysis, the effects of multipliers during peak prices could be more pronounced in reality in this case.

In Chapter 5, the internal and external effects of multipliers are analysed with the numerical simulation model TIGER. The model was developed at the Institute of Energy Economics (EWI) at the University of Cologne and simulates the gas dispatch in Europe under perfect competition and perfect foresight. For the analysis presented in this chapter, cost of capacity booking is included in the objective function of the TIGER model and the corresponding restrictions are specified. Analogous to Chapter 4, the model considers the interaction between traders, producers, storage operators, TSOs, as well as consumers. The model is also formulated as a linear optimisation problem with the objective to minimise total system costs. In contrast to Chapter 4, the model analyses the effects of multipliers in a more realistic setting. This is due to higher spatial and temporal resolution, and incorporation of detailed historical data on production capacities, demand profiles, pipelines, storages, and LNG terminals. In the analysis the gas dispatch for the gas year of 2017-2018 is simulated with a monthly resolution. As in the analysis presented in Chapter 4, gas demand is assumed to be perfectly inelastic. Hence, the argument regarding this assumption applies here as well. The analysis considers a simplified spatial structure with aggregated regions to better isolate and identify effects. In the real world the spatial structure with numerous interconnected transit countries is more complex. As a result the effects of multipliers can be more amplified due to the so-called tariff pancaking effect. It should also be noted that the assumption of perfect foresight in the model results in ex-post optimal capacity booking. In reality, because of risks and forecast errors, ex-ante booked capacities and ex-post needed capacities commonly differ, resulting in a higher share of unused capacities compared to findings in this chapter. In this regard, extending the model by including stochasticity in capacity demand to represent the realistic situation of imperfect information could be a part of future research.

Beyond this discussion, the respective chapters provide comprehensive descriptions of the methodological approaches.

2. How curtailment affects the spatial allocation of variable renewable electricity - What are the drivers and welfare effects

2.1. Introduction

Variable renewable electricity (VRE), generated for instance by wind or solar power plants, is characterised by negligible variable costs. Another characteristic is that the availability of VRE sources is determined by external factors, such as wind speed or solar radiation, which vary over time and space. The product of the availability and the installed capacity defines the potential supply. If the potential supply can neither be used locally nor exported, it must be curtailed. Such electricity, which could be provided free of charge, cannot be used to generate welfare. In the year 2020, according to Yasuda et al. (2022), less than five percent of the potential supply of VRE was curtailed in most countries. However, curtailment is found to increase as VRE increases in several markets in Europe, America and Asia. In Ireland and Denmark, where VRE from wind already meets 35% and 45% of demand respectively, curtailment reaches 11% and 8% (Yasuda et al., 2022). The increase in curtailment is plausible because, as VRE penetration increases, VRE production more often exceeds demand and must be curtailed when it cannot be exported or stored. Sinn (2017), who extrapolates the German VRE penetration, finds that curtailment increases exponentially if no additional measures are taken. When the VRE share doubles from 30% to 60%, VRE curtailment is found to increase from zero to 16%. For a VRE share of 90%, more than 60% of the total potential VRE supply is curtailed in the analysis of Sinn (2017). In other words, meeting 90% of demand with VRE would require capacity with a potential supply of more than 200% of demand. These figures highlight the increasing importance of curtailment in the context of VRE expansion.

To maximise welfare, curtailment should be reduced to an appropriate level: An appropriate level of curtailment balances the costs of curtailment and the costs of mitigating curtailment. The costs of curtailment arise from actions which compensate for the curtailed electricity, such as investing in additional VRE capacity. The costs of mitigating curtailment occur from actions which mitigate curtailment. Actions to mitigate curtailment are investments in storage and demand flexibility (e.g., Müller, 2017, Sinn, 2017, Zerrahn et al., 2018), network expansion (e.g., Fürsch et al., 2013), or a network-friendly allocation of VRE (e.g., Schmidt and Zinke, 2020). In this paper, I focus on the relationship between curtailment and the spatial allocation of VRE.

The spatial allocation decision when investing in VRE is driven by potential supply. As the weather differs between sites, the potential supply of VRE differs. Placing VRE capacity at sites with the highest availabilities maximises the potential supply. However, it is well known in the literature that system constraints may imply that unfavourable site choices from a potential supply perspective may still be optimal from a total system cost perspective (e.g., Green, 2007, Obermüller, 2017, Pechan, 2017, Schmidt and Zinke, 2020). The system constraints and their relevance are likely to depend on the features of the system. In the remainder of the text, I denote the features of the system as system topology. For VRE, relevant parameters describing the system topology are the transmission capacity, the spatial distribution of demand, the VRE penetration, the correlation, the average and the variance of VRE availabilities, as well as the capacities of storage and demand flexibility. Similar observations regarding the effect of the spatial allocation on the total system costs apply to any investment in generation, storage, or demand (e.g., Czock et al., 2022, Green, 2007, Müller, 2017). Therefore, it has been shown that first best investments require nodal prices that take into account system constraints arising from the system topology (Schweppe et al., 1988). Vice-versa, it follows, and has been demonstrated in numerous case studies, that market designs which do not reflect nodal prices, such as uniform pricing, typically fail to identify optimal site choices (e.g., Green, 2007, Obermüller, 2017, Pechan, 2017, Schmidt and Zinke, 2020).

Against this backdrop, I shed more light on the impact of various parameters of the system topology on the spatial allocation of VRE in the social optimum and under uniform pricing. The existing literature lacks a comprehensive understanding of these issues. Instead, most papers analyse either the effect of single system topology parameters or the effect of the market design, i.e. nodal versus uniform pricing. In addition, most studies consider a specific real-world setting. For example, Elberg and Hagspiel (2015) analyse the effect of increasing VRE penetration on the market value of VRE for the case of Germany. The authors find that market values decrease most for regions with high availability, suggesting that for high VRE penetration, it may be welfare-enhancing to allocate some capacity to regions with average or low availability. Schmidt and Zinke (2020) analyse the spatial allocation of wind capacity in Germany under nodal and uniform pricing for investments in the years 2020 to 2030. The authors find that 95% of the wind capacity added is allocated inefficiently, resulting in a welfare loss of 1.5% in terms of variable production costs. The most comprehensive analysis is provided by Pechan (2017). She analyses the impact of the correlation, the average and the variance of VRE availability on spatial allocation. She calculates the allocation in the social optimum and under uniform pricing and considers a 6-node network. Pechan (2017) finds that, under nodal pricing, producers increasingly concentrate capacity at high-availability nodes when the correlation increases and when the variance in the high-availability node is low. However, Pechan (2017) only analyses a setting where VRE serve 50% of demand, and she does not vary the transmission capacity or the demand distribution. She also performs a numerical analysis with few scenarios. Therefore, her results cannot be generalised.

To contribute to closing the research gap I analyse the following research questions:

1. From a theoretical perspective, under which states of the system topology is it welfare-enhancing to allocate some VRE capacity to sites with unfavourable potential supply?
2. How does the spatial allocation differ between a uniform pricing regime and a first-best nodal pricing regime, and what are the resulting welfare effects?

To analyse these research questions, I develop a stylised theoretical model. The model depicts the spatial allocation of VRE sources in a two-node network with limited transmission capacity. At the two nodes, consumers have a constant demand that must be satisfied by producers who can use a conventional and a VRE technology. The central element of the model is that producers decide how to spatially allocate VRE capacity. I model the spatial allocation of all VRE capacity as a one-shot game where producers consider a specific system topology, i.e. one specific configuration of transmission capacity, spatial distribution of demand, VRE penetration and VRE availability. This differs from reality, where VRE penetration and other system topology parameters dynamically evolve over time. The implications of assuming a one-shot game I discuss in Chapter 2.5. The model considers availabilities which vary over time and between the nodes. The temporal sequence of availabilities I refer to as availability profile. The average availability I assume to be higher in one node (i.e. high-availability node) compared to the other node (i.e. low-availability node). The effect of storage and demand flexibility I do not analyse in the model itself to ensure an analytical solution. The analytical solution is crucial to gain a profound theoretical understanding. To still shed light on the effect of storage and demand flexibility, I discuss the effects qualitatively based on the model results and findings from other papers in Chapter 2.5. To analyse the relationship between spatial allocation under a first-best nodal pricing regime and a uniform pricing regime, I solve the model for both market designs.

The main findings of the analysis regarding the first research question are as follows: The optimal spatial allocation can be grouped into three spatial allocation ranges that are valid for different levels of VRE penetration. For low levels of VRE penetration, all VRE capacity should be allocated to the high-availability node (i.e. *high-availability deployment range*). For such levels of VRE penetration, it is not welfare-enhancing to allocate some VRE capacity to sites with unfavourable potential supply. For higher levels of VRE penetration, resulting in curtailment that eliminates the advantage arising from higher average availability, it is optimal to allocate the marginal capacity only to the low-availability node (i.e. *low-*

availability deployment range). This is because marginal capacity allocated to the high-availability node would result in increasing marginal curtailment at the high-availability node, while small capacities at the low-availability node do not need to be curtailed. For even higher levels of VRE penetration, resulting in curtailment at the low-availability node, it is optimal for producers to split marginal capacity between the two nodes (i.e. *split capacity deployment range*). Thus, at higher levels of VRE penetration, it is welfare-enhancing to allocate (some) VRE capacity to less favourable sites from the perspective of potential supply. The VRE penetration levels, which mark the cut-off points between the three ranges, I derive analytically. The results imply, that the cut-off points depend on the parameter configuration of the system topology.

Therefore, the system topology affects the width of the *high-* and *low-availability deployment range* and the capacity split under the *split capacity deployment range*. Increasing the transmission capacity and demand share at the high-availability node widens the *high-availability deployment range* and narrows the *low-availability deployment range*. In the *split capacity deployment range*, more capacity is allocated to the high-availability node. A higher correlation between the nodal availability profiles increases the share of capacity allocated to the high-availability node in the *split capacity deployment range*. Higher availabilities at the low-availability node narrow the *high-availability deployment range* so that the overall share of VRE allocated to the low-availability node increases. The impact of nodal availability profiles is found to be influenced by transmission capacity. In the case of correlation, the impact of changes in correlation increases with increasing transmission capacity. The direction of the effect of changes in the availability and variance in the *split capacity deployment range* even depends on the transmission capacity. Increasing the average and decreasing the variance of the nodal availability increases the nodal share when the transmission capacity is high. The opposite happens when the transmission capacity is low. Therefore, the availability profiles alone are not sufficient to indicate the optimal spatial allocation but need to be considered in combination with the level of transmission capacity.

Regarding the second research question, my analysis provides the following insights: Under uniform pricing, producers allocate capacity only to the high-availability node for higher VRE penetration levels than socially optimal. This is because network constraints that would induce producers to allocate capacity more in line with demand are ignored. Welfare losses occur when marginal curtailment due to limited transmission capacity exceeds the average availability advantage of the high-availability node. Welfare losses increase with increasing VRE penetration until VRE penetration is sufficiently high that differences in availability profiles provide an incentive to allocate some capacity to the low-availability node. Welfare losses under uniform pricing decrease with the level of transmission capacity and increase with the need for transmission.

From a theoretical perspective, my contribution is threefold: First, using a highly stylised model, I show that the optimal spatial allocation can be grouped into three ranges. I analytically derive the VRE penetration levels that separate the three ranges, so that the results can be applied to any feasible configuration of the system topology. Second, I identify the effect of various parameters of the system topology on the optimal spatial allocation of the ranges. And third, I identify the allocation under uniform pricing and the resulting welfare loss, and show how the welfare loss is affected by the different parameters of the system topology.
Due to my model's simplicity, I analyse a highly stylised setting. When analysing a more realistic setting additional effects will occur. Such effects from considering a more realistic setting on my theoretical findings are discussed in Chapter 2.5. Combining the findings from the theoretical analysis with the considerations from the discussion can help policymakers when designing policies that affect the spatial allocation or when considering a change in the market design. Investors can use the results when trying to find the profit-maximising allocation of VRE investments.

2.2. Model

I develop a theoretical model to analyse the effect of the VRE penetration, the transmission capacity, the demand distribution, the VRE availabilities, and the market design on the spatial allocation of VRE capacity. The effect of storage and demand flexibility I do not analyse in the model itself to ensure an analytical solution. To shed light on the effect of storage and demand flexibility, I discuss the effects qualitatively based on the model results and findings from other papers in Chapter 2.5.

The model considers the interaction between profit-maximising producers in a perfectly competitive environment, consumers, and a regulator. The players act in a network consisting of two nodes, h and l (i.e., $i \in (h, l)$), which are connected by a transmission line with the transmission capacity t. Furthermore, I define the model to have three stages, the regulation stage (τ_1), the spatial allocation stage (τ_2), and the market clearing stage (τ_3). As the model is solved by backward induction, the explanation starts with the last stage.

The market clearing stage (τ_3) takes place for a time interval ranging from 0 to 1. The time interval is divided into n periods with equal length, where n goes towards infinity ($n \to \infty$). I assume the consumers' demand at each node (d_i) is constant among all n periods, inelastic and exceeds the transmission capacity (i.e., $d_i > t$). I denote the demand in terms of power[1], such that the demand for a specific period r and the total demand in stage τ_3 coincide. Due to assuming a

[1] Demand denoted in terms of power defines the rate at which electricity is retrieved from an electrical network. A well known unit for power is Watt.

time interval ranging from 0 to 1, the demand in stage τ_3 in terms of energy[2] and power coincide as well (i.e., $\sum_{r=1}^{n} \frac{1}{n} d_i = d_i$).

I assume that the producers can satisfy the demand with one conventional and one VRE technology. This assumption differs from the situation in most countries, where several conventional and at least two VRE technologies, namely wind and solar, are employed. On the one hand, the assumption of one conventional and one VRE technology allows me to derive the effect of the spatial allocation of VRE analytically. This provides a general understanding of the impact of the system topology. On the other hand, the simplification of multiple production technologies into one conventional and one VRE technology highlights the stylised nature of the model. The implications of this simplification are discussed in Chapter 2.5. The conventional technology induces constant marginal production costs. When supplying one unit throughout the time interval ranging from 0 to 1 the costs are equal to c.[3]. I assume the conventional capacity at nodes h and l to exceed the respective demand. As the producers operate in a competitive environment, they cannot charge prices above the marginal costs c of the conventional technology. Therefore, producers cannot make profits by investing in the conventional technology and have no incentive to invest in this technology.[4]

The VRE technology induces zero marginal costs. However, the capacity is limited to the total VRE capacity v (also denoted as VRE penetration). The capacity at nodes h and l is represented by V_h and $V_l = v - V_h$. The VRE capacity cannot always produce at full capacity, but only at the availability ($avail_{i,r}$) times the installed capacity (V_i). The availability varies between the nodes (denoted by the index i) and the periods (denoted by the index r). Within the time interval ranging from 0 to 1, I assume the availabilities to constantly change. The temporal sequence of all availabilities occurring during the interval I define as availability profile (i.e., $AVAIL_i$). I assume the availability profile to be a deterministic sequence which is characterised by an average (μ_i) and variance (σ_i^2). The values within the availability profile are assumed to be beta distributed. The beta distribution is chosen because it features positive densities only for values in the interval $[0, 1]$, as VRE availabilities do in reality.[5] Hence, the density function describing the distribution of the availabilities within the availability profile is given by:

$$f_{AVAIL_i} = \int_0^1 \frac{1}{B(\alpha_i, \beta_i)} x^{\alpha_i - 1} (1 - x)^{\beta_i - 1} dx \tag{2.1}$$

[2] Demand denoted in terms of energy defines the sum of electricity retrieved from an electrical network.

[3] These costs include fuel costs as well as other all relevant variable costs such as costs for carbon emission allowances.

[4] In reality, investment in conventional capacity can be observed. There are two main reasons for this. First, conventional capacity often does not exceed demand, so producers can make profits by offering capacity in periods of scarcity. Such scarcity tends to persist as older plants are retired. Second, the marginal cost of building new conventional technologies tends to fall over time, so that new capacity can be profitable even in the absence of scarcity.

[5] In accordance with the Moivre-Laplace theorem, assuming n to converge towards infinity allows to represent the discrete binomial beta distribution by the continuous beta distribution.

Choosing the parameters α_i and β_i appropriately, results in density functions similar to the realised densities of wind or solar power plants, as shown in Chapter A.2.

The potential supply in terms of power at node i and period r (denoted by $PS_{i,r}$) is given by the product of the respective availability and the installed capacity:

$$PS_{i,r} = V_i \cdot avail_{i,r} \tag{2.2}$$

The potential supply within the time interval ranging from 0 to 1 (both defined in terms of power and energy) is defined by:

$$PS_i = V_i \cdot \frac{1}{n} \sum_r avail_{i,r} \tag{2.3}$$

As the beta distribution describes the deterministic availability profile's distribution, the potential supply can be expressed as follows:

$$\begin{aligned} PS_i &= V_i \cdot f_{AVAIL_i} \\ PS_i &= V_i \int_0^1 \frac{1}{B(\alpha_i, \beta_i)} x^{\alpha_i - 1} (1-x)^{\beta_i - 1} dx \end{aligned} \tag{2.4}$$

By integrating Equation 2.4 and by considering $\mu_i = \frac{\alpha_i}{\alpha_i + \beta_i}$ the PS_i can be simplified as follows[6]:

$$PS_i = V_i \mu_i \tag{2.5}$$

I assume $\mu_h > \mu_l$, so that I call node h as *high-availability node* and node l as *low-availability node*. Further, I assume the joint distribution of nodal availability profiles to be deterministic as well. This implies, that the availabilities occurring jointly at the high- and low-availability node are known at any period (r) by the producers. The availabilities at the two nodes can be correlated, with $\rho_{h,l} \in (-1, 1)$ being the correlation coefficient. Highly correlated availabilities (i.e., $\rho_{h,l}$ close to 1) imply that a high availability in h tends to coincide with a high availability in l and vice versa. If the availabilities are barely correlated (i.e., $\rho_{h,l}$ close to 0), high availabilities in h are similarly likely to be accompanied by high or low availabilities in l. Independent of the correlation coefficient, I assume that within the sequence of n periods (with $n \to \infty$) there is a period with an availability equal to 1 at both nodes simultaneously (i.e., $avail_{h,r} = avail_{l,r} = 1$).

The potential VRE supply which can neither be consumed locally nor be exported is curtailed. The curtailment I denote with K. The difference in global

[6]Next to mean the variance is defined by $\sigma_i^2 = \frac{\alpha_i \beta_i}{(\alpha_i+\beta_i+1)(\alpha_i+\beta_i)^2}$. Both parameters, α and β, affect the average and the variance simultaneously. However, for $\mu_i \in [0.2, 0.4]$, increasing α_i primarily increases μ_i, while σ_i is barely affected. A numerical example showing these effects I provide in Chapter A.5.

potential supply ($\sum_i PS_i$) and curtailment (K) I define as usable supply (US):

$$US = \sum_i PS_i - K \tag{2.6}$$

The curtailment (K) can be grouped into to types. First, curtailment can arise when limited transmission capacity prevents the export of potential supply to the neighbouring node (K_i^t). Second, curtailment can arise when the global potential supply excluding the curtailment from limited transmission capacity exceeds the global demand (K^d).

$$K = \sum_i K_i^t + K^d \tag{2.7}$$

In the spatial allocation stage (τ_2), the producers allocate the VRE capacity between nodes h and l. The respective capacities I define as V_h and $V_l = v - V_h$. I model the spatial allocation of all VRE capacity as a one-shot game where producers consider a particular system topology, i.e. transmission capacity, spatial distribution of demand, VRE penetration and VRE availability. This differs from reality, where VRE penetration increases continuously and other system topology parameters also evolve over time. The implications of assuming a one-shot game I discuss in Chapter 2.5.

Producers choose the allocation between h and l such that their profits are maximised. When deciding on the spatial allocation, the producers have perfect foresight, i.e., they know the nodal demand and the nodal availability profiles. In addition, producers take into account the underlying market design as well as the total level of VRE capacity. All parameters presented, namely the demand (d_i), the transmission capacity (t), the VRE penetration (v), and the parameters determining the nodal availability profiles (μ_i, σ_i, $\rho_{h,l}$) define the system topology.

In this paper, I focus on the optimal spatial allocation of VRE and not on the optimal capacity (v). Therefore, I assume that the regulator defines the total VRE capacity in the regulation stage (τ_1).[7, 8]

In addition, the regulator defines the market design. The market design options the regulator can implement are nodal and uniform pricing. Under nodal pricing the nodal demand and supply define the nodal price, considering the network transmission capacity. The prices at both nodes may differ, as shown in the following example: Assume the potential VRE supply at node h at period r exceeds the nodal demand plus the transmission capacity ($PS_{h,r} > d_h + t$). In

[7] To identify the optimal v in the model at hand one would have to consider the capital costs of the VRE technology and minimise the total costs with respect to the total VRE capacity.

[8] When the regulator defines v, she may use auctions which allow for negative prices and contain an obligation to build the purchased capacity. Such a process would ensure the total capacity is sold in this stage, built and allocated in the spatial allocation stage and used in the market clearing stage.

that case, some VRE at node h needs to be curtailed, and the VRE technology sets the nodal price. As the VRE technology features zero marginal costs the price at node h is zero (i.e., $p_{h,r} = 0$). At the same time, the potential VRE supply at node l plus the VRE imports from node h is below the demand in l (i.e., $PS_{l,r} < d_l - t$). As a result, some conventional supply is required to satisfy the demand at node l, such that the conventional technology sets the price (i.e., $p_{l,r} = c$). Depending on the system topology, the price at both nodes will be c in some periods and 0 in others. The proportion of periods where the conventional technology sets the price represents the average nodal price $\overline{p_i}$.

Under uniform pricing, the price producers receive is determined by the global demand ($d_h + d_l = d_{h+l}$) and global supply. Thus, the market design implicitly ignores transmission constraints and yields identical prices at both nodes (i.e., $p_{h,r} = p_{l,r}$). The conventional technology sets the price (i.e., $p_i = c$) when the global demand exceeds the global VRE potential. When the global VRE potential exceeds the global demand, the VRE technology sets the price (i.e., $p_{i,r} = 0$). The proportion of periods where the conventional technology sets the price represents the average price $\overline{p_i}$. Under uniform pricing, periods may arise where supply sold with VRE supply cannot be dispatched to the consumers due to limited transmission capacity. I assume such VRE supply is curtailed, but producers still receive the market price. This is similar to the compensation applied in multiple countries with uniform pricing, such as Germany, Denmark, Italy or Japan (Bird et al., 2016). To ensure demand is met, the curtailed VRE supply is replaced by an additional conventional supply at the other node. Such conventional supply I denote by redispatch. These costs are assumed to be borne by the consumers. Figure 2.1 schematically represents the model setup including the three stages.

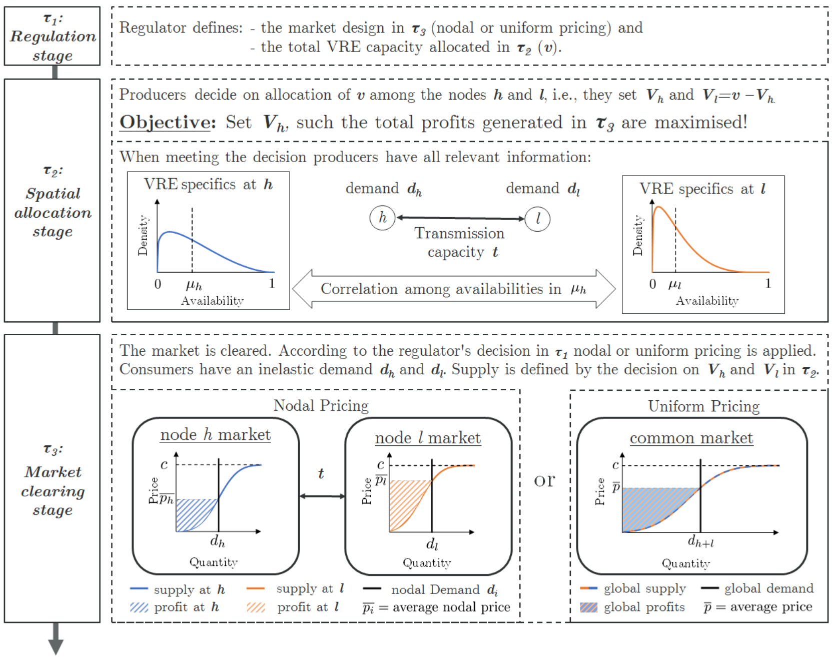


Figure 2.1.: Schematic representation of the model setup.

2.3. Spatial allocation under nodal pricing

In this chapter, I analyse the socially optimal spatial allocation under nodal pricing. One crucial aspect influencing the spatial allocation is the occurrence and the level of curtailment at the high- and low availability node. This curtailment depends on the relative level of VRE penetration with respect to the other parameters, such as nodal demand or transmission capacity. In a first step, I identify three ranges of capacity allocation, valid for different relative levels of VRE penetration. In a second step, I assess the effect of transmission capacity, nodal demand distribution, and VRE availability profiles on the width of the ranges and spatial allocation within the ranges. Within the analysis, I consider the interactions among the parameters.

2.3.1. Capacity allocation ranges

In this subchapter, I derive the capacity allocation ranges under nodal pricing for different relative levels of VRE penetration. I conclude that:

Finding NP 1. *The optimal spatial allocation can be grouped into three spatial allocation ranges which are valid for different levels of VRE penetration. For low VRE penetration levels, producers allocate marginal capacity to the high-availability node (high-availability deployment range). For higher VRE penetration levels, producers allocate marginal capacity to the low-availability node (low-availability deployment range), and for even higher VRE penetration levels, producers split marginal capacity among the two nodes (split capacity deployment range).*

Explanation. Under nodal pricing, perfect competition and perfect information, the producers' maximisation of profits coincides with the minimisation of the total costs (TC). Hence, producers spatially allocate the capacity such that the total costs are minimised. Within the total costs, the capacity level at node l (V_l) can be substituted with $v - V_h$, such that V_h is the only decision variable. The total costs are given by the conventional production times their marginal costs (c). The level of conventional production is given by the global demand (d_{h+l}) minus the usable supply (US) generated by VRE capacity. As the usable supply is defined by the difference between the global potential supply ($\sum_i PS_i$) and curtailment (K), the total costs are given by:

$$\min_{V_h} \; TC = \left(d_{h+l} - \underbrace{\left(\sum_i PS_i(V_h) - K(V_h) \right)}_{\text{usable supply}} \right) c \quad \text{for } V_h \in [0, v] \tag{2.8}$$

By substituting $PS_i(V_h)$ with the definition given in Equation 2.5 the objective function can be rewritten as follows:

$$\min_{V_h} TC = \left(d_{h+l} - \underbrace{\Big(V_h \mu_h + (v - V_h)\mu_l - K(V_h)\Big)}_{\text{usable supply}} \right) c \quad \text{for } V_h \in [0, v] \tag{2.9}$$

The global demand (d_{h+l}) is a constant and hence independent of V_h. Further, the cost parameter c is positive by definition. Hence, the total costs are minimised by maximising the level of usable supply (US).

Case 1: No curtailment

In the absence of curtailment the usable supply coincides with the potential supply. Hence the objective function is minimised by maximising the potential supply. Since $\mu_h > \mu_l$ by definition the total costs are minimised by allocating all capacity to node h.

Curtailment is absent if the installed capacity at node h does not exceed the nodal demand plus the transmission capacity (i.e., $v < d_h + t$), such that:

$$V_h^*(v) = v \quad \text{if: } v \leq d_h + t \tag{2.10}$$

Case 2: Curtailment at node h

When VRE penetration exceeds the nodal demand at the high-availability node plus the transmission capacity ($v > d_h + t$), increasing the capacity at node h induces curtailment (K). The curtailment in this case arises at node h due to limited transmission capacity (K_h^t). The existence of curtailment implies that the usable supply is smaller than the potential supply. The optimal allocation in this case depends on the level of curtailment. For VRE penetration levels represented in Case 2, marginal curtailment which only occurs at node h is strictly monotonically increasing in the VRE penetration, given that capacity is only allocated to node h (i.e. $V_h = v$).[9] Due to the strict monotonicity of marginal curtailment occurring at node h two sub cases arise:

Case 2a: Allocating capacity to node h

In Case 2a it is optimal to allocate all capacity at node h. This requires the marginal usable supply to be higher at node h than at node l when increasing the VRE penetration (v) marginally. Hence, in this case, a marginal increase in v evaluated at $V_h = v$ results in a marginal curtailment which is smaller than the delta in marginal potential supply between the high- and the low-availability node:

$$\left.\frac{\partial K}{\partial v}\right|_{V_h = v} \leq \mu_h - \mu_l \tag{2.11}$$

[9] I derive the strict monotonicity mathematically in Supplementary Material A.1.2.

The strict monotony of curtailment in v, given $V_h = v$ implies that there is a VRE penetration level at which marginal curtailment reaches the delta in marginal potential supply between the high- and the low-availability nodes (i.e., $\frac{\partial K}{\partial v}\big|_{V_h=v} = \mu_h - \mu_l$). Hence, the strict monotony of curtailment implies the existence of a cut-off point between Case 2a and Case 2b which I discuss below. The VRE penetration which induces $\frac{\partial K}{\partial v}\big|_{V_h=v} = \mu_h - \mu_l$ marks the highest VRE penetration at which producers allocate all capacity to the high-availability node. The cut-off point I denote by:

$$v^{H|L} = v\left[\frac{\partial K}{\partial v}\bigg|_{V_h=v} = \mu_h - \mu_l\right] \tag{2.12}$$

Case 2b: Allocating capacity to node l

When v exceeds $v^{H|L}$, it is no longer optimal to allocate all capacity to node h. For such higher level of v a marginal increase in VRE penetration evaluated at $V_h = v$ results in a marginal curtailment which is higher than the delta in marginal potential supply between the high- and the low-availability node:

$$\frac{\partial K}{\partial v}\bigg|_{V_h=v} > \mu_h - \mu_l \tag{2.13}$$

This implies the marginal usable supply when allocating a marginal capacity unit to node h is lower than the marginal potential supply at node l. At node l curtailment is absent for initial capacity allocations. Hence, marginal potential supply at node l coincides with marginal usable supply. Hence, for such VRE penetration levels it is optimal to allocate marginal capacity to node l.

As stated in Chapter 2.2, I assume that an availability of 1 occurs simultaneously at both nodes (i.e., $avail_{h,r} = avail_{l,r} = 1$). Further I know from Case 1 and 2a that $V_h \geq d_h + t$. Hence, when VRE penetration at node l reaches $d_l - t$, there is a period at which supply can fully serve the demand at node h and l.[10] Allocating further capacity to node l would yield to curtailment. Hence, $V_l = d_l - t$ marks the end of Case 2b. The overall VRE penetration at this cut-off point is given by:

$$v^{L|S} = v^{H|L} + (d_l - t) \tag{2.14}$$

Summarising the results of Case 2, results in the following optimal spatial allocation:

$$V_h^*(v) = \begin{cases} v & \text{if: } \; d_h + t < v \leq v^{H|L} \\ v^{H|L} & \text{if: } \; v^{H|L} < v \leq v^{L|S} \end{cases} \tag{2.15}$$

[10] When both $avail_{h,r}$ and $avail_{l,r}$ are equal to 1, the VRE supply at node h can serve the nodal demand (d_h). Further VRE exports from node h to node l of t are feasible. The residual demand at node l, which is given by $d_l - t$ can be served by $V_l = d_l - t$.

Case 3: Curtailment at both nodes

When VRE penetration exceeds $v^{L|S}$ curtailment occurs at both nodes. In addition to curtailment due to limited transmission capacity occurs (K_h^t), which occurs in Case 2, also curtailment due to global potential supply exceeding global demand (K^d) occurs in this case.[11]

At a penetration level of $v^{L|S}$, which marks the lower boundary of Case 3, the marginal curtailment at node h exactly compensates for the advantage in the nodal potential supply. As a result, the marginal usable supply at node h and l coincide. When adding marginal capacity, such that $v^{L|S}$ is exceeded, it can either be allocated to node h, to node l, or split among the nodes. When allocating the capacity to the high-availability node the nodal marginal usable supply would decrease as marginal curtailment increase. First, marginal curtailment due to limited transmission capacity would increase as such curtailment is strictly monotonically increasing in nodal VRE penetration (see Supplementary Material A.1.2). Second, curtailment due to global supply exceeding global demand would start to occur. Allocating the capacity to the low-availability node would also decrease the nodal marginal usable supply as curtailment due to global supply exceeding global demand would start to occur. Hence, in order to achieve an marginal usable supply which coincides at both nodes, marginal capacities have to be split among the two nodes.

When splitting of capacity is optimal, the first-order condition is fulfilled:

$$\frac{\partial TC}{\partial V_h} = -\left(\mu_h - \mu_l - \frac{\partial K}{\partial V_h}\right)c = 0 \tag{2.16}$$

The optimal spatial allocation is then given by:

$$V_h^*(v) = V_h[\frac{\partial TC}{\partial V_h} = 0] \quad \text{if: } v > v^{L|S} \tag{2.17}$$

For relatively low VRE penetration levels within Case 3, the majority of the marginal capacity should be allocated to node l. This is because for $V_l \in (d_l - t, d_l + t)$ curtailment at node l only occurs when global supply exceeds global demand, while curtailment due to limited transmission capacity remains absent. With increasing v, the capacity at the low-availability node (V_l) eventually exceeds $d_l + t$, also inducing curtailment due to limited transmission capacity at node l. For such level of VRE penetration capacity should be split more equally among the two nodes.

[11] The level of K^d depends on the overall VRE capacity (v), the capacity allocation (V_i), the density of the nodal availability profiles (f_{AVAIL_i}), and the deterministic joint distribution of the nodal availability profiles. The joint availability distribution depends on the nodal beta distribution parameter α_i, β_i and the correlation $\rho_{h,l}$ among the nodal availability profiles and cannot be derived analytically. Hence, K^d cannot be derived analytically.

Defining the capacity allocation ranges

Based on the results from the case analysis the ranges arise:

$$V_h^*(v) = \begin{cases} v & \text{if: } v \leq v^{H|L} \\ v^{H|L} & \text{if: } v^{H|L} < v \leq v^{L|S} \\ V_h[\frac{\partial TC}{\partial V_h} = 0] & \text{if: } v > v^{L|S} \end{cases} \quad (2.18)$$

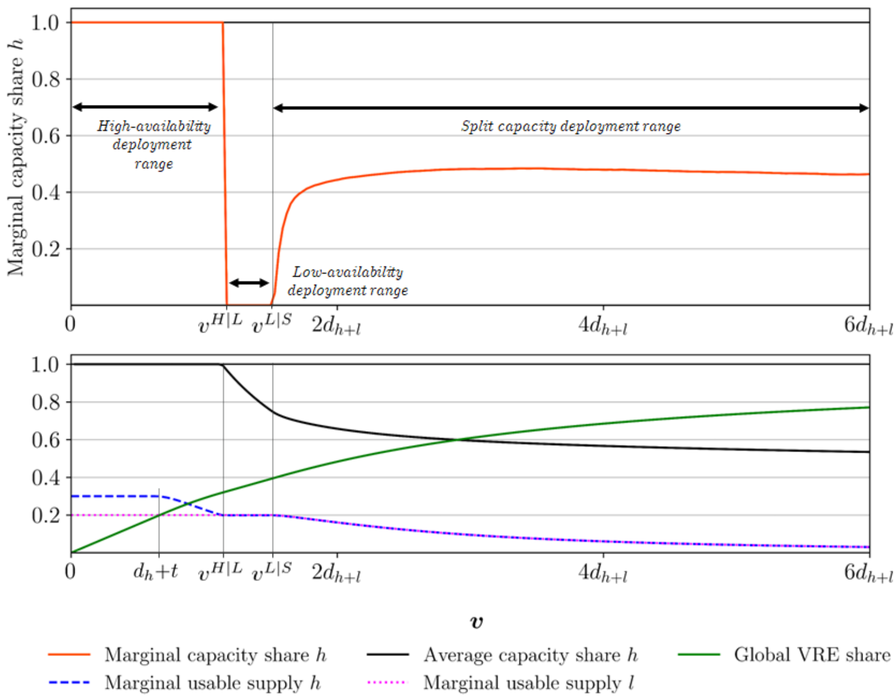
For relatively low VRE penetration ($v < v^{H|L}$) capacity is allocated only to the high-availability node. I denote this range as *high-availability deployment range*. For higher VRE penetration levels ($v^{H|L} < v \leq v^{L|S}$), producers allocate capacity to the low-availability node. I denote this range as *low-availability deployment range*. For even higher VRE penetration levels ($v > v^{L|S}$), producers split the capacity among the two nodes. This range I denote as *split capacity deployment range*.

End of Explanation.

Figure 2.2 demonstrates the insights from Finding NP 1 numerically. The optimisation problem solved numerically to generate this and the remaining figures in this chapter resembles the optimisation problem described in Chapter 2.2 and solved in this chapter. Differences in the formulation arising from using a numerical instead of an analytical approach are described in Supplementary Material A.4. The availability density parameters are chosen to resemble the availabilities for wind in the north (h) and south (l) of Germany. A comparison of the assumed and the historical German availability density for wind power is presented in Supplementary Material A.3. The demand is split equally among the nodes, and the transmission capacity can transmit one-fourth of the nodal demand. The orange line in the upper diagram displays the marginal capacity share allocated to node h. A value of 1 indicates that producers allocate marginal capacity units to node h, while 0 indicates the marginal capacity unit is allocated to node l. Values in between imply that producers split marginal capacity among the two nodes.

For low levels of VRE penetration, all capacity is allocated to node h (i.e., *high-availability deployment range*). For $v < d_h + t$ curtailment is absent. For $v = d_h + t$, VRE supplies roughly 20% of the demand. For higher VRE capacity levels, marginal curtailment is higher such that the marginally usable supply is lower. For $V^h = v^{H|L} = 1.1d_{h+l}$, the marginal usable supply is identical at both nodes, and the *high-availability deployment range* ends. At this point, VRE supplies 33% of the demand.

When the global VRE share ranges between 33-40%, capacity is allocated solely to node l. The VRE penetration marking the shift from the *low-availability deployment range* to the *split capacity deployment range* is given by $v^{L|S} = 1.1d_{h+l}+(d_l-h)$. At this VRE penetration level $V_l = d_l - l$ holds. For higher VRE shares, the capacity is split among the nodes. For low levels of VRE penetration within the *split capacity deployment range* (40-45% global VRE share) the majority of marginal capacity is allocated to node l . When VRE share rises above 45%, producers allocate roughly half of the capacity to each node.



Parameter values: d_i=50, t=$\frac{1}{4}d_i$, $B_h(1.071, 2.5)$→μ_h=0.3, σ_h=0.21, $B_l(0.625, 2.5)$→μ_l=0.2, σ_l=0.20, $\rho_{h,l}$=0.6.
X-axis values: $v^{H|L} = 1.1d_{h+l}$ and $v^{L|S} = 1.1d_{h+l} + (d_l - t)$.

Figure 2.2.: Spatial allocation, marginal usable supply, and VRE share at different VRE penetration levels under nodal pricing.

Additionally, the figure highlights the relevance of curtailment with increasing VRE penetration. At $v = 6d_{h+l}$, the potential supply exceeds 1.5 times the global demand. However, the VRE share remains below 80%. This is because roughly 50% of VRE supply is curtailed (not shown in the figure). The marginal curtailment at such a high VRE penetration level even reaches 85% at node l and 90% at node h (not shown in the figure).

Based on these observations, the question arises of how the transmission capacity, demand distribution, and availability profiles affect the capacity allocation ranges.

2.3.2. Effect of changes in the transmission capacity

In this subchapter, I derive the effect of changes in the transmission capacity (t) on the width of the capacity allocation ranges and the capacity split in the *split capacity deployment range*. Thereby the assumption $t < d_i$ (stated in Chapter 2.2) is relaxed. Based on the analysis, I conclude:

Finding NP 2. *Under nodal pricing, increasing the transmission capacity t widens the high-availability deployment range and narrows the low-availability deployment range. For $t \geq d_l$, the low-availability deployment range disappears. In the split capacity deployment range, the share of the high-availability node increases with increasing t.*

Explanation. Increasing the transmission capacity widens the *high-availability deployment range* as producers are willing to allocate capacity solely to the high-availability node for higher VRE penetration levels. This arises due to two effects: First, when increasing t by one unit, it is possible to add one unit of capacity at node h without inducing curtailment (refers to Case 1 of Finding NP 1).

Second, for v valid in Case 2 of Finding NP 1, marginal curtailment increase at a lower rate with increasing t.[12] The lower increase in marginal curtailment implies that the marginal usable supply is reduced at a lower rate. As a result the cut-off point between the *high-* and *low-availability deployment range* ($v^{H|L}$) is reached for higher levels of v. As both parts of the *high-availability deployment range* are widened, the range is widened as a whole.

With increasing the transmission capacity, the *low-availability deployment range* is narrowed. This is because the width of the range is given by $v^{H|L} - v^{L|S} = d_l - t$. For $t \geq d_l$, the *low-availability deployment range* disappears because supply produced at node h can serve the entire demand at node l. As a result, curtailment at node l already occurs for initial capacities. Hence, there is no VRE penetration level when producers are incentivised to allocate additional VRE units solely to node l.

In the *split capacity deployment range*, producers increasingly allocate capacity to the high-availability node when transmission capacity increases. This is because, with increasing t, network restrictions get less relevant, such that producers can increasingly exploit the more favourable VRE conditions at node h.

End of Explanation.

Figure 2.3 demonstrates the insights from Finding NP 2 numerically. The numerical example displays the marginal allocation share at node h and the marginal usable supply. Assumptions regarding the nodal demand (d_i), the availabilities ($B(\alpha_i, \beta_i)$) and the correlation among the availabilities ($\rho_{h,l}$) are identical to Figure 2.2.

First, marginal usable supply is constant for $V_h \leq d_h + t$, such that the first part of the *high-availability deployment range* gets wider with increasing t. Second, the increase in marginal curtailment (i.e., the reduction in marginal usable supply) is dampened with increasing t. Hence, the second part of *high-availability deployment range* gets wider. While for $t = \frac{1}{4}d_i$ the width the range is roughly $1.1d_i$, it is 50% wider for $t = \frac{3}{4}d_i$.

[12] I show this effect mathematically in Supplementary Material A.1.3.

When transmission capacity exceeds the nodal demand ($t \geq d_i$), only differences in production patterns (arising when $\rho_{h,l} < 1$) incentivise the allocation of capacity to node l. For the given numerical example, this is relevant only when the $v > 5d_{h+l}$.

Figure 2.3 also shows the shorting of the *low-availability deployment range*. While for $t = \frac{1}{4}d_i$ the range has a width of $\frac{4}{5}d_l$, the *low-availability deployment range* is narrowed to $\frac{1}{4}d_l$ if $t = \frac{3}{4}d_l$, and disappears if $t \geq d_l$.

Lastly, the figure shows the shift towards the high-availability node in the *split capacity deployment range*. While for $t = \frac{1}{4}d_i$ roughly 50% of marginal capacity is placed to node h when $v = 4d_{h+l}$, the nodal marginal capacity share at node h increases to 70% for $t = \frac{3}{4}d_i$.

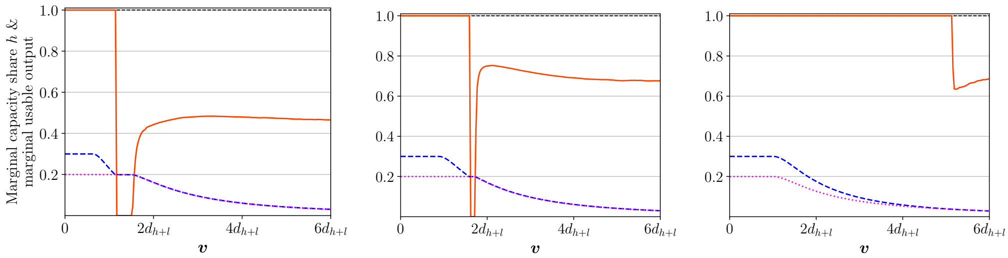


Figure 2.3.: Effect of changes in the transmission capacity (t) on the spatial allocation ranges under nodal pricing.

2.3.3. Effect of changes in the demand distribution

In this subchapter, I derive the effect of the demand distribution on the capacity allocation ranges. Based on the analysis, I conclude:

Finding NP 3. *Under nodal pricing, increasing the demand at the high-availability node d_h widens the high-availability deployment range, while increasing d_l widens the low-availability deployment range. In the split capacity deployment range, the nodal share increases with the nodal demand.*

Explanation. Increasing the demand at node h widens the *high-availability deployment range*. This implies producers are willing to allocate capacity solely to the high-availability node for higher VRE penetration levels. This has two reasons: First, when increasing d_h by one unit, it is possible to add one unit of capacity at node h without inducing curtailment (refers to Case 1 of Finding NP 1). Second, for v valid in Case 2 of Finding NP 1, the marginal curtailment increases at a lower rate with increasing d_h. As a result the cut-off point between the high- and low-availability deployment range ($v^{H|L}$) is reached for higher levels of v.[13]

With increasing d_l the *low-availability deployment range* is widened. This is because the width of the range is defined by $V_l > d_l - t$.

In the *split capacity deployment range*, increases in nodal demand motivate producers to increase the share of capacity they allocate to the node. This is because the more nodal demand and nodal supply are aligned, the lower the need for transmission and the lower the resulting curtailment from limited transmission capacity. Hence, when the demand increases at one node, curtailment can be reduced by shifting capacity to that node. The reduction in curtailment implies increased usable supply from VRE and decreased need for costly conventional power.

End of Explanation.

Figure 2.4 demonstrates the insights from Finding NP 3 numerically. Assumptions regarding the transmission capacity and the availability profiles are identical to Figure 2.2. When demand is mainly allocated to node l (i.e., $d_h = 25$ & $d_l = 75$), the *high-availability deployment range* is 5% smaller than the *low-availability deployment range*. Shifting demand from node l to node h widens the *high-availability deployment range* and narrows the *low-availability deployment range*. When demand is mainly allocated at node h (i.e., $d_h = 75$ & $d_l = 25$), the *high-availability deployment range* is 12 times as long as the *low-availability deployment range*. In the *split capacity deployment range*, the capacity share at node h increases from roughly 25% to 75% when shifting 50% of global demand from node l to node h.

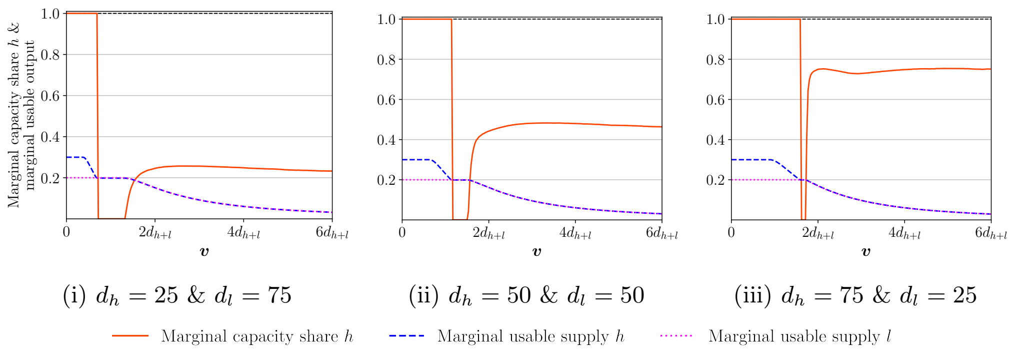


Figure 2.4.: Effect of changes in the demand distribution on the spatial allocation ranges under nodal pricing.

2.3.4. Effect of changes in the availability profiles

In this subchapter, I derive the effects arising from different features of the availability profiles on the capacity allocation ranges under nodal pricing. To do

[13] I show this effect mathematically in Supplementary Material A.1.3.

so, I analyse the effects of changes in the correlation among nodal availability profiles and changes in the average and the variance of nodal availability profiles.

Correlation

In this subchapter, I derive the effect of the correlation among availability profiles on the capacity allocation ranges. Based on the analysis, I conclude:

Finding NP 4. *Under nodal pricing, changing the correlation among availability profiles $\rho_{h,l}$ does not affect the width of the high and low-availability deployment range. In the split capacity deployment range, increasing $\rho_{h,l}$ increases the capacity share allocated to the high-availability node. The effect increases with increasing t.*

Explanation. Under nodal pricing, changing the correlation among availability profiles $\rho_{h,l}$ does not affect the width of the *high-availability deployment range* because $\rho_{h,l}$ does not affect the nodal curtailment at node h when capacity at node l is absent (refers to Case 1 and 2a of Finding NP 1). Changing the correlation does also not affect the width of the *low-availability deployment range*. This is because the width of the range is defined by $d_l - t$ as shown in Finding NP 1 in Case 2b. The correlation between availabilities does not affect the producers' allocation decision as long as VRE penetration is sufficiently low, so curtailment at node l is absent.

In the *split capacity deployment range* when $K^d > 0$ the joint distribution of the availabilities affects the spatial allocation of VRE (refers to Case 3 of Finding NP 1). Increasing correlation shifts capacity towards the *high-availability node*. This is because the incentive for producers to allocate capacity to node l, namely exploiting the differences in availability profiles, is weakened with increasing correlation.

The extent of the effect increases with increasing transmission capacity (t). This is because, for low levels of t, the optimal allocation is mainly driven by the network restrictions. Producers reduce curtailment arising from limited transmission capacity to an appropriate level by allocating capacity relatively even among the nodes (see Finding NP 2). In such a case, the effect of correlation on the allocation is limited. Network restrictions are less relevant for high levels of t, and the availability profiles, including the differences in availability profile patterns, mainly drive the optimal allocation. Hence, the relevance of correlation on the producer's allocation decision in the *split capacity deployment range* increases with increasing t.

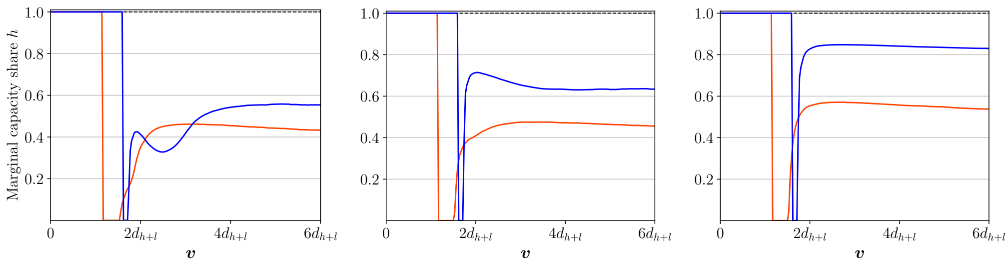
End of Explanation.

Figure 2.5 demonstrates the insights from Finding NP 4 numerically. Assumptions regarding the demand and the availability profiles are identical to Figure 2.2.

The correlation varies in the interval 0 and 1 (i.e., $\rho_{h,l} \in [0, 1]$) for the case of low and high transmission capacity (i.e., $t = \frac{1}{4}d_i$ and $t = \frac{3}{4}d_i$).

Independent of the transmission capacity, the width of the *high* and *low-availability deployment range* are not affected by changes in the correlation.

In the *split capacity deployment range*, increasing $\rho_{h,l}$ increases the capacity share of the high-availability node. Analysing the nodal marginal capacity shares at $v = 6d_{h+l}$ shows that the effect of correlation on node-h-capacity-share increases with increasing transmission capacity. For low levels of transmission capacity (i.e., $t = \frac{1}{4}d_i$), the node-h-capacity-share increases by only 10% from 43% to 53% when $\rho_{h,l}$ is increased from 0 (uncorrelated) to 1 (perfectly correlated). When transmission capacity is high (i.e., $t = \frac{3}{4}d_i$), the node-h-capacity-share increases by almost 30% (i.e., from 55% to 83%) when $\rho_{h,l}$ is increased from 0 to 1.



(i) $\rho_{h,l} = 0$ (uncorrelated) (ii) $\rho_{h,l} = 0.5$ (iii) $\rho_{h,l} = 1$ (perfectly correlated)

Low transmission capacity ($t = \frac{1}{4}d_i$) High transmission capacity ($t = \frac{3}{4}d_i$)

Figure 2.5.: Effect of the correlation among availability profiles on the spatial allocation ranges under nodal pricing.

Average

In this subchapter, I derive the effect of the average availability on the capacity allocation ranges. The density of availabilities is described by the parameters α_i and β_i. As stated in Chapter 2.2, increasing α_i primarily increases the average availability, while the variance remains rather constant. To assess the effect of changes in the average availability, I analyse the effects arising from changes in α_i.[14] Based on the analysis, I conclude:

Finding NP 5. *Under nodal pricing, increasing the average availability μ_l by increasing α_l narrows the high-availability deployment range. The effect of μ_h on the high-availability deployment range and the effect of μ_i on the split capacity deployment range is ambiguous and depends on the system topology.*

Explanation. With increasing average availability at node l the *high-availability deployment range* is narrowed, because the difference in the average availability

[14] Changing μ_i while keeping the variance fully constant, i.e. also altering β_i, does not alter the findings. However, such an approach does not allow for analysing the effects analytically.

between the high- and the low availability node is reduced. As a result, producers only tolerate lower marginal curtailment levels at node h when allocating capacity to the node (refers to Case 2a of Finding NP 1). This implies a narrowing of the *high-availability deployment range.*

Increasing the average availability at node h can either narrow or widen the *high-availability deployment range.* This is due to two opposing effects which occur for v valid in Case 2a of Finding NP 1: On the one hand, an increase in the average availability at node h increases the difference between the average availability at node h and h. This effect allows for higher curtailment at node h and incentives producers to widen the sole capacity allocation to node h. On the other hand, increases in nodal availability increase the marginal curtailment arising from limited transmission capacity. The increased relevance of network restrictions and the resulting increase in marginal curtailment incentive producers to narrow the sole capacity allocation to node h. Subtracting both effects yields the effect on the marginal usable supply. Depending on the parameters t, d_i, and $\beta_i(\alpha_i, \beta_i)$ as well as the nodal VRE capacity, V_i, the effect on marginal usable supply can either be positive or negative.[15] When t is low, the term tends to be negative in the relevant domain. Hence, when network restrictions are tight, the increase in marginal curtailment due to limited transmission capacity outweighs the increase in marginal potential supply. In such a situation, producers reduce the amount of capacity they solely allocate to node h, with increasing μ_h. When t is high, the term tends to be positive in the relevant domain. Hence, with increasing μ_h, producers increase the capacity they solely allocate to node h.

The width of the *low-availability deployment range* is given by $d_l - t$, such that it is not affected by μ_i (refers to Case 2a of Finding NP 1).

The effect of increasing average availability on the *split capacity deployment range* is ambiguous. This is also due to the two opposing effects of increased potential supply and curtailment. While producers tend to increase the nodal with increasing μ_i when t is high, the opposite is true for low levels of t.

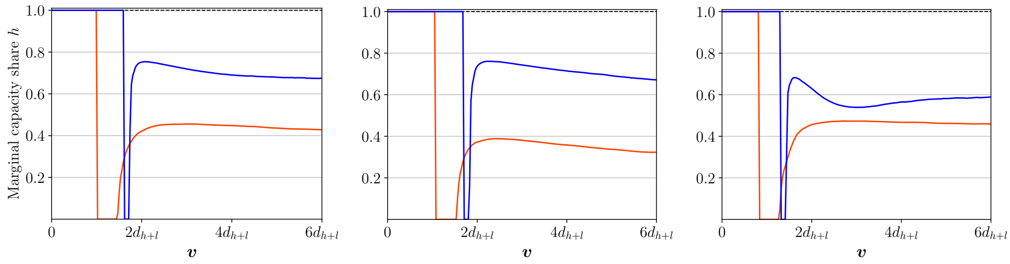
End of Explanation.

Figure 2.6 displays the effects of changes in the average nodal availability on the capacity allocation in a numerical example. Assumptions regarding the demand, q_i, and the correlation are identical to Figure 2.2. To analyse the effect for low and high levels of transmission capacity, the marginal capacity shares are calculated for $t = \frac{1}{10}d_i = 5$ and $t = \frac{3}{4}d_i = 37.5$.

When increasing μ_l from 0.2 to 0.25, the *high-availability deployment range* narrows by roughly 17% for both cases of transmission capacity (compare Figure 2.6a and c). All other effects when changing μ_i highly depend on the level of transmission capacity. In the numerical example, this can be observed best

[15] I mathematically derive the effect of changes in α_h on the marginal usable supply in Supplementary Material A.1.3.

for the *split capacity deployment range*. When transmission capacity is low, a higher μ_h decreases the share of capacity allocated to node h by roughly 10-15 percentage points and increasing μ_l decreases the share of capacity allocated to node l by roughly 10-20 percentage points. When transmission capacity is high, the opposite effects occur. Higher μ_h increases the share of capacity allocated to node h by roughly five percentage points. Higher μ_l decreases the share of capacity allocated to node l by roughly 10-20 percentage points.



(i) $\mu_h = 0.3$ & $\mu_l = 0.2$ (ii) $\mu_h = 0.4$ & $\mu_l = 0.2$ (iii) $\mu_h = 0.3$ & $\mu_l = 0.25$

Low transmission capacity ($t = \frac{1}{10}d_i$) High transmission capacity ($t = \frac{3}{4}d_i$)

Figure 2.6.: Effect of the average in the availability profile on the spatial allocation ranges under nodal pricing.

Variance

In this subchapter, I derive the effect of the variance in availability profiles (σ_i^2) on the capacity allocation ranges. The variance is defined by the availabilities density function, i.e., $B(\alpha_i, \beta_i)$. To analyse the effect of σ_i^2, the parameter values α_i and β_i are changed such that the average supply potential μ_i remains constant. Based on the analysis, I conclude:

Finding NP 6. *Under nodal pricing, higher variance at the high-availability node* σ_h^2 *narrows the high-availability deployment range. The effect of increases in the nodal variance on the split capacity deployment range is twofold: The nodal share decreases for moderate VRE penetration or high transmission capacity, while the nodal share increases when VRE penetration is high, and transmission capacity is low.*

Explanation. Under nodal pricing higher σ_h^2 narrows the *high-availability deployment range* due to higher marginal curtailment at node h. The increase in marginal curtailment results in the fact that Case 2a in Finding NP 1 is valid only for lower level of v. This can be explained as follows: The volatile potential supply decreasingly matches the constant demand, such that the potential supply at node h more often exceeds $d_h + t$ requiring curtailment. As marginal curtailment is higher the usable supply at node h is equal to the potential supply at node l for lower levels of v.

The width of the *low-availability deployment range* is given by $d_l - t$, such that it is not affected by σ_i^2 (refers to Case 2b of Finding NP 1). The effect of increases in variance on the *split capacity deployment range* is twofold: On the one hand, increases in the variance reduce the overall usable supply as the potential supply decreasingly matches the nodal demand. As a result, increasing the variance reduces the marginal usable supply for low and moderate levels of VRE penetration. Producers are incentivised to allocate less VRE to the node with increased variance. On the other hand, increasing the nodal variance can increase the marginal usable supply for high levels of VRE penetration. This is because increases in the nodal variance lower the nodal VRE share. When nodal VRE shares are high (e.g. close to 100%), additional VRE can be barely used to serve the nodal demand. With lower VRE shares, due to increases in the variance, a higher share of the additional VRE can be used to serve the nodal demand and thereby increase the marginal usable supply. Hence, producers are incentivised to allocate more VRE to the node with increased variance when VRE penetration is high. The higher the VRE penetration, the stronger the effect. The effect gets weaker with increasing transmission capacity because VRE supply can be increasingly integrated by exports and high nodal VRE shares get less relevant.

End of Explanation.

Figure 2.7 displays the insights from Finding NP 6 numerically for the case of low transmission capacity (i.e., $t = \frac{1}{10}d_i$).[16] Assumptions regarding the demand, the average availability, and the correlation are identical to Figure 2.2. When the variance is increased at node h (compare Figure 2.7a and b), the *high-availability deployment range* is narrowed from $1.3d_{h+l}$ to $0.9d_{h+l}$. Additionally, the figure confirms that the effect on the *split capacity deployment range* is twofold. In the case of moderate VRE penetration levels (i.e., roughly $v \leq 4d_{h+l}$), increasing σ_h^2 lowers the share of capacity allocated to node h. Such changes occur for VRE shares below 70%, as indicated by the green line. For high VRE penetration levels (i.e., $v \leq 4d_{h+l}$ or a VRE share above 70%), increasing σ_h^2 increases the share of capacity allocated to node h. Comparing the green lines also illustrates the decrease in the global VRE share. When the variance is increased at node l (compare Figure 2.7a and c)), the *high-availability deployment range* is not affected. The effect on the *split capacity deployment range* is the same as in the case of increases in σ_h^2. For moderate VRE penetration levels, higher σ_l^2 lower the share of capacity allocated to node l (i.e., more capacity is allocated to node h). For high VRE penetration levels, higher σ_l^2 increase the share of capacity allocated to node l.

[16] A numerical analysis for the case of high transmission capacity is shown in Chapter A.6.

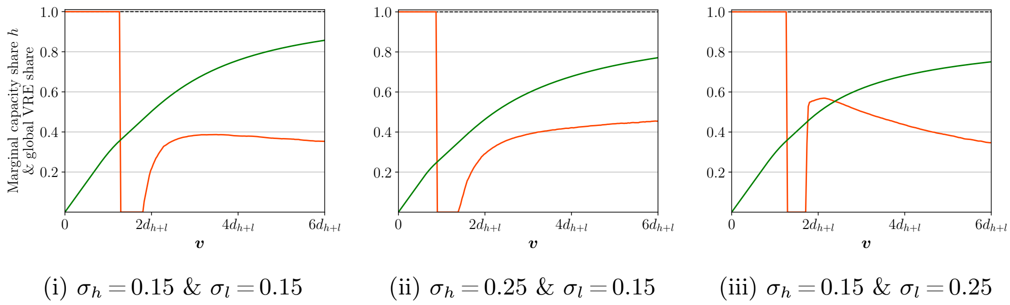


Figure 2.7.: Effect of the variance in the availability profile on the spatial allocation ranges under nodal pricing.

2.4. Spatial allocation under uniform pricing

In this chapter, I analyse the spatial allocation under uniform pricing. The structure is similar to the previous chapter. First, I identify the two ranges of capacity allocation, valid for different relative levels of VRE penetration. Within the analysis, I assess the inefficiency by comparing the results to the optimal spatial allocation I derived in Chapter 2.3. Second, I analyse how various parameters of the system topology, namely the transmission capacity, the demand distribution, and the characteristics of the VRE availability profile, drive the spatial allocation and the resulting inefficiencies. Within the analysis, I assess the interactions among the parameters.

2.4.1. Capacity allocation ranges

In this subchapter, I derive the capacity allocation ranges under uniform pricing for different relative levels of VRE penetration. I conclude that:

Finding UP 1. *Under uniform pricing, the spatial VRE allocation encompasses the high-availability deployment range and the split capacity deployment range. Allocation is efficient when VRE penetration is low. With increasing VRE penetration inefficiencies emerge. These resulting welfare losses increase until marginal capacity is split among nodes.*

Explanation. Uniform pricing implicitly ignores network constraints when deriving market prices. Hence, the producers' profit maximisation problem coincides with minimising the total costs when ignoring network constraints, denoted by DTC. DTC is given by the global supply of the conventional technology before redispatch times the marginal costs of the conventional technology (c). The global supply of the conventional technology before redispatch arises from the global demand minus the global supply of the VRE technology sold to the market. This supply I also denote as saleable supply (SS). The saleable supply is given by the global potential supply minus the supply, which cannot be sold to the market (i.e., $SS = \sum_i PS_i - \overline{K^c}$). The supply which cannot be sold to the market is denoted as commercial curtailment ($\overline{K^c}$). Hence, the objective function for the case of uniform pricing is given by:

$$\min_{V_h} DTC = \Bigg(d_{h+l} - \underbrace{\Big(V_h\mu_h + (v - V_h)\mu_l - \overline{K^c}(V_h)\Big)}_{\text{saleable supply}}\Bigg)c \quad \text{for } V_h \in [0, v] \quad (2.19)$$

The global demand (d_{h+l}) is a constant and hence independent of V_h. Further, the cost parameter c is positive by definition. Hence, the distorted total costs are minimised by maximising the level of saleable supply (SS).

Case 1: No commercial curtailment

In the absence of commercial curtailment the saleable supply coincides with the potential supply. Hence the objective function is minimised by maximising the potential supply. Since $\mu_h > \mu_l$ by definition the total costs are minimised by allocating all capacity to node h.

Supply is commercially curtailed when the global supply exceeds the global demand. As stated in Chapter 2.2, I assume that there is a period r with an availability equal to 1 at both nodes simultaneously (i.e., $avail_{h,r} = avail_{l,r} = 1$). Hence, commercial curtailment is absent if the installed capacity does not exceed the global demand:

$$V_h^{UP*}(v) = v \quad \text{if: } v \leq d_{h+l} \quad (2.20)$$

Case 2: Presence of commercial curtailment

When VRE penetration exceeds the global demand ($v > d_{h+l}$), increasing the capacity at either node induces commercial curtailment ($\overline{K^c}$). This implies that the saleable supply is smaller than the potential supply. The optimal allocation of marginal capacity in this case depends on the level of marginal commercial curtailment. This results in two sub cases:

Case 2a: Allocating capacity to node h

In Case 2a it is optimal to allocate all capacity to node h. This requires the marginal saleable supply to be higher at node h than at node l when increasing the VRE penetration (v) marginally. Hence, a marginal increase in v, evaluated at $V_h = v$, yields in a marginal commercial curtailment which is smaller than the delta in marginal potential supply between the high- and the low-availability node minus the marginal curtailment that would happen when placing the marginal capacity at node l.

For which level of v this is the case cannot be derived analytically, as I cannot derive $\overline{K^c}$ analytically.[17] However, I can explain the general rationale present in Case 2a. Later a numerical example will confirm the explanation:

At $v = d_{h+l}$ commercial curtailment is zero, i.e., there is no period in which $\sum_i PS_{i,r} > d_{h+l}$. When adding marginal capacity at node h such that d_{h+l} is exceeded commercial curtailment occurs. The commercial curtailment induced by the marginal capacity allocated to node h is very low for v, only slightly exceeding d_{h+l}. This is because $\overline{K^c_r}$ is positive only in those periods r with an $avail_{h,r}$ close to 1.

The marginal saleable supply decreases faster at node h than at node l, when the availability profiles are not perfectly correlated $\rho_{h,l} < 1$. This is because VRE production follows the availability profile of node h. When allocating additional capacity to node h the high additional potential supply at node h occurs in periods with commercial curtailment. When allocating additional capacity to node l and nodal availability profile patterns differ, the periods with potential supply at node l less often occur in periods with commercial curtailment. The VRE penetration v, which induces the marginal supply to coincide between the high- and low availability node, marks the cut-off point ($v^{UP/H|S}$). On this occasion, the first-order condition is fulfilled for $V_h = v$:

$$\left.\frac{\partial DTC}{\partial V_h}\right|_{V_h=v} = -\left(\mu_h - \mu_l - \left.\frac{\partial \overline{K^c}}{\partial V_h}\right|_{V_h=v}\right)c = 0 \tag{2.21}$$

Hence, the cut-off point can be expressed as follows:

$$v^{UP/H|S} = v[\left.\frac{\partial DTC}{\partial V_h}\right|_{V_h=v} = 0] \tag{2.22}$$

Case 2b: Splitting capacity among node h and l

For $v > v^{UP/H|S}$ profit maximising producers split the capacity among the high- and low-availability node. This can be explained as follows: At $v = v^{UP/H|S}$ the marginal saleable supply is identical at the high- and low availability node. When adding one unit of capacity, such that $v^{UP/H|S}$ is exceeded, it can either be allocated to node h, to node l or split among the nodes. In line with the explanation given in Case 2a, adding VRE capacity solely to node h would result in marginal saleable supply of node h to drop below the one of node l. Hence, producers will not allocate the additional capacity solely to node h. However, adding the VRE capacity solely to node l would result in marginal saleable supply at node l dropping below the one at node h. This is because also in when allocating capacity to node l commercial curtailment would increase. Hence, profit

[17] The level of $\overline{K^c}$ depends on the overall VRE capacity (v), the capacity allocation (V_i), the density of the nodal availability profiles (f_{AVAIL_i}), and the deterministic joint distribution of the nodal availability profiles. The joint availability distribution depends on the nodal beta distribution parameter α_i, β_i and the correlation $\rho_{h,l}$ among the nodal availability profiles, which cannot be derived analytically.

maximising producers split the capacity among the high- and low-availability node, such that the marginal saleable supply is identical at the high- and low-availability node.

When the splitting of capacity results in a profit maximising allocation the first-order condition is fulfilled:

$$\frac{\partial DTC}{\partial V_h} = -\left(\mu_h - \mu_l - \frac{\partial \overline{K^c}}{\partial V_h}\right)c = 0 \tag{2.23}$$

Summarising the results of Case 2 yields in the following profit maximising spatial allocation under uniform pricing:

$$V_h^{UP*}(v) = \begin{cases} v & \text{if: } \; d_{h+l} < v \leq v^{UP/H|S} \\ V_h[\frac{\partial DTC}{\partial V_h} = 0] & \text{if: } \; v > v^{UP/H|S} \end{cases} \tag{2.24}$$

Defining the capacity allocation ranges

Summarising the results of Case 1 and 2, the profit maximising spatial allocation is given by the following section-wise defined function:

$$V_h^{UP*}(v) = \begin{cases} v & \text{if: } \; v \leq v^{UP/H|S} \\ V_h[\frac{\partial DTC}{\partial V_h} = 0] & \text{if: } \; v > v^{UP/H|S} \end{cases} \tag{2.25}$$

For VRE penetrations below $v \leq v^{UP/H|S}$ capacity is allocated only to the high-availability node. I denote this range as *high-availability deployment range*. For higher VRE penetration levels ($v > v^{UP/H|S}$), producers split the capacity among the two nodes. This range I denote as *split capacity deployment range*.

Inefficiencies resulting from UP

Allocation is efficient when the highest usable supply possible for a given VRE penetration is achieved. Such a usable supply is achieved under nodal pricing, which I denote with $US^{NP}(v)$. For some VRE penetration levels (v) the allocation under uniform pricing differs from the efficient allocation, resulting in a lower usable supply (denoted with US^{UP}). This can be explained as follows: As $t < d_i \forall i$ by assumption (see Chapter 2.2) the actual curtailment at node h exceeds the commercial curtailment. Hence, there is a lower incentive for producers to allocate capacity to node l under uniform pricing compared to nodal pricing. As a result, producers allocate capacity to node h for higher VRE penetration levels under uniform pricing compared to nodal pricing (i.e., $v^{UP/H|S} > v^{H|L}$).

When VRE penetration is sufficiently low (i.e., $v \leq v^{H|L}$) capacity is allocated to node h under uniform and nodal pricing, such that the usable supply is identical under both regimes and inefficiencies are absent. For $v > v^{H|L}$ inefficiencies occur. The resulting welfare loss is given by the reduction in usable supply compared to the optimum times the marginal costs for conventional technology:

$$\text{Welfare Loss}(v) = \big(US^{NP}(v) - US^{UP}(v)\big)c \tag{2.26}$$

As the global VRE share is given by dividing the usable supply with the global demand, the following relationship between the welfare loss and the reduction in the global VRE share exists:

$$\text{Reduction in global VRE share}(v) = \frac{\text{Welfare Loss}(i)}{d_{h+l} \cdot c} \tag{2.27}$$

Welfare loss increase with increasing v as long as producers solely allocate capacity to node h, i.e. for $v^{H|L} < v \leq v^{UP/H|S}$.

When capacity is split among the two nodes (i.e., *split capacity deployment range* which is valid for $v > v^{UP/H|S}$), welfare losses are partly mitigated. This is because capacity allocated to node l is not curtailed, such that the average marginal usable supply of both nodes exceeds the marginal usable supply under nodal pricing.

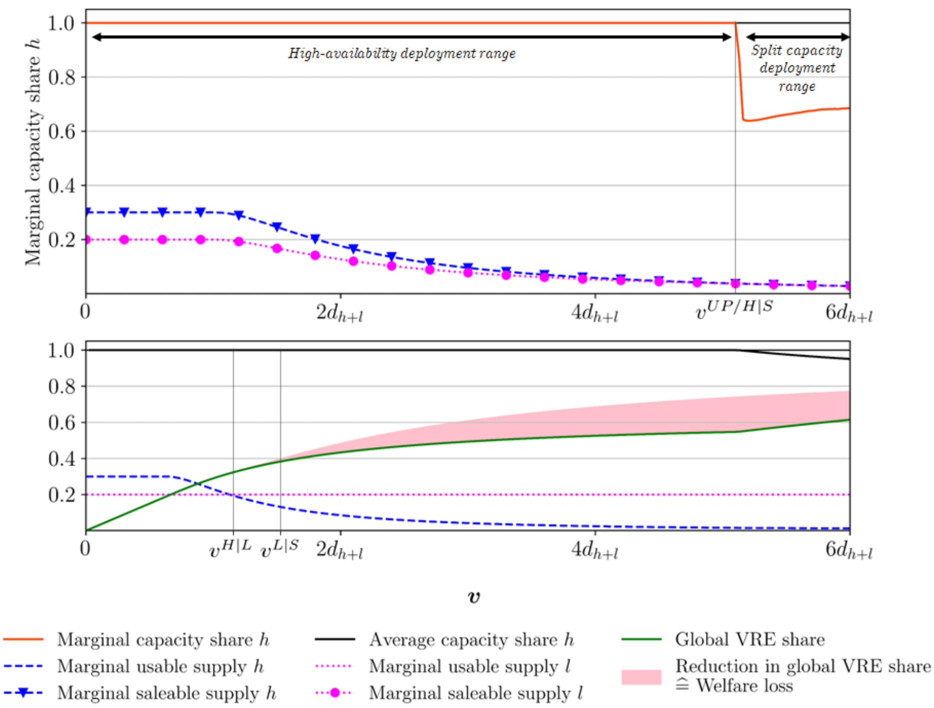
End of Explanation.

Figure 2.8 demonstrates the insights from Finding UP 1 numerically.[18] The parameters are identical to Figure 2.2. By setting $c = \frac{1}{d_{h+l}}$, the welfare loss coincides with the reduction in the global VRE share.

The upper diagram shows that capacity is allocated solely to the high-availability node when $v < 5.1d_{h+l}$. Welfare losses arise for $v > v^{H|L}$, where $v^{H|L}$ defines the value separating the *high-* and *low-availability deployment range* under nodal pricing. For $v > v^{H|L}$ the marginal usable supply of node h subceeds the one of node l, such that welfare would be increased when some VRE would be shifted to node l. The welfare losses grow with increasing VRE penetration. At $v = 5.1d_{h+l}$, welfare losses reach their maximum. Due to the inefficient allocation, less than 55% of demand can be served with VRE, compared to 75% under an optimal allocation. Hence, conventional power needs to serve an additional 20% of demand, inducing costs of $0.2d_{h+l}c = 0.2$. This is the case even though the potential supply of VRE is 20% higher under uniform pricing than under nodal pricing. These numbers imply 65% of VRE is curtailed on average and 95% of marginal supply is curtailed.

When $v > 5.1d_{h+l}$ capacity is split among the nodes h and l in a roughly 70/30 ratio. As no supply is curtailed at node l, welfare losses compared to the social optimum slightly decline.

[18] The optimisation problem solved numerically to generate this and the remaining figures in this chapter resembles the optimisation problem described in Chapter 2.2 and solved in this chapter. Differences in the formulation arising from using a numerical instead of an analytical approach are described in Supplementary Material A.4.



$v^{H|L}$ and $v^{L|S}$ mark the cut-off points between respective ranges arising under nodal pricing.

Figure 2.8.: Spatial allocation, marginal usable and saleable supply, VRE shares and welfare losses at different VRE penetration levels under uniform pricing.

2.4.2. Effects of changes in the transmission capacity and demand distribution

In this subchapter, I derive the effect of changes in the transmission capacity t and the demand distribution on the capacity allocation and welfare. Based on the analysis, I conclude:

Finding UP 2. *Under uniform pricing, the transmission capacity and the demand distribution do not affect the capacity allocation. With increasing t, welfare losses decrease. Allocation is efficient when $t > d_i$. Distributing demand more according to potential supply also reduces welfare losses.*

Explanation. The transmission capacity and the demand distribution do not affect the capacity allocation. This is because curtailment arising from network restrictions is ignored under uniform pricing.

As curtailment arising from network restrictions affect the socially optimum capacity allocation, the transmission capacity and the demand distribution affect the efficiency. The deployment under uniform pricing coincides with the one under nodal pricing, when transmission capacity allows to serve the demand at node l, even when capacity is allocated solely at node h, i.e., $t > d_i$. Hence, welfare losses are absent when $t \geq d_i$. For such level of t the commercial curtailment $(\overline{K^c})$ and the physical curtailment (K) coincide. With decreasing t, welfare losses increase. This is because with decreasing t, it is optimal to allocate more capacity to node l to reduce curtailment.

Distributing nodal demand more according to potential supply reduces the need for transmission. Hence, distributing demand more according to potential supply also reduces the level of curtailment arising from network restrictions and welfare losses.

End of Explanation.

Figure 2.9 displays the effect of changes in the transmission capacity and the demand distribution.[19]

Independent of the transmission capacity and the demand distribution, capacity is allocated to node h until $v = 5.1d_{h+l}$. For the same VRE penetration level, welfare losses are highest. The welfare loss decreases with increasing t. For low transmission capacity, i.e., $t = \frac{1}{4}d_{h+l}$, welfare losses equivalent to the variable costs, when serving 20% of global demand with conventional power, occur. The welfare loss is more than halved when transmission capacity is tripped, and welfare losses disappear for $t > d_i$. Furthermore, the higher t, the later welfare losses emerge.

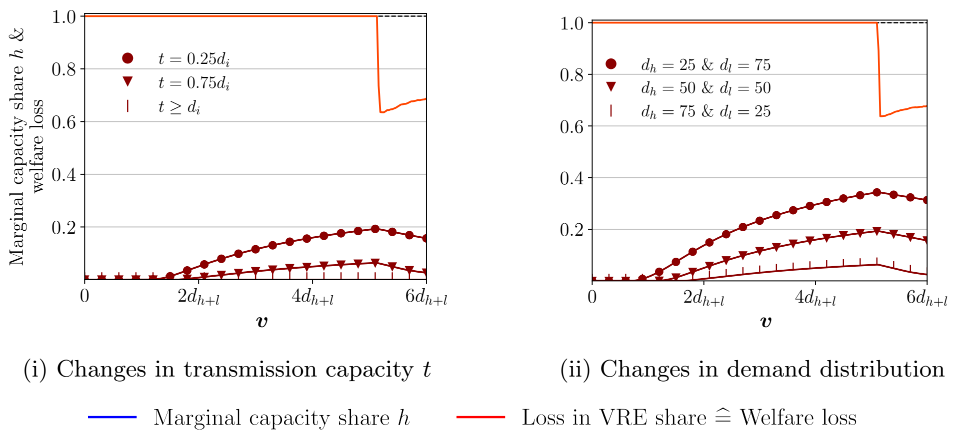


Figure 2.9.: Effects of changes in the transmission capacity and demand distribution on the spatial allocation ranges and welfare losses under uniform pricing.

With demand increasingly located at node h, the welfare losses decreases. When 75% of demand is located at node h, welfare losses are only one-fifth compared to the case when 75% of demand is located at node l.

[19]The parameters in Figure 2.9a are identical to Figure 2.3, while the parameters in Figure 2.9b are identical to Figure 2.4. By setting $c = \frac{1}{d_{h+l}}$, the welfare loss coincides with the reduction in VRE share.

2.4.3. Effect of changes in the availability profiles

In this chapter, I derive the effects arising from different features of the availability profiles on the ranges of the capacity allocation under uniform pricing. To do so, I analyse the effects of changes in the correlation among nodal availability profiles and changes in the average and the variance of nodal availability profiles. Based on the analysis, I conclude:

Finding UP 3. *Under uniform pricing, the availability profiles affect the spatial allocation. The effect on the high-availability deployment range and the split capacity deployment range is identical to the case of nodal pricing and* $t \geq d_i$*. For* $t < d_i$*, the effect on the spatial allocation are stronger under uniform pricing. Changes in the availability profiles, which incentivise capacity to be allocated more according to demand, reduce welfare losses.*

Explanation. The availability profiles affect the spatial allocation because the global supply is affected. The effect on the spatial allocation of VRE is stronger than under nodal pricing if transmission capacity is binding. This is because, under nodal pricing the effects on the capacity allocation arising from availability profiles are mitigated by curtailment arising from limited transmission capacity. The higher t, the lower the mitigation and the higher the impact of availability profiles. For $t \geq d_i$, the availability profiles affect the allocation in the same way under uniform and nodal pricing.

Changes in the availability profiles also affect the level of inefficiency. If changes in the availability profiles incentivise a capacity allocation which induces potential supply to be allocated more according to demand, welfare losses are reduced. This is because, with the increasing alignment of potential nodal supply and nodal demand, less potential supply needs to be transmitted, and less supply is curtailed due to limited transmission capacity. Thereby the redispatch-level of conventional power plants decreases, reducing costs and increasing welfare and the global VRE share. Depending on the demand distribution, the availability profiles, which minimise the welfare loss differ.

End of Explanation.

Correlation

When availabilities are perfectly correlated, producers allocate capacity solely to the high-availability node. The lower the correlation, the more often the availability at node l exceeds the availability at node h. Producers exploit these differences in the availability profile by allocating some capacity to node l. Hence, with decreasing correlation, producers allocate more capacity to node l. The effect on welfare depends on the demand distribution. When demand $d_l \geq d_h$, decreasing the correlation decreases the need for transmission and curtailment arising from limited transmission capacity. Hence, low $\rho_{h,l}$ lead to low welfare losses compared to the social optimum. However, when demand is located mainly at node h, the need for transmission is lowest, and resulting curtailment is lowest when the correlation is high. Hence, high $\rho_{h,l}$ lead to low welfare losses compared to the social optimum.

Figure 2.10 illustrates these results.[20] Capacity is solely allocated to node h for all levels of analysed VRE penetration when availability profiles are perfectly correlated. In contrast, when availability profiles are uncorrelated, 38% to 100% of marginal VRE capacity is allocated to node l for $v > 2.5d_{h,l}$. Which level of correlation yields the lowest welfare loss depends on the demand distribution. When demand is equally distributed, welfare losses are lowest when availability profiles are uncorrelated. In contrast, when 95% of demand is concentrated at node h, welfare loss remains absent in the analysed VRE penetration domain when the correlation is perfect.

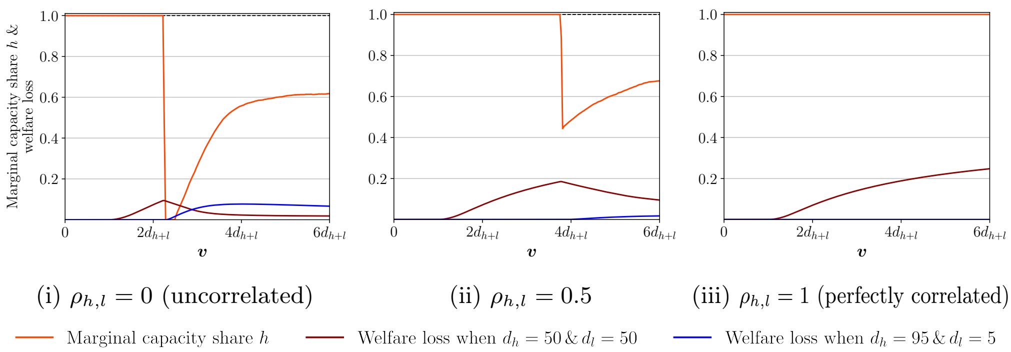


Figure 2.10.: Effect of the correlation among availability profiles on the spatial allocation ranges and welfare losses under uniform pricing.

Average

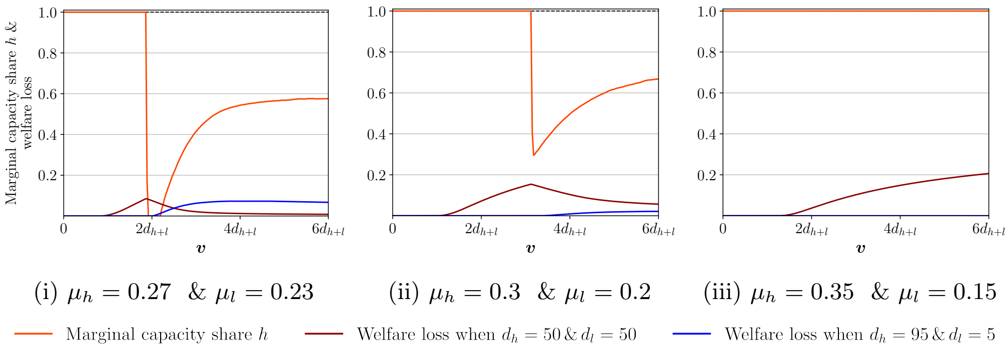
Under uniform pricing, a higher average nodal availability incentivises producers to allocate more VRE capacity to the respective node. This is because producers only consider the increase in the potential nodal supply. The increasing curtailment arising from limited transmission capacity are ignored. As $\mu_h > \mu_l$, producers allocate more capacity to node h. With an increasing difference in availabilities the share of capacity allocated to node h increases.

Hence, when demand $d_l \geq d_h$, decreasing the differences in availability decreases the need for transmission and hence decreases the welfare loss. However, when demand is concentrated at node h, the need for transmission and the resulting welfare losses from curtailment is lowest when differences in availability are substantial.

[20]The average and the variance of the availability profiles are identical to Figure 2.8. To better identify the effect on welfare losses, $t = 5$ is assumed. By setting $c = \frac{1}{d_{h+l}}$, the welfare loss coincides with the reduction in the VRE share.

Figure 2.11 illustrates these results.[21] Capacity is solely allocated to node h for all levels of analysed VRE penetration when the average availability at node h is 2.3 times as high than at node l (i.e., $\mu_h = 0.35$ and $\mu_l = 0.15$). In contrast, 40-100% of marginal VRE capacity is allocated to node l for $v > 1.9d_{h,l}$, when the average availability at node h is only 17% higher than at node l (i.e., $\mu_h = 0.27$ and $\mu_l = 0.23$).

Which nodal average availabilities yield the lowest welfare loss depends on the demand distribution. When demand is equally distributed, welfare losses are lowest when average availabilities barely differ. In contrast, when 95% of demand is concentrated at node h, welfare losses remain absent in the analysed VRE penetration domain when availabilities are 2.3 times higher at node h than at node l.



(i) $\mu_h = 0.27$ & $\mu_l = 0.23$ (ii) $\mu_h = 0.3$ & $\mu_l = 0.2$ (iii) $\mu_h = 0.35$ & $\mu_l = 0.15$

Figure 2.11.: Effect of the average availability on the spatial allocation ranges and welfare losses under uniform pricing.

Variance

Under uniform pricing, a higher nodal variance incentivises producers to allocate less VRE capacity to the respective node. This is because, with increasing nodal variance, the potential nodal supply exceeds the global demand is higher share of periods. In such periods, prices are zero, and some potential supply cannot be sold. To reduce the share of such situations, producers allocate less VRE to the node with increased variance.

Hence, when demand $d_l \geq d_h$, increasing the variance at node h or decreasing the variance at node l decreases the need for transmission and curtailment arising from limited transmission capacity. Such changes in the variance also decrease the welfare loss compared to the social optimum. When demand is concentrated at node h, the need for transmission and the resulting welfare losses from curtailment is lowest when the variance at node h is low compared to the variance at node l.

[21] The variance of the availability profiles varies between X and Y, which is very close to the variance assumed in Figure 10. To better identify the effect on welfare losses, a low transmission capacity of $t = \frac{d_{h+l}}{10}$ and a moderate correlation of $\rho_{h,l} = 0.4$ is assumed. By setting $c = \frac{1}{d_{h+l}}$, the welfare loss coincides with the reduction in VRE share.

These results are illustrated in Figure 2.11.[22] Capacity is solely allocated to node h for all levels of analysed VRE penetration when availability profiles at both nodes share the same variance (i.e., $\sigma_h = \sigma_l = 0.2$). In contrast, 65-100% of marginal VRE capacity is allocated to node l for $v > 3d_{h,l}$, when the variance at node h is 50% higher than at node l (i.e., $\sigma_h = 0.24$ and $\sigma_l = 0.16$).

Which nodal variances yield the lowest welfare loss depends on the demand distribution. When demand is equally distributed, welfare losses are lowest when the variance is 50% higher at node h. In contrast, when 95% of demand is concentrated at node h, welfare losses remain absent in the analysed VRE penetration domain when the variance is identical at both nodes.

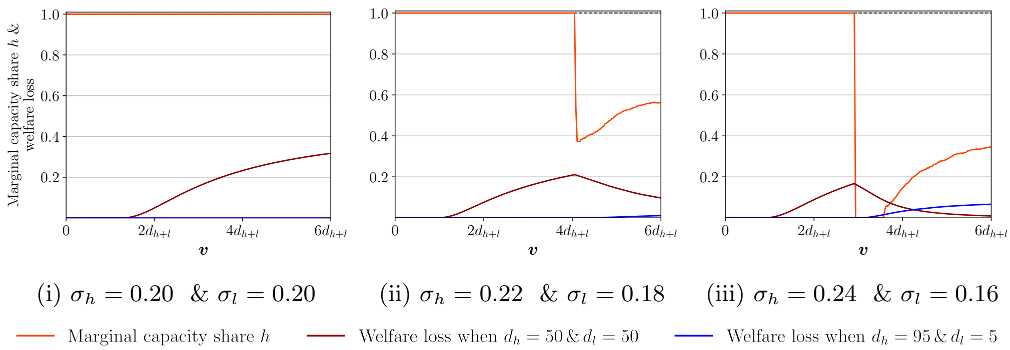


Figure 2.12.: Effect of the average availability on the spatial allocation ranges and welfare losses under uniform pricing.

2.5. Discussion

This paper shows that optimal VRE allocation can be grouped into three ranges. At low levels of VRE penetration, capacity should be allocated to the node with higher average availability (i.e. *high-availability deployment*). When curtailment fully removes the advantage in usable supply of the high-availability node, marginal capacity should be allocated to the node with lower availability (i.e., *low-availability deployment range*). When curtailment is present at both nodes, marginal capacity should be split (i.e., *split capacity deployment range*). Policymakers designing instruments to expand the VRE capacity should consider the range they are in. Countries starting to deploy in VRE should incentivise the placement of initial capacities in regions with high availability. Countries that already have significant VRE capacity in regions with a high average availability may be better off when incentivising (some) investment in regions with lower availability.

The width of the *high* and the *low-availability deployment range*, as well as the nodal capacities shares in the *split capacity deployment range*, are found to

[22] The assumed average in the availability profiles is identical to Figure 2.7, and the assumed transmission capacity, correlation, as well as variable costs of the conventional technology, are identical to Figure 2.11.

depend on the transmission capacity, the demand distribution, and the availability profiles. These characteristics vary among countries. In the UK, compared to Germany, the average wind availability is higher, the regional difference is lower, and the correlation among the availability profiles is lower (Sinden, 2007, Staffell and Pfenninger, 2016). As a result, the *high* and *low-availability deployment range* are narrower in the UK than in Germany due to lower differences in availability and higher average availabilities when assuming similar transmission capacity and demand distribution.

Under uniform pricing, the dominant market design, producers are found to allocate capacity to the high-availability node for higher VRE penetration levels than socially optimal. This is because curtailment arising from limited transmission capacity, which would encourage producers to allocate capacity more according to demand, is ignored. Welfare losses occur when curtailment from limited transmission capacity fully diminish the advantage in usable supply of the high-availability node. The welfare losses increase until differences in availability profiles incentivise allocating some capacity to the low-availability node.

Hence, countries with uniform pricing that start to deploy VRE or feature low VRE shares do not have to implement additional measures to improve the spatial allocation. In line with the findings of this paper, in Japan, VRE only serves 6% of demand, and support schemes do not differentiate spatially (IEA, 2022b). Countries with uniform pricing, which already deploy substantial VRE capacity, such as Germany, should consider measures encouraging producers to invest in areas with lower availabilities.

The welfare losses under uniform pricing decreases in the level of transmission capacity and increases in the need for transmission. The latter is found to be influenced by the demand distribution and the availability profiles. Welfare losses are, for instance, small when transmission capacity is high compared to nodal demand or demand is allocated mainly to the high-availability node. In contrast, welfare losses are found to be high when transmission capacity is low, demand is concentrated in the low-availability node and availability profiles incentive an allocation to the high-availability node (e.g., high difference in nodal availabilities). Such circumstances are, for example, present in Germany. This is in line with the finding from ACER (2022), who show that splitting Germany into two market zones would yield larger welfare increases than splitting market zones in other EU countries.[23] Hence, policymakers should take into account the given transmission capacities, the demand distribution, and regional availability profiles when considering to split their market zone or to implement spatially differentiated VRE subsidies.

[23]Splitting a country into two market zones allows prices to differ when transmission capacities between the new market zones are congested. Such a market design is an intermediate design of uniform pricing and nodal pricing.

The findings of my analysis confirm and extent the findings of the papers presented in Chapter 2.1. The results of Kies et al. (2016) suggest that with an increasing VRE penetration, it is optimal to increasingly allocate capacity to regions with a low average availability. My findings extend the result by showing that the optimal spatial allocation of VRE can be grouped into three ranges.

Pechan (2017) finds that under nodal pricing, producers increasingly concentrate capacity at high-availability nodes when the correlation increases and when the variance in high-availability nodes is low. This is in line with my analysis. My research adds the finding that the effect of correlation becomes more relevant with increasing transmission capacity.

In the case of uniform pricing, Pechan (2017) does finds no effect of correlation and variance on the allocation under uniform pricing. This is because she only analyses a case with a moderate VRE share. I can show that the capacity allocation is affected once the VRE penetration reaches a certain threshold.

In line with my analysis, a welfare loss arises in Schmidt and Zinke (2020) due to an inefficient allocation of VRE. The identified welfare loss of 1.5% seems low compared to numerical results in my analysis. This is because, in my analysis, a case with similar correlation, similar demand distribution and similar VRE penetration leads to a welfare loss of roughly 15%. [24] The difference in welfare loss mainly arises due to the following two aspects: First, Schmidt and Zinke (2020) only assess the allocation of wind onshore capacities added in the years 2020 to 2030. These capacities produce less than 20% of the VRE supply. The remaining 80% come from onshore wind built before the year 2020, offshore wind and solar power. These capacities are distributed identically under uniform and nodal pricing in their analysis. Second, Schmidt and Zinke (2020) consider regional VRE potentials, which limit the capacity allocation to regions with a high average availability. This limitation increases the capacity allocation to nodes with lower average availability under uniform pricing compared to my analysis. Hence, if the authors would ignore limited potentials and allocate all VRE capacities, welfare losses would likely be substantially higher.

The model's simplicity allows to fundamentally understand the effect of crucial system topology parameters on the spatial allocation of VRE, and the inefficiencies arising under uniform pricing. Despite the model's simplicity, I consider main elements which influence the spatial allocation. While the rationales I identify should remain, additional effects may occur when considering a more realistic setting. In the following paragraphs I discuss central simplifications and potential impacts.

I model the spatial allocation decision as a one-shot game in which producers can observe a fixed system topology. Based on this topology, producers allocate

[24] In Figure 10b the case with a 25% demand allocation to high-availability node and a VRE penetration of $2d_{h+l}$ roughly depicts the setting in Schmidt and Zinke (2020).

capacity between the two nodes. In reality, the parameters of the system topology, such as VRE penetration, transmission capacity and demand distribution, change continuously over time in a dynamic process. If only the VRE capacity increases continuously over time, the three ranges identified can be translated into three phases. Namely, initial VRE capacity is allocated at the high-availability node, then capacity is allocated at the low-availability node, and when a high VRE penetration level is reached, capacity is split between the two nodes. In the more likely case of multiple parameters evolving over time, the optimal spatial allocation becomes much more complex. A still simple example could be a continuous increase in VRE capacity and a discrete transmission capacity at one point in time. In such a case, generators allocating VRE capacity need to consider the proportion of the lifetime of the plant before and after the increase in transmission capacity. A possible outcome could be that in a period with moderate VRE penetration, which is well before the transmission capacity increase, it is optimal to allocate a high proportion of capacity to the low-availability node. As the date of the transmission capacity increase approaches, it would be optimal to increasingly allocate VRE capacity to the high-availability node. In reality, therefore, the three ranges identified are unlikely to translate into three phases of capacity expansion. Nevertheless, the results improve the general understanding of the impact of changes in the system topology on the spatial allocation of VRE.

Second, to ensure an analytical solution and to gain a profound theoretical understanding I do not analyse the effect of storage and demand flexibility in the model. However, storage and demand flexibility represents important elements of the system topology and influence the spatial allocation of VRE. This is because storage and demand flexibility provide means to better align VRE supply with demand, by shifting the time of supply provision or shifting the time of demand. The IEA (2022c) assumes storage and demand flexibility to provide a quarter of the required flexibility each in the year 2050 in the Announced Policy Scenario. In my model, storage operators would maximise profits by injecting during periods of high VRE supply (i.e., VRE technology sets the price) and withdrawing during periods of low VRE supply (i.e., conventional technology sets the price). Similarly, operators of demand flexibility would maximise profits form flexibility when shifting demand from periods with high VRE supply to periods with low VRE supply. This implies, with increasing storage and demand flexibility the sum of supply from VRE and storage minus flexible demand becomes less volatile. This is similar to a decrease in the variance of the availability profile. An increase in flexible capacity should therefore have similar effects like a decrease in the variance, which I analyse in Chapter 2.3 and 2.4. Czock et al. (2022), who analyse the optimal storage allocation find storage to be predominantly built at transmission bottlenecks, such that curtailment before bottlenecks decreases. This corresponds to building storage capacity mainly at the high-availability node in my model. Considering such a spatial allocation of storage would increase in the VRE capacity at the high-availability node for most VRE penetration levels compared to my analysis.

Furthermore, I only consider a two-node network. Considering a complex network with multiple nodes yield supply to be transported via multiple nodes. These nodes' remaining available transmission capacity is reduced in such a case. Hence, in case of multiple nodes, not only does the transmission capacity of the producing or importing node affect the spatial allocation of VRE, but also the transmission capacity of all nodes in between.

Furthermore, the analysis only considers one VRE technology and one conventional technology. In most countries, at least two VRE technologies, namely wind and solar, are employed. The coexistence of the VRE technologies likely induces additional interaction effects. For instance, when demand is regionally equally distributed, solar conditions are similar across a country, and wind capacities are located mostly in the north, it would be optimal to allocate more solar capacity to the south than to the north. In contrast, it would be optimal to allocate solar capacity would be predominantly to the north, when most wind capacities are located in the south. Thereby the effects also depend on the penetration level of each VRE technology and the correlation of availabilities among the different VRE technologies. Similar dependencies likely arise from multiple conventional technologies with differing variable costs.

Another simplification is the assumption of constant and inelastic demand. In reality, demand is neither constant nor inelastic. Instead, demand is fluctuating and slightly positively correlated with VRE availability. This is because demand tends to be higher during the day than at night, which is also the case for solar availability. Demand also tends to be higher in winter than in summer, which is also the case for wind availability. Furthermore, household and industrial electricity demand features some level of price elasticity (Cialani and Mortazavi, 2018). Taking such demand characteristics into account is likely to affect the results under nodal pricing in the following way: The *high-availability deployment range* is likely to be valid also for higher VRE penetration levels. First, because curtailment at node h would be lower due to the positive correlation between demand and VRE availabilities. And second, because the remaining curtailment would be partly offset by an increase in elastic demand due to the lower average price level at node h. The *low-availability deployment range* is likely to be narrowed. This is because the range cut-off point is reached when curtailment occurs at node l, and with non-constant demand, low demand may coincide with high VRE availabilities, triggering curtailment. Under *split capacity deployment range*, the share of capacity allocated to the high-availability node is likely to increase if elastic demand is considered. This is because average prices at node h are on average lower than at node l, so the share of demand at node h increases. The increase in demand translates into an increase in electricity prices at node h, which then increases the willingness of investors to allocate capacity to node h.

Considering demand elasticity would also yield in demand being lower under uniform compared to nodal pricing for higher VRE penetration levels. This is because consumers on average bear higher electricity prices under uniform pricing due to the less efficient spatial allocation of VRE and the assumption that the costs for redispatch are borne by the electricity consumers. In addition, due to the lack of a regional price signal under uniform pricing, the regional distribution of demand would be less consistent with the VRE production compared to the case under nodal pricing.

While the two most common market designs are uniform and nodal pricing, the exact regulation usually differs from the two cases I analyse. A prominent example is the compensation of VRE capacity in case of redispatch under uniform pricing. I assume, like Schmidt and Zinke (2020) and Pechan (2017), curtailed producers of VRE are compensated with the market price. In some countries, like Spain, the compensation of VRE capacity in case of redispatch is below market prices, such that producers consider the curtailment, when deciding on the spatial allocation (Bird et al., 2016). The lower the compensation, the closer the capacity allocation to the one arising under nodal pricing. However, studies that analyse the effects of reduced compensations on spatial allocation are lacking. Such studies would get increasingly relevant as countries, such as the UK, consider reducing their compensations (Cholteeva, 2020).

2.6. Conclusion

To date, there is a lack of theoretical literature that provides a comprehensive understanding of the implications of VRE allocation. This paper contributes to this research gap by developing a theoretical model that depicts the spatial allocation of VRE in a two-node network. Using the model, I analyse under which conditions it is welfare enhancing to allocate some VRE capacity to locations with unfavourable potential supply. Furthermore, I assess how the spatial allocation under uniform pricing differs from the optimum and derive the resulting welfare effects.

From a theoretical perspective, my contribution is threefold: First, I show analytically that the optimal spatial allocation can be grouped into three spatial allocation ranges. Second, I show how the width of each range and the allocation when capacity is split is determined by the different parameters of the system topology. And third, I identify the allocation under uniform pricing and the resulting welfare loss, and show how the welfare loss is affected by the different parameters of the system topology. In addition, my study can assist policymakers when designing policies that affect the spatial allocation, or investors trying to identify the profit-maximising allocation of VRE investments.

I develop a stylised model which provides a fundamental understanding of the dynamics and interactions in the allocation of VRE. However, additional effects are likely to occur when considering a setting with a realistic network, multiple VRE and conventional technologies, as well as storage and demand elasticity. The same holds true when considering endogenous investments not only in VRE but also in additional technologies, such as transmission capacity. Taking into account real-world constraints, such as limited regional VRE potentials, is likely to reduce

the inefficiencies observed under uniform pricing.
Further research could extend the model to include additional technologies. For example, investigating a second VRE technology would allow understanding the interdependencies of expanding wind and solar capacity at the same time. The inclusion of elastic demand and storage would allow the analysis of how flexibility affects the spatial distribution of VRE. It could also identify combinations of VRE and storage capacity sufficient to meet all demand with VRE. The implementation of endogenous investments in transmission capacity would make it possible to identify the effects of such investments on the spatial allocation of VRE. This would provide insights into the trade-off between a network-friendly allocation of VRE and the expansion of transmission capacity. Finally, the implementation of a more realistic network would allow to study the impact of curtailment occurring at nodes between production and consumption nodes on the spatial allocation.

3. Complementing carbon prices with Carbon Contracts for Difference in the presence of risk - When is it beneficial and when not?

3.1. Introduction

The decarbonization of the industrial sector requires substantial investments throughout the next decade (IEA, 2021). These investments are typically irreversible decisions that firms have to take in the presence of risk. The risk of an investment's profitability in a decarbonizing world mainly stems from two sources:

First, the profitability of investments in low-carbon or emission-free technologies depends on carbon prices. These technologies are only competitive with conventional technologies if the carbon price throughout the asset's economic life reaches a certain level. However, carbon prices may feature risk. One reason is that the expected carbon damage may change as new scientific evidence on climate change emerges.[25] Another reason is the potentially changing public valuation of carbon damage, shown by court rulings on climate policy in 2021 in Germany (Bundesverfassungsgericht, 2021, Economist, 2021). Both circumstances create a *damage risk.* Firms facing irreversible investments are exposed to such a damage risk as the regulator may adjust the carbon price according to these changes. In fact, Chiappinelli et al. (2021) report that four out of five firms state that the lack of effective and predictable carbon pricing mechanisms is a major barrier to low-carbon investments. López Rodríguez et al. (2017) or Dorsey (2019) provide further empirical analysis that firms reduce their investments due to environmental regulation-related risks.

Second, there is a *variable cost risk.* Variable costs of low-carbon technologies are not fully known, as adopting innovative production processes may involve novel input factors. The markets for some of these input factors are highly immature, the most prominent example being green hydrogen. The production costs of hydrogen might vary depending, e.g., on the costs of electricity or transport (Brändle et al., 2021). Additionally, there is an active and ongoing market ramp-up involving multiple stakeholders to facilitate technological learning (Schlund et al., 2021). Hence, the market for hydrogen is still at the beginning of organising itself (Agency, 2019).

[25] For instance, the Sixth Assessment Report of the Intergovernmental Panel on Climate Change concludes that the climate system is warming faster than previously estimated (IPCC, 2021). Furthermore, OECD (2021) highlight the risks to predict the environmental damage due to the complex climate dynamics.

Firms' possibilities to hedge against these risks are limited or prohibitively costly.[26] For instance, in the European Emission Trading System (EU ETS), the availability of futures contracts with a maturity longer than three years is low (Newbery et al., 2019).[27] Similarly, there are limited hedging possibilities against variable cost risk from novel input factors traded on immature markets (OEIS, 2021). The described risks and the missing hedging possibilities deter firms from investing, which, in turn, poses a challenge to decarbonization.

To nevertheless facilitate and incentivize large-scale investments in the presence of such risks, the European Commission's Hydrogen Strategy and the reform proposal for a *Fit for 55* package, suggest Carbon Contracts for Differences (CCfDs) as a support scheme for firms in the industry sector (European Commission, 2021a). CCfDs are contracts between the government and a firm that pay out the difference between a guaranteed price, the so-called *strike price*, and the actual carbon price, per tonne of emission reduction delivered by the firm through a low-carbon project. The contracts can be interpreted as a short position in a forward on emission permits. Therefore, CCfDs are effectively a hedging instrument to reduce the firms' risk when making investment decisions. Besides their hedging properties, CCfDs may contain a subsidy for decarbonization investments.[28] Such subsidies may be justified by, e.g., positive externalities. In this paper, we do not consider such externalities, and, hence, CCfDs mainly serve as hedging instrument in our setup. So far, there is only a limited understanding of how regulators should design such instruments and under which circumstances the introduction of CCfDs is welfare-enhancing.

In this paper, we analyze how different sources of risk affect the efficiency of CCfDs and when these contracts are preferable to other policies, like committing to a carbon price early on or a flexible carbon pricing regime. We develop an analytical model in which a regulator sequentially interacts with a continuum of risk-averse firms. These firms can either supply the market with a conventional technology, which causes carbon emissions subject to carbon pricing, or invest in an emission-free technology. The valuation of environmental damage from carbon emissions and the variable costs of the emission-free technology may be subject to risk. The firms are heterogeneous regarding their investment costs when adopting the emission-free technology. Firms invest if they increase their expected utility by adopting the emission-free technology. The regulator maximises social welfare by choosing one out of three carbon pricing regimes: 1) setting a carbon price flexibly after the actual damage or costs are revealed (*Regulatory Flexibility*),

[26]If markets were complete, a perfect hedge of all relevant factors determining an investment's profitability would always be possible (Arrow and Debreu, 1954). Thereby, the profitability of abatement investments would not be volatile, and investments would be made as long as they are profitable in expectation without the impact of risk.

[27]There are several reasons why forward markets for emission allowances are incomplete (e.g. Tietjen et al., 2020, for a survey).

[28]This is the case for the German and EU Hydrogen Strategy, as well as 'Fit for 55' package.

2) committing to a carbon price early (*Commitment*)[29], and 3) a hybrid policy regime containing a CCfD and flexible carbon pricing (*CCfD*). We compare these three carbon pricing regimes against the social optimum.

We find that under perfect foresight, i.e. in the absence of risk, all carbon pricing regimes result in the social optimum. In all regimes, the carbon price equals the marginal environmental damage of production. The marginal firm investing in the emission-free technology balances the marginal costs and the marginal benefit of abatement. This finding arises from two effects: First, because the regulator has perfect foresight, she can set the optimal carbon price level at any time. Second, firms do not face a risk in profits. Any risk would hamper firms' willingness to invest if they are risk averse.

We then assess the effect of risk and risk aversion on the performance of the three carbon pricing regimes. In a first setup, we assume that production of the emission-free technology is always socially optimal given the actual damage and variable costs. In these cases, offering a CCfD results in the social optimum irrespective of the source of risk. The regulator can incentivize socially optimal investments via the CCfD and adjust the carbon price according to the actual damage valuation. In contrast, both *Regulatory Flexibility* and *Commitment* fall short of reaching the social optimum. Which of the two regimes is welfare-superior depends on the source of risk. In case of damage risk, the welfare ranking is ambiguous and depends on the level of the firms' risk aversion (with high risk aversion favouring *Commitment*) and the elasticity of demand (with high elasticity favouring *Regulatory Flexibility*). In contrast, committing to a carbon price is welfare-superior to *Regulatory Flexibility* in settings with variable cost risk, as the regulator can incentivize additional investments under *Commitment*.

Lastly, we assess the effects of emission-free production that is potentially welfare reducing given the actual damage and variable costs. In this case, we find that offering a CCfD does not reach the social optimum. If the regulator offers a CCfD, the firms' production decision does not depend on the actual carbon price. Thereby, the regulator safeguards emission-free production even if it is socially not optimal ex-post. The same holds for *Commitment*. In contrast, under *Regulatory Flexibility*, the firm faces a carbon price equal to the social costs of carbon, such that it does not distort the production decision. Depending on the level of risk aversion and the probability of ex-post socially not optimal production, either *Regulatory Flexibility* or offering a CCfD is welfare superior.

Our paper contributes to two broad streams of literature in the context of irreversible investments in low-carbon technologies in the presence of risk.

The first literature stream focuses on policy options when firms face irreversible decisions. Baldursson and Von der Fehr (2004) analyze policy outcomes in a model in which firms choose between an irreversible long-term investment in

[29]Literature suggests that regulators may have an incentive to deviate from announced carbon prices ex-post, implying regulators may not be able to credibly commit (e.g. Helm et al., 2003).

abatement under risk and a short-term abatement option after the risk resolves. In the presence of risk aversion, the authors show that committing to a carbon tax ex-ante outperforms flexible carbon prices stemming from tradable permits because the latter increase the firms' risk exposure. Jakob and Brunner (2014) show that regulators can combine the advantages of flexibility and commitment by not committing to a specific climate policy level but a transparent adjustment strategy in response to climate damage shocks. In reality the regulator may need to address not only the optimal level of an irreversible investment decision but also the optimal consumption level. Höffler (2014) points out that regulators should address each target with a separate instrument. Therefore, a hybrid policy, i.e. the combination of two policies may be necessary. Offering a CCfD in addition to carbon prices constitutes a hybrid policy in the sense that the CCfD targets the firms' investment decisions while the complementary carbon price targets the optimal consumption level. Closely linked to our paper, Christiansen and Smith (2015) extend the analysis of Baldursson and Von der Fehr (2004) to hybrid policy instruments. The authors analyze a sequential setting in which firms initially have to decide on an investment in a low-carbon technology under risk and subsequently adjust output after the risk resolves. If a carbon tax commitment is the only instrument, the regulator sets the tax higher than the expected damage to incentivize more appropriate investments.[30] Supplementing the carbon tax with a state-contingent investment subsidy increases welfare as it allows for incentivizing investment without setting a carbon tax that is too high. In a similar vein, Datta and Somanathan (2016) analyze a carbon tax and a permit system and examine the role of research and development (R&D) subsidies. They conclude that using only one instrument cannot be welfare-optimal if the regulator aims to address two targets - the internalisation of external effects from R&D and carbon damage. This is in line with our finding that a hybrid policy, in our case a CCfD, can improve welfare in a setting with an irreversible investment decision.

The second literature stream examines the role of hedging instruments for incentivising investments in low-carbon technologies under risk. Within this literature stream, the introduction of hedging instruments are found to increase investments in the presence of risk aversion. Borch (1962), who analyzes reinsurance markets, demonstrates that players are willing to share risks according to their level of risk aversion by trading reinsurance covers which act as hedging instruments. This finding is supported by Willems and Morbee (2010), who examine investments in energy markets. The authors find that the availability of hedging opportunities increases investments of risk-averse firms and welfare. Habermacher and Lehmann (2020) analyze the interaction between a regulator aiming to maximise welfare and firms facing an investment decision in low-carbon technologies. Similar to our paper, the authors assess carbon damage and variable costs risk. They find

[30]This result resembles the insights from the real options literature where risk, combined with investment irreversibility, gives rise to an option value of waiting, e.g., Dixit et al. (1994). Chao and Wilson (1993) find an option value for emission allowances. Purchases of emission allowances provide flexibility to react to risk in a way that irreversible investments do not. The price of emission allowances may therefore exceed the marginal cost of abatement.

that the introduction of stage-contingent payments which partly hedge the risks of the regulator and the firm improve welfare compared to committing to carbon price or setting it flexibly. Those findings are in line with our result that a CCfD as an instrument for firms to hedge their risk leads to more investment and may increase welfare. Furthermore, hedging instruments may improve welfare even in the absence of risk aversion. An early example is Laffont and Tirole (1996), who show that the introduction of options solves the problems arising from strategic behaviour between the regulator and a firm.[31] If the regulator faces incomplete information, Unold and Requate (2001) show that offering options in addition to permits is welfare-enhancing. In contrast to this stream of literature, Quiggin et al. (1993) find that hedging instruments may also be welfare-deterring, as they may foster undesired behaviour. This result resembles our findings in the case of potentially ex-post welfare-reducing production in Chapter 3.4.

CCfDs combine the effects of a hybrid policy and a hedging instrument. They recently gained attention from academic literature. Richstein (2017) focuses on the optimal combination of CCfDs and investment subsidies to lower policy costs and support investment decisions under risk and risk aversion. However, the study does not include the regulator's decision on the carbon price regime. To the best of our knowledge, Chiappinelli and Neuhoff (2020) provide the only study that explicitly analyzes CCfDs in the context of multiple carbon pricing regimes. The authors model firms which face an irreversible investment decision and behave strategically, which influences the regulator's decision on the carbon price. In this setup, higher investments in abatement technologies lead to lower carbon prices so that firms strategically under-invest to induce higher carbon prices. Offering CCfDs can alleviate such a hold-up problem. We build on the model developed in Chiappinelli and Neuhoff (2020) but change the focus of analysis. We analyze a setup with a large number of small firms in a competitive market. Chiappinelli and Neuhoff (2020) show how CCfDs can alleviate the hold-up problem that results from regulation and, hence, mitigate regulatory risk. In contrast, we focus on the impact of CCfD in an environment of risks that are outside the control of regulator and firms, i.e., damage and variable cost risk. We also present the first paper in this literature stream to point out that CCfDs can cause a lock-in in technologies that are ex post not socially optimal.

3.2. Carbon pricing regimes in the absence of risk

This section introduces the model setup to analyse the effects of CCfDs. In the model, we assess the interactions between a regulator and firms in the absence of risk. The regulator can apply three carbon pricing regimes to reduce emissions while firms face an irreversible investment decision to abate emissions during production.

[31] This type of expropriation game constitutes a type of climate policy risk but mainly includes strategic behaviour.

3.2.1. Model framework in the absence of risk

We model the market for a homogeneous good G in which three types of agents participate - namely, consumers, firms, and a regulator. Consumers have an elastic demand $Q(p_G)$ for the good at a market price p_G. Demand decreases in the good's price, i.e., $Q'(p_G) < 0$.

A continuum of firms supplies the good in a competitive market. Each firm produces one unit. Initially, all firms produce the good with a conventional technology. Using the conventional technology to produce one unit of G induces constant marginal production costs ($c_0 \geq 0$) that are identical among all firms. The production process emits one unit of carbon emission. The emission causes constant marginal environmental damage d, which lowers the overall welfare, and is subject to a carbon price ($p \geq 0$). The resulting total marginal costs of the conventional technology equal $c_c = c_0 + p$.[32]

Firms can invest in an emission-free technology to produce G at carbon costs of zero. Investing implies that firms adopt new production processes within their existing production sites. As a result, the production capacity of the firms remains unaffected by an investment.[33] The investment decision is irreversible and induces investment costs as well as higher marginal production costs. We assume firms face heterogeneous investment costs, similar to the approach in Harstad (2012) or Requate and Unold (2003).[34] This heterogeneity may stem from several sources, e.g., because firms can adopt different technologies, have different access to resources, or have different R&D capacities. In our model, firms are ranked from low to high investment costs, such that they can be placed within an interval ranging from $[0, \chi_{max}]$.[35] We assume the firm-specific investment costs to be the product of the firm-specific position on the interval χ and a positive investment cost parameter c_i that is identical among firms. Hence, the investment costs of the firm positioned at χ equal $C_i(\chi) = \chi c_i$. Firms invest if they increase their profit by adopting the emission-free technology. Otherwise, they produce conventionally. We identify the firm which is indifferent between the two technologies by $\overline{\chi}$. As $C_i'(\chi) > 0$, all firms with $\chi \leq \overline{\chi}$ invest. In other words, $\overline{\chi}$ refers to the marginal firm investing in the emission-free technology. The position of a firm on the interval χ not only defines the firm-specific investment costs but also corresponds to the cumulative production capacity of all firms facing investment costs lower than the respective firm. In consequence, $\overline{\chi}$ defines the emission-free production capacity. In the following, we refer to $\overline{\chi}$ interchangeably either as the emission-free production capacity or as the marginal firm.

[32]We discuss the implication of assuming constant marginal damage in Chapter 3.5.

[33]This does not exclude market entry of new firms; however, we do not model entry or exit decisions explicitly, as adopting new processes in established installations is likely less costly then investing in new installations.

[34]Empirical evidence shows that firms differ with respect to their costs of investing in pollution abatement Blundell et al. (2020).

[35]χ_{max} represents the production capacity of all firms and is assumed to exceed the demand $Q(p_G)$ for all possible values of p_G.

Emission-free production has additional marginal production costs c_v. This technology may, for instance, require more expensive input factors compared to the conventional technology. Hence, the total marginal production costs of firms using the emission-free technology equal $c_f = c_0 + c_v$. In Chapter 3.2 and 3.3, we assume the marginal production costs of the emission-free technology to be lower than the carbon price (i.e., $c_v < p$). We alleviate the assumption in Chapter 3.4. Additionally, we adopt the normalisation $c_0 = 0$. Considering investment and production costs, the profit of investing in the emission-free technology equals $\pi(\chi) = p_G - (c_0 + c_v + c_i\chi)$.

The regulator aims at maximising the welfare resulting from the market for G. For this, the regulator can choose among the three different carbon pricing regimes. Firstly, she can opt for *Regulatory Flexibility* (short: *Flex*), in which she sets the carbon price flexibly after the investment decisions of the firms took place. Secondly, she can make a *Commitment* (short: *Com*) and commit to a carbon price before the investment takes place. The third option *CCfD* is a hybrid policy of offering CCfDs to the firms before the investments take place and setting the carbon price afterwards. The CCfD sets a strike price p_s that safeguards firms against carbon price volatility. If the carbon price, which realises after the investments, is lower than the strike price, the regulator pays the difference $(p_s - p)$ to the firm. If the carbon price is higher than the strike price, firms have to pay the difference to the regulator.

Before introducing the sequence of actions, we discuss the model approach and its main assumptions. First, a price-elastic demand, a competitive market structure, and the provision of homogeneous goods resemble many industries for which CCfDs are proposed, e.g., steel and chemicals (e.g. European Commission, 2021b, Fernandez, 2018, OECD, 2002). Second, these industries likely face a discrete, irreversible investment decision to decarbonise the production in combination with increased marginal production costs of the low-carbon technology. Currently, a switch of production processes from the coal- and coke-based blast furnace to hydrogen-based direct reduction is seen as the most promising way to decarbonise the primary steel sector (e.g. IEA, 2021). This switch in the production process induces a shift in input factors from coal to more expensive hydrogen (Vogl et al., 2018). Hence, our model captures many characteristics of industries, for which policymakers propose the use of CCfDs.

The agents in our model can take actions in four stages, namely the Early Policy stage t_1, the Investment stage t_2, the Late Policy stage t_3, and the Market Clearing stage t_4. Figure 3.1 depicts these stages. The sequence of actions differs between the carbon pricing regimes that we analyse in this paper. We subsequently discuss the agents' actions during the various stages of the game. As we derive the sub-game perfect Nash equilibrium by backward induction, we begin by presenting the last stage of the game.

	Firms	**Regulator** *Com*	**Regulator** *Flex*	**Regulator** *CCfD*	**Social Planner** *Opt*
Early Policy (t_1)		Sets p		Sets p_s	
Investment (t_2)	Invest up to $\overline{\chi}$				Sets $\overline{\chi}$
Late Policy (t_3)			Sets p	Sets p	Sets p
Market Clearing (t_4)	$p_G^* = p$ and $Q(p_G^*) = Q(p)$				

T

Figure 3.1.: Sequence of actions in the different carbon pricing regimes.

Market Clearing stage: In t_4, the market clearing takes place. Firms produce the good with the respective technologies and serve the demand. In this stage, the carbon price p and the resulting emission-free production capacity $\overline{\chi}$ are already determined.

Late Policy stage: In t_3, the regulator sets the carbon price under *Regulatory Flexibility* and *CCfD*, given the previously determined production capacity of the emission-free technology.

Investment stage: In t_2, the firms decide whether to invest in the emission-free technology or not. Firms with $\chi \leq \overline{\chi}$ invest as they increase their profit by adopting the emission-free technology, while the others ($\chi > \overline{\chi}$) maintain the conventional technology.

Early Policy stage: In t_1, the regulator can take actions in two of the three carbon pricing regimes. Under *Commitment*, she announces and commits to a carbon price for the subsequent stages. Under *CCfD*, the regulator offers firms CCfDs and determines the strike price.

In contrast to the other stages, the market clearing in t_4 is independent of the carbon pricing regime, such that we present the result upfront. We assume the investment costs to be sufficiently high compared to the demand, such that investments in the emission-free capacity cannot supply the overall demand, i.e., $\overline{\chi} < Q(p_G)$. This assumption implies that the demand for the good is partially served by firms that invested in the emission-free technology and by firms producing conventionally.[36] As demand exceeds the emission-free production capacity and marginal production costs of the emission-free technology are lower than of the conventional technology, the latter sets the market price. Due to the normalisation of $c_0 = 0$, the market price is defined by $p_G = p$ and the demand is equal to $Q(p_G) = Q(p)$, i.e., the carbon price fully determines the product price. Figure 3.2 illustrates the market clearing.

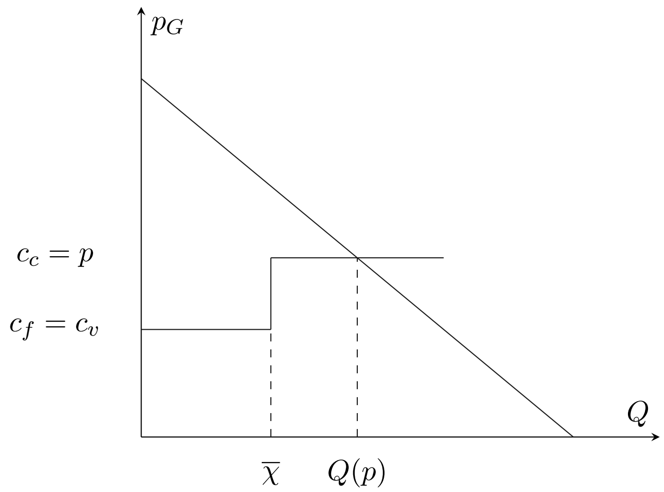


Figure 3.2.: Market clearing.

Firms producing the good with the conventional technology do not generate profits as marginal revenue equals marginal costs, which are constant. The marginal profit of production of the firms investing in the emission-free technology equals $p - c_v$. Together, the assumptions $\chi < Q(p_G)$ and $c_v < p$ ensure that some firms will invest in the emission-free technology. The first assumption addresses the fixed investment costs and the second the variable costs of the emission-free technology. These assumptions also ensure that some firms continue producing conventionally. Chapter 3.5 discusses why CCfDs can only be beneficial in this setting.

To evaluate the carbon pricing regimes, we compare the respective outcomes to the social optimum (short: *Opt*). In this hypothetical benchmark, a social planner sets the socially optimal investment in t_2 and the carbon price level in t_3. The social planner's objective is, identical to the regulator, to maximise social welfare stemming from the market for the product G. Social welfare comprises four elements: 1) net consumer surplus (CS), 2) producer surplus, 3) environmental damage, and 4) policy costs/revenues from carbon pricing and the CCfD.

The producer surplus is defined as the margin between marginal revenue and marginal costs. It differs before and after the irreversible investment. Before the

[36]We discuss this assumption in Chapter 3.5, as it is crucial for the outcome of the market clearing and the resulting incentives to invest in the emission-free technology.

investment, i.e., in t_1 and t_2, the marginal costs comprise investment and marginal production costs. After the investment, i.e., in t_3 and t_4, the investment costs are sunk, such that the marginal costs only comprise the marginal production costs. Equation 3.1 displays the welfare before the investment takes place. The welfare representation after the investment takes place does not contain the investment costs $\int_0^{\overline{\chi}}(c_i z)dz$.

$$\mathcal{W}^{Flex/Com/Opt} = \underbrace{\int_p^{\infty} Q(z)dz}_{\text{consumer surplus}} + \underbrace{\int_0^{\overline{\chi}}(p - c_v - c_i z)dz}_{\text{producer surplus}} - \underbrace{d\big[Q(p) - \overline{\chi}\big]}_{\text{environmental damage}} + \underbrace{p\big[Q(p) - \overline{\chi})\big]}_{\text{revenues from carbon pricing}}$$

$$\mathcal{W}^{CCfD} = \underbrace{\int_p^{\infty} Q(z)dz}_{\text{consumer surplus}} + \underbrace{\int_0^{\overline{\chi}}(p_s - c_v - c_i z)dz}_{\text{producer surplus}} - \underbrace{d\big[Q(p) - \overline{\chi}\big]}_{\text{environmental damage}} + \underbrace{p\big[Q(p) - \overline{\chi})\big]}_{\text{revenues from carbon pricing}} - \underbrace{(p_s - p)\overline{\chi}}_{\text{CCfD payment}} \tag{3.1}$$

Payments arising from the CCfD do not affect the overall welfare as they only shift payments between firms and the regulator.[37] Hence, we can simplify welfare with and without CCfDs before investment to:

$$\mathcal{W} = \int_p^{\infty} Q(z)dz + (p - d)Q(p) + \int_0^{\overline{\chi}}(d - c_v - c_i z)dz \tag{3.2}$$

This simplified representation illustrates that welfare can be grouped into two elements. On the one hand, welfare is defined by consumption, the associated environmental damage, and the carbon pricing revenue. On the other hand, welfare stems from the level of emission-free production capacity $\overline{\chi}$ and the related costs and benefits from abatement.

3.2.2. Policy ranking in the absence of risk

In the following, we derive the optimal emission-free production capacity $\overline{\chi}$ and the optimal carbon price p in the absence of risks (i.e., under perfect foresight) under the assumption of a social planner. The solution serves as a hypothetical benchmark for the three carbon pricing regimes. To solve the optimisation of the

[37] Note that we do not assume shadow costs of public funds. We discuss this assumption in Chapter 3.5.

social planner, we derive the first-order conditions of the welfare function:

$$\begin{aligned} \max_{\overline{\chi},p} \mathcal{W} &= \int_p^\infty Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}} (d - c_v - c_i z)dz \\ \frac{\partial \mathcal{W}}{\partial \overline{\chi}} &= (d - c_v - c_i\overline{\chi}) \longrightarrow \overline{\chi}^{Opt} = \frac{d - c_v}{c_i} \\ \frac{\partial \mathcal{W}}{\partial p} &= -Q(p) + Q(p) + Q'(p)(p-d) \longrightarrow p^{Opt} = d \end{aligned} \tag{3.3}$$

The social planner chooses the emission-free production capacity such that the abatement costs (i.e., the investment and production costs) of the marginal firm ($\overline{\chi}^{Opt}$) equal the damage avoided by the investment in and the utilisation of the emission-free technology. The optimal carbon price (p^{Opt}) equals the marginal damage, i.e., the Pigouvian tax level (Pigou, 1920), as the marginal unit of the good is produced with the conventional technology, associated with an environmental damage of d. With this carbon price, the social planner inhibits all consumption with a lower benefit than damage to society.

We provide the optimal solutions under the different carbon pricing regimes in Supplementary Material B.1. We find that

Proposition 3.2.1. *In the absence of risk, all carbon pricing regimes reach the social optimum. In all regimes, the carbon price is equal to the marginal environmental damage of production, i.e., $p = d$. The marginal firm using the emission-free technology balances the marginal investment costs and the respective marginal benefit of abatement, i.e., $\overline{\chi} = {}^{(d-c_v)}/_{c_i}$.*

In the absence of risk, i.e., under perfect foresight, the optimisation rationales in t_1 (before investing) and t_3 (after investing) regarding balancing the damage from carbon emission and the costs of abatement are identical. Therefore, it does not make a difference if the regulator commits to a carbon price before the firms invest or sets the carbon price flexibly afterward. Under all regimes, Pigouvian taxation is optimal. Hence, offering a CCfD in t_1 does not improve social welfare.

This result regarding the welfare ranking of carbon pricing regimes and, notably, CCfDs differs from Chiappinelli and Neuhoff (2020). In their model, firms also face an irreversible investment decision but behave strategically and influence the regulator's decision on the carbon price. Thereby, firms under-invest to induce higher carbon prices, leading to a hold-up problem. In this setting, CCfDs can alleviate the investment-hampering effect of flexible carbon prices and increase welfare. In contrast, firms do not have market power in our model and cannot affect the regulator's carbon pricing decision. Hence, it does not make a difference if the firms invest before or after the regulator sets the carbon price under perfect foresight.

Proof. We provide the proof of Proposition 3.2.1 in Supplementary Material B.1.
■

3.3. Carbon pricing regimes in the presence of risk

In this chapter, we analyse the impact of damage and variable cost risk on the welfare ranking of the carbon pricing regimes in the presence of risk aversion.

3.3.1. Model framework in the presence of risk and socially optimal production

We integrate risk into the model by redefining the marginal environmental damage and the variable production costs of the emission-free technology from the model introduced in section 3.2.1 as random variables D and C_v. Both random variables realise after the firms invest in abatement (t_2), but before the late policy stage (t_3) and the market clearing (t_4). We denote the realisation of D and C_v by $\hat{d}$ and $\hat{c_v}$. In this chapter, we assume the production with the emission-free technology to be socially optimal under all circumstances, i.e., the environmental damage is always larger than the variable costs of abatement $P(D > C_v) = 1$. For this assumption to hold, we define the random variables to follow a truncated normal distribution, i.e., $D \sim TN(\mu_D, \sigma_D^2, \underline{\theta_D}, \overline{\theta_D})$ and $C_v \sim TN(\mu_{C_v}, \sigma_{C_v}^2, \underline{\theta_{C_v}}, \overline{\theta_{C_v}})$ with $\underline{\theta_D} > \overline{\theta_{C_v}}$, where μ denotes the mean value, σ^2 the variance and $\underline{\theta}$ and $\overline{\theta}$ the lower and upper limit of the distribution, respectively. Hence, the lowest possible damage is larger than the highest possible realisation of variable costs.[38] As in Chapter 3.2.1, we assume $\chi < Q(p(d))$, such that for all $\hat{d} \in D$ the total demand in the market exceeds the emission-free production capacity.

[38]We assess a setting in which the social costs of damage are potentially smaller than the variable costs of abatement, i.e., $P(D > C_v) < 1$, in Chapter 3.4 by assuming an non-truncated normal distribution.

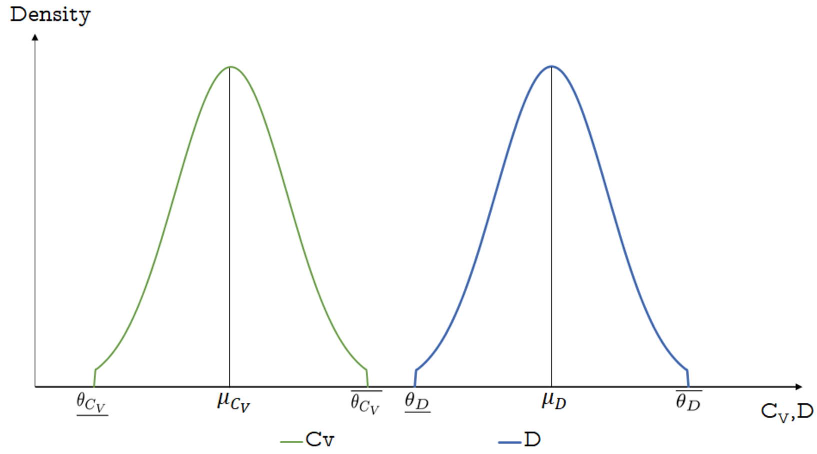


Figure 3.3.: Density of D and C_v following a truncated normal distribution with $P(C_v > D) = 0$.

We assume that firms are risk averse, facing a utility that is exponential in profits. Whether or not risk aversion is a real-world phenomenon for firms and how it manifests in actions is debated within the broad literature of economics and the context of energy and environmental economics (Meunier, 2013). Diamond (1978) argues that even if markets were incomplete, firms should act as if they were risk neutral, and shareholders could hedge their risks at the capital markets. However, there are several reasons why firms may act aversely to risk (see e.g. Banal-Estañol and Ottaviani (2006) for a review). These reasons include non-diversified owners, liquidity constraints, costly financial distress, and nonlinear tax systems. Additionally, and independently of the owners' risk aversion, the delegation of control to a risk-averse manager paid based on the firm's performance may cause the firm to behave in a risk-averse manner.

How the firms' risk aversion can be modelled depends on the distributional assumptions of the underlying risks. Markowitz (1952) show that for non-truncated normally distributed profits, the mean-variance utility could express firms' optimisation rationale. However, this simplification is not appropriate for our model in which the distribution of firms' profits is truncated due to distributional assumptions on damage and variable cost risk. Norgaard and Killeen (1980) show that the optimisation rationale of an agent facing an exponential utility and truncated normally distributed profits can be approximated by a mean-standard deviation decision rule containing a risk aversion parameter λ.[39] We apply this approximation by using a mean-standard deviation utility in our model. Firms invest in the emission-free technology if their expected utility is positive. The expected utility of the marginal firm investing in the emission-free technology is

[39] In the context of energy and environmental economics, Alexander and Moran (2013) apply this approach to assess the impact of perennial energy crops income variability on the crop selection of risk-averse farmers.

equal to zero:

$$\begin{aligned} EU(\pi(\overline{\chi})) &= \mu_\pi(\overline{\chi}) - \lambda\sigma_\pi(\overline{\chi}) \\ &= (\mu_p - \mu_{C_v} - c_i\overline{\chi}) - \lambda\sigma_{p,C_v} \\ &= 0 \end{aligned} \tag{3.4}$$

In contrast to the firms' risk aversion, we assume the regulator to be risk neutral. There are several reasons why environmental regulation is determined on a risk-neutral basis (see e.g. Kaufman (2014) for an extensive review). In the context of public economics, Arrow and Lind (1970) argue that with a sufficiently large population, the risk premiums converge to zero because they can be spread out among constituents. Fisher (1973) discusses the principles of Arrow and Lind in the context of risks stemming from environmental externalities. [40] Hence, we assume the regulator to maximise the expected welfare:

$$E[\mathcal{W}] = E\left[\int_p^\infty Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}}(d - c_v - c_i z)dz\right] \tag{3.5}$$

3.3.2. Policy ranking with damage risk

In the following, we focus on damage risk and neglect the risk of the variable production costs. Therefore, we set $\mu_{C_v} = c_v$ with $\sigma^2_{C_v} = 0$. We derive and compare the outcomes of the three carbon pricing regimes in terms of the emission-free production capacity $\overline{\chi}$ and carbon price p in the presence of damage risk. We contrast the three regimes to the social optimum and conclude that

Proposition 3.3.1. *In the presence of damage risk and firms' risk aversion, only the hybrid policy of offering a CCfD and setting the carbon price flexibly yields an socially optimal level of* p *and* $\overline{\chi}$*. A pure carbon pricing regime reaches either a socially optimal carbon price through allowing for flexibility or optimal investment through early commitment.*

As the valuation of environmental damage is not known before investing (t_1), while it is known after investing (t_3), the timing of the carbon pricing regimes changes the carbon prices and the resulting market outcomes. When setting the carbon price flexibly in t_3, all relevant information is available for the regulator. Hence, the *Regulatory Flexibility* regime results in the socially optimal carbon price for the market clearing. However, in this regime, firms face a risk regarding their revenues. Due to their risk aversion, firms consequently invest less than

[40] Besides the risk neutrality of the regulator, we assume that her welfare maximisation is also not affected by the firms' risk aversion. This corresponds to the concept of the literature on non-welfarist taxation, which is common practice in public economics (e.g. Heutel (2019), Kanbur et al. (2006)). In essence, the regulator's *ignorance* of the risk-averse utility of the firms can stem from either paternalistic behaviour or an insufficiently large proportion of the firms on the market.

socially optimal. When committing to a carbon price in t_1, the regulator cannot take into account the information becoming available in t_3. Hence, the carbon price under *Commitment* is ex-post either too high or too low. However, the carbon price commitment incentivises socially optimal investments. It accounts for the risk in the valuation of environmental damage; that is, the firms and the regulator face the same problem. Offering a CCfD removes the impact of damage risk for the firms and enables socially optimal investments. Furthermore, socially optimal consumption is reached as the regulator sets the carbon price in t_3, having complete information on the damage valuation.

Proof. For the proof of proposition 3.3.1, we compare the socially optimal carbon price and emission-free production capacity to the three carbon pricing regimes. Supplementary Material B.2 presents a complete derivation of the respective optimal solutions. In the following, we provide the main results and the intuition behind the finding in proposition 3.3.1.

Social optimum

In the social optimum, the social planner sets the carbon price p after the actual environmental damage revealed. Following the rationale of the risk-free setting, the socially optimal carbon price equals the realised marginal damage, i.e., $p^{Opt} = \hat{d}$. As the social planner knows the actual damage level when setting the carbon price, the damage risk does not impact her decision.

In contrast, investments are due before the actual damage reveals. Hence, the social planner must set the emission-free production capacity $\overline{\chi}$ in the presence of damage risk. The social planner sets $\overline{\chi}^{Opt}$ such that it maximises the expected welfare gain from abatement investments.

$$\overline{\chi}^{Opt} = \frac{\mu_D - c_v}{c_i} \tag{3.6}$$

The emission-free production capacity balances the expected benefit of abatement, i.e., the expectation of the avoided environmental damage and the abatement costs, consisting of variable production costs and investment costs.

Regulatory flexibility

Similar to the social planner case, the regulator sets the carbon price after the actual damage revealed when she chooses *Regulatory Flexibility.* As the regulator and the social planner have the same objective function, both settings result in a carbon price at $p^{Flex} = p^{Opt} = \hat{d}$, i.e. the Pigouvian tax level.

In t_2, the firms choose to invest if their expected utility is positive, anticipating the carbon price set by the regulator in the following stage. However, the price is stochastic to firms, as it depends on the realised damage.

$$\overline{\chi}^{Flex} = \frac{\mu_{p^{Flex}} - c_v - \lambda\sigma_{p^{Flex}}}{c_i} = \frac{\mu_D - c_v - \lambda\sigma_D}{c_i} \quad (3.7)$$

Unlike in the case of a (risk-neutral) social planner, firms not only account for the expected revenues and costs of abatement but also consider a risk term stemming from the abatement revenue risk. This risk term reduces the firms' expected utility and consequently the emission-free production capacity, as firms aim to avoid situations where their investments are unprofitable. The dampening effect of risk on investments increases with the volatility of expected carbon prices and the firms' risk aversion.

Commitment

Under *Commitment*, the firms' investment rationale is based on the carbon price known at the time of taking their decision:

$$\overline{\chi}^{Com} = \frac{p^{Com} - c_v}{c_i} \quad (3.8)$$

Following the intuition of the setting without risk, those firms invest which increase their profit by adopting the emission-free technology. As revenues are not subject to risk, the firms' risk aversion does not impact their investment decisions in t_2 and the resulting emission-free technology balances the marginal revenue and the marginal costs of abatement.

In t_1, the regulator sets the carbon price maximising expected welfare and taking into account that firms solely invest if the investment is profitable. As a result, the regulator sets the carbon price to $p^{Com} = \mu_D$, i.e., the expected Piguvian tax level. Substituting the optimal carbon price p^{Com} into Equation 3.8 yields $\overline{\chi}^{Com} = \frac{\mu_D - c_v}{c_i}$, which is equal to the solution of the social planner. However, the carbon price to which the regulator commits herself in t_1 is ex-post not optimal. If the revealed damage is greater than expected, the carbon price is too low, and vice versa.

CCfD

When the regulator can offer the firms a CCfD, the regulator faces the same objective function for setting the carbon price in t_3 as under *Regulatory Flexibility*. Hence, she chooses the Pigouvian tax level $p^{CCfD} = p^{Flex} = p^{Opt} = \hat{d}$.

In t_2, the firms' problem is identical to the one under *Commitment*. Here, the firms receive the strike price:

$$\overline{\chi}^{CCfD} = \frac{p_s - c_v}{c_i} \tag{3.9}$$

The rationale for investments is the same as without risk: Firms invest in the emission-free technology if it increases their profits. In t_1, the regulator chooses the strike price that maximises expected social welfare. She accounts for the firms' reaction function to the announced strike price and faces damage risk. The resulting strike price equals the expected marginal damage, i.e., $p_s = \mu_D$. By substituting p_s in Equation 3.9, we see that under a *CCfD* regime, the emission-free production capacity equals the one under *Com* (and the social planner), i.e., $\overline{\chi}^{CCfD} = \overline{\chi}^{Com} = \overline{\chi}^{Opt}$). ■

Welfare Comparison

We calculate and compare the ex-ante social welfare in the different carbon pricing regimes in terms of welfare.[41] We find that:

$$E[\mathcal{W}_{\sigma_D}^{Opt}] = E[\mathcal{W}_{\sigma_D}^{CCfD}] \geq E[\mathcal{W}_{\sigma_D}^{Com}] \lesseqgtr E[\mathcal{W}_{\sigma_D}^{Flex}] \tag{3.10}$$

First, the carbon price and the emission-free production capacity are identical in the social optimum and the *CCfD* regime. Consequently, the *CCfD* regime results in the social optimum.

Second, we compare offering a CCfD against *Regulatory Flexibility* and *Commitment*. While the *CCfD* regime achieves the socially optimal emission-free production capacity, investments in *Flex* are lower. As the expected welfare increases in χ as long as $\chi \leq \overline{\chi}^{CCfD} = \frac{\mu_D - c_v}{c_i}$, the welfare under the *Flex* regime is lower than the social optimum or offering a CCfD. The welfare loss increases in the firms' risk aversion and the standard deviation of environmental damage. However, if firms are risk neutral, the *Flex* regime reaches the socially optimal emission-free production capacity. Figure 3.4i shows these results numerically. Note that these parameter values are illustrative and do not correspond to empirical estimates.[42] In contrast to the case of *Regulatory Flexibility*, the policy regimes *Commitment* and *CCfD* both result in the socially optimal emission-free production capacity. However, these regimes differ concerning the carbon price level and the resulting utility from consumer surplus. Under the *Com* and *CCfD* regimes, consumers bear the same carbon prices in expectation. However, the

[41] The subscript σ_D represents the welfare in the presence of damage risk.

[42] Both Figure 3.4i and Figure 3.4ii share the parameters regarding the distribution of the environmental damage $D \sim TN(\mu_D = 4, \sigma_D^2 = 0.25, \underline{\theta_D} = 2.5, \overline{\theta_D} = 5.5)$ and the cost parameters of the emission-free technology $c_v = 2$ and $c_i = 4$.

consumer surplus is a convex function of the respective carbon price. I.e., a higher carbon price decreases the consumer surplus less than an equivalently lower carbon price would lead to an increase of the consumer surplus.[43] Hence, the difference in expected consumer surplus is positive, i.e., $E[\int_{p^{CCfD}}^{\infty} Q(z)dz] > \int_{p^{Com}}^{\infty} Q(z)dz$. With an increase in demand elasticity, the difference in consumer surplus of the *Com* and *CCfD* regimes increases. Therefore, the greater the demand elasticity, the higher the loss in ex-ante welfare arising from not setting the carbon price according to the actual marginal damage under *Com*. We illustrate this finding numerically in Figure 3.4ii.

Third, it is unclear whether *Com* or *Flex* is welfare superior. *Flex* results in socially optimal carbon pricing, while *Com* allows for socially optimal emission-free production capacity. Which regime is welfare superior depends on the relevance of the two variables. In case of damage risk, setting a flexible carbon price is welfare superior to *Com* if demand elasticity is sufficiently high and the share of emission-free production is sufficiently low. The same holds vice versa for *Com*.

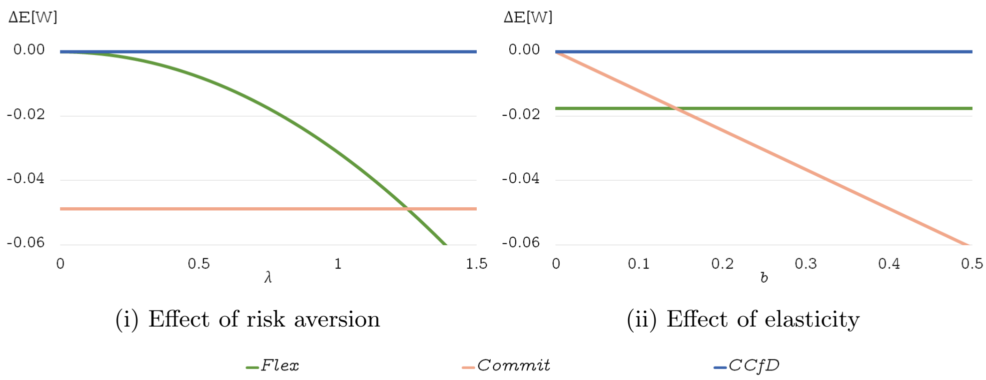


Figure specific parameters in (a): $\lambda \in [0, 1.5]$, $Q(p) = 5 - 0.4p$ and (b): $\lambda = 0.75$, $Q(p) = 5 - bp$ with $b \in (0, 0.5]$.

Figure 3.4.: Difference in welfare compared to social optimum in the presence of damage risk.

3.3.3. Policy ranking with variable cost risk

In this chapter, we focus on variable cost risk and set $\mu_D = d$ with $\sigma_D^2 = 0$. We derive the outcomes of the three carbon pricing regimes in terms of emission-free production capacity $\overline{\chi}$ and carbon price p when the firms do not know the variable costs of the emission-free technology when investing. We contrast the three regimes with the social optimum and conclude that

Proposition 3.3.2. *In the presence of variable cost risk, only the hybrid policy of offering a CCfD and setting the carbon price flexibly yields a socially optimal level*

[43]This relation is also known as the Jensen gap stemming from Jensen's inequality.

of p and $\overline{\chi}$. A pure carbon price in a regime with Regulatory Flexibility reaches a socially optimal carbon price p but falls short of the socially optimal emission-free production capacity $\overline{\chi}$. Commitment reaches neither the socially optimal level of p nor $\overline{\chi}$.

When firms face a variable abatement costs risk, risk aversion reduces the utility from investing in the emission-free production technology. Depending on the carbon pricing regime, the regulator can mitigate this effect. The regulator can encourage firms to increase investments by setting the carbon price above the Pigouvian tax level when committing to a carbon price. However, the price increase results in inefficient consumption levels. Hence, the regulator faces a trade-off between high consumer surplus and low environmental damage, resulting in a deviation from the social optimum. When the regulator can offer a CCfD in addition to a carbon price, she does not face this trade-off. Instead, the regulator can offer a CCfD, which sufficiently compensates firms for facing risk regarding their revenue and enable socially optimal investments. Furthermore, the regulator achieves the socially optimal consumption level. She can set the carbon price to the Pigouvian tax level, indicating the benefit of having two instruments for different objectives. If the regulator cannot offer a CCfD and sets the carbon price flexibly, the regulator achieves the socially optimal consumption level but cannot alter the firms' investment decisions. Consequently, fewer firms invest than socially optimal.

Proof. For the proof of proposition 3.3.2, we compare the socially optimal carbon price and emission-free production capacity to the three carbon pricing regimes. Supplementary Material B.3 presents a complete derivation of the respective optimal solutions. In the following, we provide the main results and the intuition behind the finding in proposition 3.3.2.

Social optimum

In the social optimum, the social planner maximises welfare by setting the carbon price p^{Opt} after the level of variable costs revealed. She chooses the Pigouvian tax level $p^{Opt} = d$, which equals the social marginal costs of production.

The social planner sets the emission-free production capacity $\overline{\chi}^{Opt}$ under risk such that it maximises the expected welfare. The emission-free production capacity balances the marginal benefit and marginal costs from abatement. The optimisation rationale resembles the one under damage risk. However, in this case, not the benefit of emission-free production but its costs are subject to risk:

$$\overline{\chi}^{Opt} = \frac{d - \mu_{C_v}}{c_i} \tag{3.11}$$

Regulatory flexibility

Under *Regulatory Flexibility*, the regulator faces the same optimisation problem as the social planner. Hence, she sets the carbon price to the Pigouvian tax level $p^{Flex} = p^{Opt} = d$.

In t_2, firms invest in the emission-free technology if the investment increases the expected utility of the firm. For this, the firms anticipate the Pigouvian tax. As firms are risk averse, the firms' utility decreases in the level of risk and risk aversion. The resulting emission-free production capacity equals:

$$\overline{\chi}^{Flex} = \frac{p^{Flex} - \mu_{C_v} - \lambda\sigma_{C_v}}{c_i} = \frac{d - \mu_{C_v} - \lambda\sigma_{C_v}}{c_i} \tag{3.12}$$

The emission-free production capacity falls short of the social optimum in case of risk aversion ($\lambda > 0$). The shortfall increases with an increasing level of risk and risk aversion.

Commitment

Under *Commitment*, in t_2, firms choose to invest given the announced carbon price level. As in the case of *Regulatory Flexibility*, firms invest if they generate a positive expected utility, such that the emission-free production capacity equals:

$$\overline{\chi}^{Com} = \frac{p - \mu_{C_v} - \lambda\sigma_{C_v}}{c_i} \tag{3.13}$$

In t_1, the regulator sets the carbon price anticipating that her choice impacts firms' investment decisions and the consumer surplus. These two effects result in a trade-off which we can express as:

$$\frac{p - d}{p} = \frac{1}{\epsilon(p)} \frac{\partial\overline{\chi}^{Com}(p)}{\partial p} \frac{1}{Q(p)}(d - c_i\overline{\chi}^{Com}(p) - \mu_{C_v}), \tag{3.14}$$

where $\epsilon(p) = -\frac{\partial Q(p)}{\partial p}\frac{p}{Q(p)}$ is the elasticity of demand.

The resulting carbon price is higher than d, which we show in Supplementary Material B.3. In fact, the optimal carbon price under commitment p^{Com} ranges from $[d, d + \lambda\sigma_{C_v}]$, depending on the configuration of parameters. Hence, the regulator sets a carbon price above the social marginal costs of the conventional technology, i.e. d, and the carbon price is higher than in the social optimum. The solution is a modified version of the Ramsey formula for monopolistic price setting under elastic demand (Höffler, 2006, Laffont and Tirole, 1996). The regulator increases the carbon price above the socially optimal level to encourage investments. This price mark-up is proportionate to the inverse price elasticity of demand and the marginal benefit from increased investments. The marginal benefit arises from the marginal increase in the share of emission-free production, i.e., $\frac{\partial \overline{\chi}^{Com}(p)}{\partial p} \frac{1}{Q(p)}$, and the benefit of the marginal emission-free production, i.e., $d - c_i \overline{\chi}^{Com}(p) - \mu_{C_v}$. In other words, the regulator balances the loss in consumer surplus and the abatement benefits.

The trade-off under *Com* with variable cost risk is different from the case with damage risk: With damage risk, the regulator commits to a carbon price that will be sub-optimal ex-post. By committing to a carbon price, the regulator takes up the firms' risk, mitigating the negative effect of the firms' risk aversion on social welfare. With cost risk, the regulator cannot take away the firms' risk, but she can compensate the firms for taking the risk. By committing to a carbon price that includes a premium, she incentivises more investments. However, this price increase has the downside of a loss in consumer surplus and, in consequence, neither consumption nor investments are socially optimal. If demand was fully inelastic, i.e., $Q'(p) = 0$, the trade-off would diminish. The regulator would set the carbon price such that she fully compensates the firms for their profit risk, i.e. $d + \lambda\sigma_{C_v}$.

CCfD

When the regulator can offer firms a CCfD in t_1, she sets the carbon price in t_3 after the actual variable costs revealed and firms invested in the emission-free technology. Her optimisation problem is the same as under *Regulatory Flexibility* and the social optimum. Hence, $p^{CCfD} = d$.

In t_2, the firms' optimisation rationale is the same as under the *Commitment*, only that they face a strike price instead of the carbon price.

$$\overline{\chi}^{CCfD} = \frac{p_s - \mu_{C_v} - \lambda\sigma_{C_v}}{c_i} \tag{3.15}$$

In t_1, the regulator chooses a strike price that maximises expected social welfare and accounts for the firms' reaction to the strike price.

$$p_s = d + \lambda\sigma_{C_v} \tag{3.16}$$

In contrast to the previous cases, the regulator sets the strike price above the expected benefit of abatement. By substituting p_s^{CCfD} in Equation 3.15, we see that under a *CCfD* regime, the emission-free production capacity equals the choice of the social planner, i.e., $\overline{\chi}^{CCfD} = \overline{\chi}^{Opt}$. The mark-up $\lambda\sigma_{C_v}$ of the strike price compensates firms for taking the risk. The strike price equals the upper limit of the carbon price under *Commitment*, i.e., the level of p^{Com} with fully inelastic demand. As the strike price does not affect the consumer surplus, the regulator can fully assume the firms' risk. In the absence of risk aversion, the regulator sets the strike price at the level of marginal damage. ■

Welfare Comparison

This subchapter compares the ex-ante social welfare of the different carbon pricing regimes to determine which regime is socially optimal in an environment with risk regarding variable costs. We see that offering a CCfD yields the social optimum, while the other regimes fall short of it. Under *Commitment*, the carbon price is too high and the emission-free production capacity too low. With *Regulatory Flexibility*, the carbon price is socially optimal, but the emission-free production capacity is too low. We find that:

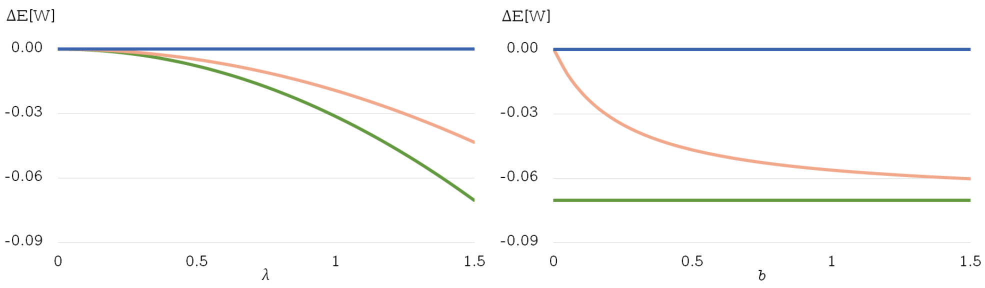
$$E[\mathcal{W}_{\sigma_{C_v}}^{Opt}] = E[\mathcal{W}_{\sigma_{C_v}}^{CCfD}] \geq E[\mathcal{W}_{\sigma_{C_v}}^{Com}] \geq E[\mathcal{W}_{\sigma_{C_v}}^{Flex}] \tag{3.17}$$

First, we compare the expected welfare in *CCfD* with the one the social planner obtains. As both the carbon price and the emission-free production capacity are identical, the *CCfD* regime results in the social optimum.

Second, we find that welfare in *Flex* falls short of the benchmark if firms are risk averse. Like in the case of damage risk, this arises due to too low investments. With increasing risk aversion, the shortfall of investments and welfare increases - a finding that can also be observed numerically in Figure 3.5i.[44]

Third, we find that welfare under *Commitment* falls short of the social optimum but is superior to *Regulatory Flexibility.* The shortfall in welfare arises as the *Com* regime reaches neither the socially optimal carbon price nor the socially optimal emission-free production capacity. The welfare superiority of *Com* compared to *Flex* emerges as the regulator can influence not only the market size but also the investments by setting the carbon price early. In contrast to the damage risk case, there is no disadvantage from setting the carbon price early as the realisation of the damage is known in t_1. When deciding on a carbon price under *Com*, the regulator balances the welfare gain from increased abatement arising from a higher carbon price against the welfare loss from decreased consumption. With an increasing elasticity of demand, e.g., due to an increasing slope of a linear demand function, the welfare loss from setting a higher carbon price increases. Hence, the higher the elasticity, the less the carbon price is increased compared to p^{Flex} by the regulator. In consequence, the relative advantage of *Com* compared to *Flex* decreases with increasing demand elasticity. Figure 3.5ii displays the finding numerically. The analytical proof showing the welfare of *Com* is superior to *Flex* can be found in Supplementary Material B.3.

[44]Both, Figure 3.5i and Figure 3.5ii, share the parameters regrading the distribution of the environmental damage and the costs related to the emission-free technology of Figure 3.4. The chosen parameter values are illustrative and do not correspond to empirical estimates.



(i) Effect of risk aversion

(ii) Effect of elasticity

Figure specific parameters in (a): $\lambda \in [0, 1.5]$, $Q(p) = 5 - 0.4p$ and (b): $\lambda = 1.5$, $Q(p) = 5 - bp$ with $b \in (0, 1.5]$.

Figure 3.5.: Difference in welfare compared to social optimum in the presence of cost risk.

3.4. Carbon pricing regimes with potentially socially not optimal production

In the previous chapter, we focused on the effects of different carbon pricing regimes in settings in which the production of the emission-free technology is always socially optimal in t_4, i.e., the variable costs of abatement are ex-post lower than the marginal environmental damage. In this chapter, we alleviate this assumption and allow for situations in which emission-free production may not be socially optimal.

3.4.1. Model framework in the presence of risk and socially not optimal production

To allow for situations in which the production of the emission-free technology is welfare reducing, we assume the environmental damage to be normally distributed instead of truncated normally distributed. That means there is a positive probability that variable costs exceed the realised damage, i.e. $P(C_V > D) > 0$ (see Figure 3.6).[45] We denote the cumulative distribution and probability density functions of D as $F_D(.)$ and $f_D(.)$. To keep investment in abatement ex-ante socially optimal in all cases, we maintain the assumption that $\mu_D > \mu_{C_V}$.

To emphasise the impact of potentially welfare-reducing production on the different carbon pricing regimes, we assume firms to be risk neutral when analysing

[45] The assumption of an untruncated normal distribution implies that $\chi < Q(p(d)$ cannot hold for all $\hat{d} \in D$. Instead, we can almost ensure that the emission-free capacity cannot cover the total demand by assuming $P(Q(p(d)) < \chi) \to 0$, such that the probability of this case is infinitesimally small and can be neglected.

the problem analytically (Chapter 3.4.2). As the three carbon pricing regimes yield the same outcome in the variable cost risk case if firms are risk neutral (see Chapter 3.3.3), we focus on the damage risk case.[46] Hence, we set $\mu_{C_V} = c_v$ with $\sigma^2_{C_V} = 0$ in the following. Being risk neutral, firms invest if their expected profits are positive, i.e., $E[\pi(\chi)] > 0$. To assess the combined effect of potentially welfare-reducing production and risk aversion, we analyse the model numerically in Chapter 3.4.3.

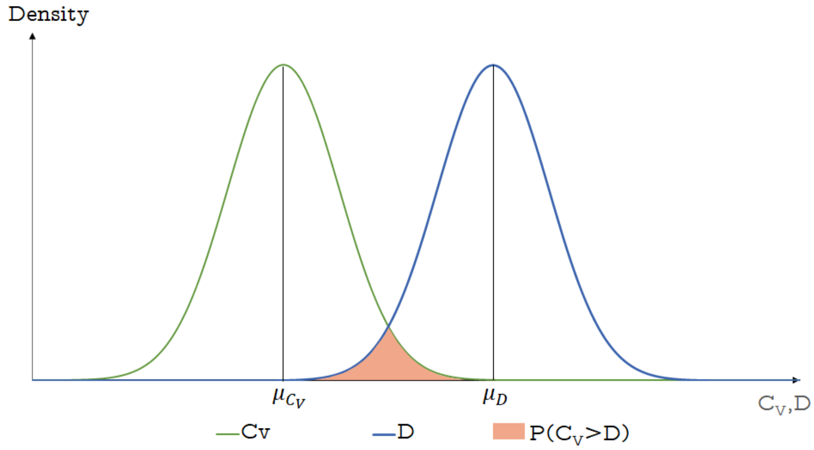


Figure 3.6.: Density of normally distributed D and C_V with $P(C_V > D) > 0$.

Due to the adjusted assumptions on the distribution of damage and costs, the carbon price applied in t_4 may be smaller than the variable costs, such that firms may not produce.[47] Firms may decide not to produce even if they invested in the emission-free technology as investment costs are sunk. The profit function can be defined as:

$$\pi(\chi) = \begin{cases} p - c_v - c_i\overline{\chi} & \text{if } c_v \leq p \\ -c_i\overline{\chi} & \text{else} \end{cases} \tag{3.18}$$

Like in Chapter 3.3, we assume the regulator to be risk neutral. Hence, she maximises the expected social welfare. As firms only produce if the carbon price exceeds the variable costs, welfare in t_4 is given by:

$$\mathcal{W} = \begin{cases} \int_p^\infty Q(z)dz + (p - \hat{d})Q(p) + \int_0^{\overline{\chi}}(d - c_v - c_i z)dz, & \text{if } c_v \leq p \\ \int_p^\infty Q(z)dz + (p - \hat{d})Q(p) - \int_0^{\overline{\chi}}(c_i z)dz, & \text{else} \end{cases} \tag{3.19}$$

[46]Supplementary Material B.5 shows that all carbon pricing regimes yield the social optimum if risk stems from variable costs and production is potentially welfare reducing.

[47]In Chapter 3.3.2, the realised carbon price by assumption is higher than the marginal costs of production, such that firms produce for any realisation of damage and costs.

3.4.2. Policy ranking with damage risk

This section analytically assess the different carbon pricing regimes when the emission-free production is potentially welfare reducing in a setting with damage risk and risk-neutral firms. We derive the outcomes of the three carbon pricing regimes regarding emission-free production capacity $\overline{\chi}$ and carbon price p. We contrast the three regimes to the social optimum and conclude that

Proposition 3.4.1. *In the presence of damage risk, potentially welfare-reducing production and risk-neutral firms, only setting a carbon price flexibly yield a socially optimal level of* p *and* $\overline{\chi}$. *Offering a CCfD or committing to a carbon price falls short of the social optimum, as these regimes safeguard emission-free production even if it is ex-post socially not optimal.*

Under *Regulatory Flexibility*, the regulator can react flexibly to the actual environmental damage and sets the socially optimal Pigouvian tax level. Concurrently, as firms are risk neutral, investments are not hampered by the risk in profits. Hence, in *Flex*, the emission-free production capacity is socially optimal. In contrast, if the regulator offers a CCfD or commits to a carbon price, the firms' production decision is independent of the actual environmental damage. Hence, these regimes safeguard emission-free production even if it is ex-post socially not optimal. Although the regulator anticipates this effect and, in the *CCfD* regime, lowers the strike price, she cannot reach the social optimum. In addition to the welfare-reducing production level, committing to a carbon price early on also sets the carbon price for consumers, which is ex-post socially not optimal. As in the previous chapter, this socially not optimal carbon price level additionally lowers welfare.

Proof. For the proof of proposition 3.4.1, we compare the socially optimal carbon price and the emission-free production capacity to the three carbon pricing regimes. Supplementary Material B.4 presents a complete derivation of the respective optimal solutions. In the following, we provide the main results and the intuition behind the finding in proposition 3.4.1. ∎

Social optimum

In t_3, the social planner sets the carbon price p^{Opt} when the level of damage revealed. She optimises Equation 3.19, anticipating that her choice of the carbon price impacts the production of the emission-free technology. Irrespective of the production decision, the social planner sets the carbon price equal to the actual environmental damage, i.e., the Pigouvian tax level $p^{Opt} = \hat{d}$. Hence, whether firms that invested in the emission-free technology in t_2 produce in t_4 or not depends on the realisation of marginal environmental damage.

In t_2, the social planner sets the emission-free production capacity $\overline{\chi}^{Opt}$ to maximise expected welfare. She considers the cases in which production of the emission-free technology may not be socially optimal, i.e., $c_v > \hat{d}$. Thereby, she knows that irrespective of the investment decision, firms will only produce if the realised damage is greater than the marginal variable costs of abatement. In the social optimum, she sets the emission-free production capacity to:

$$\overline{\chi}^{Opt} = \frac{\int_{c_v}^{\infty}(z - c_v) f_D(z) dz}{c_i} \tag{3.20}$$

The solution balances the expected benefit of abatement with its investment costs. The expected benefit of abatement is equal to the benefit from reduced environmental damage minus variable costs weighted by its probability of realisation represented by the integral over the distribution function. The integral is limited to c_v as there is no emission-free production for $c_v > \hat{d}$.

Regulatory flexibility

Under *Regulatory Flexibility*, the regulator sets the carbon price after the actual damage revealed. Hence, in t_3, the regulator faces the same optimisation problem as the social planner, such that $p^{Flex} = p^{Opt} = \hat{d}$.

Sunk investment costs from t_2 or whether the emission-free technology produces or not in t_4 are irrelevant for the regulator's decision.

In t_2, firms choose to invest if their expected utility is positive, anticipating that the Pigouvian carbon tax depends on the damage level that is not yet revealed.

The firms anticipate that they will only produce if the damage (and the respective carbon price) is large enough, i.e., $c_v \leq \hat{d}$. Thereby, the marginal firm investing in the emission-free technology is defined by

$$\overline{\chi}^{Flex} = \frac{\int_{c_v}^{\infty}(z - c_v) f_D(z) dz}{c_i} \tag{3.21}$$

In the absence of risk aversion, the investment rationales of firms and the social planner are aligned, such that *Flex* reaches the social optimum. This result extends the findings from Chapters 3.3.2 and 3.3.3 with $\lambda = 0$ to the case in which emission-free production can be ex-post welfare reducing.

Commitment

Under *Commitment*, firms choose to invest in the emission-free technology in t_2 given the announced carbon price level. The investment decisions are identical to those under *Regulatory Flexibility*, only that the firms know the carbon price when making their decision. Hence, the marginal firm investing in the emission-free technology is characterised by

$$\overline{\chi}^{Com} = \begin{cases} \frac{p^{Com}-c_v}{c_i} & \text{for } c_v \leq p \\ 0 & \text{else} \end{cases} \tag{3.22}$$

In t_1, the regulator sets the carbon price anticipating that her choice impacts the firms' investment decision. She chooses a carbon price equal to the expected environmental damage, i.e., $p^{Com} = \mu_D$. As in Chapter 3.3.2 the carbon price is either too high or too low. By assumption, the expected damage is greater than the variable costs, i.e., $\mu_D > c_v$, which implies that investments and production occur. In cases where $\hat{d} < c_v$, the emission-free technology should not produce but does so in response to a too high carbon price. Furthermore, plugging in p^{Com} in Equation 3.22 and subtracting the socially optimal investment level shows that the investment level under *Com* falls short of the social optimum:

$$\begin{aligned} \overline{\chi}^{Com} - \overline{\chi}^{Opt} &= \frac{\int_{-\infty}^{\infty}(z-c_v)f_D(z)dz}{c_i} - \frac{\int_{c_v}^{\infty}(z-c_v)f_D(z)dz}{c_i} \\ &= \frac{\int_{-\infty}^{c_v}(z-c_v)f_D(z)dz}{c_i} \\ &\leq 0 \end{aligned} \tag{3.23}$$

This result shows that the regulator incentivises less investments than socially optimal in order to limit the welfare loss arising from potentially welfare-reducing production.

CCfD

When the regulator offers a CCfD in t_1, the optimisation rationale in t_3 is the same as in the social optimum and under *Regulatory Flexibility* (compare Equation 3.19). The solution yields the socially optimal Pigouvian tax level

$$p^{CCfD} = p^{Opt} = p^{Flex} = \hat{d} \tag{3.24}$$

In t_2, the investment decision of firms is identical to the rationale under the other regimes and hence:

$$\overline{\chi}^{CCfD} = \begin{cases} \frac{p_s-c_v}{c_i}, & \text{for } c_v \leq p_s \\ 0, & \text{else} \end{cases} \tag{3.25}$$

If the strike price, i.e., the firms' marginal revenue, is larger than their variable costs, they invest in the emission-free technology. Otherwise, it is not worthwhile for firms to enter a CCfD and invest.

In t_1, the regulator chooses a strike price that maximises social welfare. She accounts for the firms' reaction to the strike price.

$$p_s = \begin{cases} \mu_D, & \text{for } c_v \leq \mu_D \\ 0 \leq p_s < c_v, & \text{else} \end{cases} \tag{3.26}$$

By assumption $\mu_D > c_v$ holds. Hence, only the first case materialises, and the regulator offers a CCfD that incentivises investments and production. The resulting emission-free production capacity and production coincide with the one under *Commitment*. Hence, socially not optimal production occurs in those cases were $\hat{d} < c_v$. Furthermore, less investments than socially optimal are incentivised ($\overline{\chi}^{CCfD} = \overline{\chi}^{Com} = \frac{\mu_D - c_v}{c_i} < \overline{\chi}^{Opt}$) in order to limit the negative welfare effects of socially not optimal production.

Welfare comparison

We now compare the welfare of the three carbon pricing regimes in a setting of damage risk, risk-neutral firms, and potentially welfare-reducing emission-free production. *Regulatory Flexibility* yields both the socially optimal emission-free production capacity and carbon price. Under the *CCfD* regime, the carbon price is socially optimal, but too few firms invest in the emission-free technology. *Commitment* falls equally short of the socially optimal investment level. In addition, it achieves a lower consumer surplus due to a sub-optimal carbon price. Hence we derive the ranking:

$$E[\mathcal{W}_{\sigma_D}^{Opt}] = E[\mathcal{W}_{\sigma_D}^{Flex}] \geq E[\mathcal{W}_{\sigma_D}^{CCfD}] \geq E[\mathcal{W}_{\sigma_D}^{Com}] \tag{3.27}$$

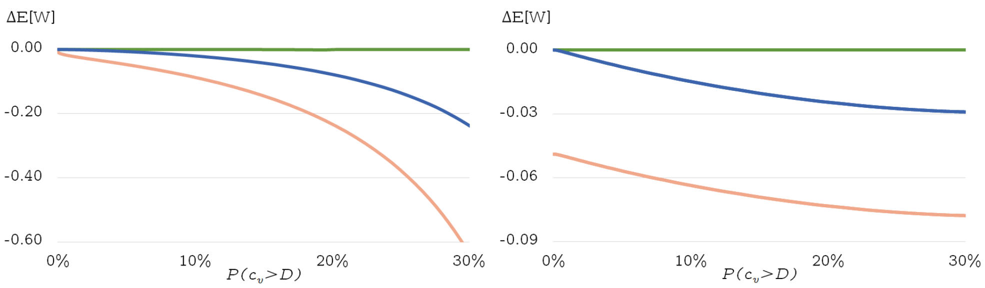
First, we find that *Regulatory Flexibility* reaches the social optimum. The firms face a carbon price equal to the marginal environmental damage and, thus, their production decision is socially optimal. Concurrently, as the firms are risk neutral, volatile profits do not impede investments.

Second, welfare falls short of the social optimum if the regulator offers a CCfD. Firms' production decision is independent of the actual carbon damage, such that emission-free production is safeguarded even if it is ex-post socially not optimal. We find that with an increasing probability of ex-post welfare-reducing production, welfare increasingly falls short of the social optimum. The probability of situations in which emission-free production is socially not optimal depends both on the variance (σ_D) and the expected value (μ_D) of the environmental damage. However, the impact of these two factors differs. As the expected value of environmental damage decreases, the welfare-deterring effect of the *CCfD* regime is partially mitigated as the socially optimal emission-free production capacity decreases, too. Figure 3.7 illustrates these findings for a numerical example.[48]

[48]These parameter values are illustrative and do not correspond to empirical estimates. Both, Figure 3.7i and Figure 3.7ii, share the parameters regrading the demand $Q(p) = 5 - 0.4p$ and the costs related to the emission-free technology $c_v = 2$ and $c_i = 1$.

We provide an analytical proof showing the welfare superiority of *Regulatory Flexibility* compared to the *CCfD* regime in Supplementary Material B.4. Figure 3.7i presents welfare changes induced by an increase of the variance of the damage, σ_D, and Figure 3.7ii welfare changes induced by an increase of the mean of the environmental damage, μ_D.

Third, confirming the results of Habermacher and Lehmann (2020), we find that *Com* likewise falls short of the social optimum. Moreover, *Com* performs worse than offering a CCfD. In addition to the welfare-reducing production, committing to a carbon price early on does not only affect producers but also consumers. Suppose the probability of socially not optimal production increases due to an increase of the damage variance, both the production and the consumption decisions are increasingly distorted. As a result, the welfare deterring effect in comparison to the *CCfD* regime increases. In turn, if the probability of socially not optimal production increases due to a reduced difference between μ_D and c_v, the shortfall in welfare is unaffected. We depict these results in Figure 3.7.



(i) Change in $P(c_v > D)$ due to an increase in σ_D

(ii) Change in $P(c_v > D)$ due to a decrease in μ_D

Flex   *Commit*   *CCfD*

Figure specific parameters in (i): $D \sim N(\mu_D = 2.75, \sigma_D^2 \in [0, 1.5]$ and (ii): $D \sim N(\mu_D \in [2.25, 3.5]), \sigma_D^2 \in (0, 1.5])$.

Figure 3.7.: Difference in welfare compared to social optimum in the presence of damage risk and potentially welfare-reducing production.

3.4.3. Numerical application with risk aversion

We complement our analytical results with a numerical application. The primary intention of this numerical exercise is to show how firms' risk aversion alters the effect of potentially welfare-reducing production in case of damage risk. Like in Chapter 3.3, we assume the firms to have a utility which is exponential in profits (i.e., $EU[\pi(\chi)] = E[1 - e^{\pi(\chi)}]$. We find that the introduction of risk aversion reduces the superiority of *Regulatory Flexibility* and generates a trade-off for the regulator between incentivising investments and triggering socially optimal production. Note that these parameter values are illustrative and do not correspond to empirical estimates.[49] For the analysis, we vary two parameters in our model: firms' risk aversion and the distribution of the environmental damage. The latter results in different probabilities of socially not optimal production, i.e., how likely it is that variable costs of abatement are ex-post higher than the marginal environmental damage.

To illustrate the effects of these two variations, we calculate the expected welfare levels of the carbon pricing regimes and compare them to the social optimum. Figure 3.8 depicts the results. In Figure 3.8i, we analyse the impact of firms' risk aversion. Extending our analytical results for the case without risk aversion, *Commitment* and *CCfD* do not result in the social optimum, whereby the *CCfD* regime is superior to *Com*, as it sets the socially optimal carbon price. Firms' risk aversion does not impact the welfare levels as both regimes remove risk for the firms. Also reflecting the results of Chapter 3.4.2, the *Flex* regime results in the social optimum if firms are risk neutral. However, as the risk aversion increases, fewer firms invest in the emission-free technology, whereby the expected welfare of this policy regime decreases. If this investment hampering effect of risk aversion becomes sufficiently large, the *Flex* regime becomes welfare inferior to *Com* and *CCfD*. Hence, there is a trade-off between the effects identified in Chapter 3.3.2 and 3.4.2.

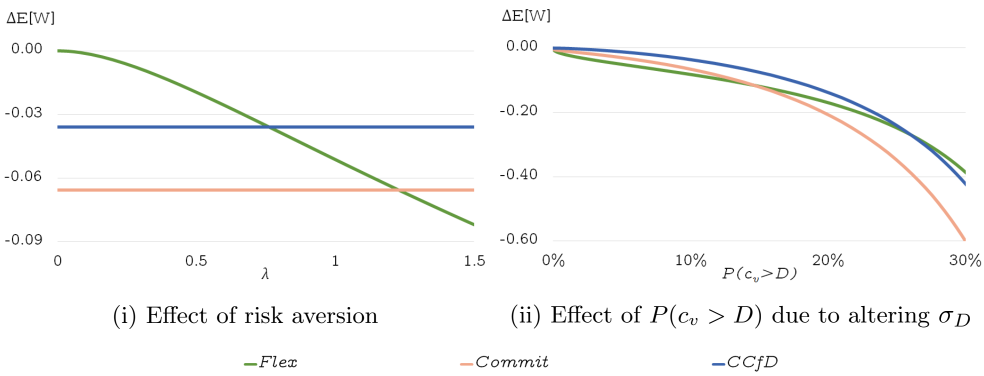
Figure 3.8ii shows a similar effect when varying the probability of socially not optimal production by altering the variance of the marginal damage as $P(C_v > D)$ increases in σ_D.[50]

With increasing volatility, *Flex* becomes less efficient as firms' risk aversion increasingly impedes investments. Offering a CCfD and committing to a carbon price, in contrast, become less efficient due to the increasing probability of welfare-reducing production arising from increased volatility. The level of risk aversion does not impact this effect. Under *Com*, the ex-post socially not optimal carbon price also applies for consumers, such that welfare is lower than in the *CCfD* regime. With an increasing probability of socially not optimal production, the welfare-deterring effect of *CCfD* and *Com* becomes more pronounced compared to the *Flex* regime. Hence, with an increasing probability of welfare-reducing production, the *Flex* regime becomes welfare superior to *Com* and *CCfD*.[51]

[49] Figure 3.8i and Figure 3.8ii share the parameters regarding the demand $Q(p) = 5 - 0.1$ and the costs related to the emission-free technology $c_v = 4$ and $c_i = 1$.

[50] In this illustrative example, all carbon pricing regimes achieve the social optimum at $P(C_v > D) = 0$. This is only the case because $\sigma_D = 0$ holds as well.

[51] When changes in the probability of socially not optimal production stem from decreasing the difference between μ_D and c_v, similar effects occur (see Supplementary Material B.6).



(i) Effect of risk aversion

(ii) Effect of $P(c_v > D)$ due to altering σ_D

Figure specific parameters in (a): $\lambda \in [0, 1.5]$, $D \sim N(\mu_D = 2.75, \sigma_D^2 = 0.7803)$ such that $P(c_v > D) = 10\%$ and (b): $\lambda = 1.5$, $D \sim N(\mu_D = 5, \sigma_D^2 \in (0, 2])$.

Figure 3.8.: Difference in welfare compared to social optimum in the presence of damage risk, potentially welfare-reducing production and risk aversion.

Both numerical simulations show that the superiority of the respective carbon price regime is ambiguous and depends on specific parameters. However, if the regulator had to choose between offering a CCfD and committing to a carbon price early on, i.e., before the risk resolves, it is always beneficial to provide a CCfD.

3.5. Discussion

In the previous chapters, we showed under which circumstances offering a CCfD can be a valuable policy measure. CCfDs could increase welfare compared to a flexible carbon price if the regulator expects that, first, firms will significantly under-invest in an abatement technology in the presence of risk and, second, the probability of welfare-reducing emission-free production is low. In other words, a CCfD is only beneficial if the benefit from the additional abatement that it incentivises outweighs the risk that it supports a technology that is socially not optimal.

There are several considerations beyond our model setup determining whether a CCfD is an efficient policy instrument. First, it matters who can enter a CCfD. While policy constraints may imply that a regulator should offer CCfDs only to limited sectors, for instance, heavy industry, our research indicates that they may be helpful in a broader range of settings in which agents make insufficient investments for decarbonisation because of the presence of risk. Second, the variance of the variable at risk may increase with a longer duration of the CCfD. Hence, the probability of supporting an ex-post welfare-reducing technology may increase with the duration. Third, the process of how the regulator grants a CCfD determines its impact on welfare. Suppose the CCfD only addresses the risk regarding the valuation of damage. In that case, the strike price should equal the regulator's damage expectation, and she can offer the CCfD to any interested party. If, however, the regulator aims to address private information, for instance, on the expected variable costs or firms' risk aversion, an auction process may be preferable to minimise costs for the regulator. Likewise, this holds if the CCfD involves an additional subsidy.

In addition to the carbon price risk, the regulator may introduce an instrument, similar to a CCfD, that assumes risks on the firms' variable costs. For instance, the proposal of the German funding guidelines for large-scale decarbonisation investments in the industrial sector includes such an extended risk assumption by the government (BMU, 2021). The extended risk-bearing could reduce complementary investment subsidies from the regulator to risk-averse firms, as shown by Richstein et al. (2021).[52] However, the regulator would safeguard firms in situations with ex-post socially not optimal production, i.e., unexpectedly high variable costs which exceed marginal damage. Thereby, the probability of financing an ex-post socially not optimal technology would increase, decreasing welfare. This measure would need a reasonable justification, for instance, a significant level of firms' risk aversion or a sufficiently low probability that the low-carbon technology is socially not optimal.

Our research relies on several assumptions that, if relaxed, might dampen the identified effects and potentially change the policy rankings. Noteworthy, we assume the absence of shadow cost of public funding. Because taxation has distortionary effects, public expenses might come at a cost (e.g. Ballard and Fullerton, 1992, for a review). Including shadow costs of public funds into our model might yield two effects. First, the carbon price would optimally be higher than the marginal environmental damage. The regulator would value one unit of revenue from the carbon price at more than one unit of consumer surplus because it allows other distortionary taxes to be reduced (see, e.g., Helm et al., 2003, for a discussion of this *weak form* of a double-dividend). Second, offering a CCfD would be more costly, and the regulator might require a premium for providing the contract and safeguarding the investments. If this is the case, the benefits of offering a CCfD would partially diminish. We expect a trade-off between the benefit of increased investments and the costs of additional public funds when comparing a *CCfD* regime with *Regulatory Flexibility* and *Commitment*.

Similarly, the regulator may also be risk averse. In this case, we can see the three carbon pricing regimes from the angle of who bears the risk (see Hepburn, 2006, for a discussion of risk-sharing between the government and the private sector). While the risk remains with the firms under *Regulatory Flexibility*, the regulator assumes the risk under *Commitment* and *CCfD*. Suppose a risk-averse regulator bears the risk in the presence of an unknown valuation of environmental damage. To reduce the negative welfare effects in case of great environmental

[52] In our model, e.g., in Chapter 3.3.3, such a scheme would lower the average strike price to the expected damage and reduce the average spending of the regulator.

damage, she would set a higher strike price when offering a CCfD or increase the carbon price under *Commitment*. In contrast, with variable cost risk, she prefers incentivising a lower level of investment to reduce her risk. This aspect may change the policy ranking of the three carbon pricing regimes.

We analyse a setting where carbon prices determined by the marginal environmental damage result in a demand that exceeds the optimal emission-free production capacity. However, we could think of settings, in which demand can be covered entirely by the emission-free production. In these settings, the conventional technology would not produce. Hence, the marginal utility of consumption, given the production capacity of the emission-free technology, would determine the product price. In consequence, if firms would assume the product price to be set by the conventional technology, some of the firms using the emission-free technology would incur a loss. Instead, firms would anticipate a product price below the carbon price and reduce their investment. The marginal firm would avoid a loss by balancing its investment costs with the contribution margin, which is reduced to lower prices. If the firm cannot pass through its investment costs, it would not invest in the first place. The model would not have an equilibrium.

Broadly speaking, if the regulator aims to fully replace the conventional technology, offering a CCfD is not an adequate policy. The instrument implicitly assumes that the profit of the emission-free technology is linked to the carbon price. This is only the case if the conventional technology sets the market price because the emission-free technology is not subject to the carbon price. For the same reason, CCfDs can only support a technology switch in an existing product market but not the market ramp up for a new product.

Our model results focus on the effects of each type of risk separately. In reality, stakeholders likely face damage and cost risk simultaneously. If the two risks are uncorrelated, their effects are additive. Variable cost risk can lead to an investment that is too low. Damage risk can affect both investment and consumption. Hence, the welfare ranking in Equation 3.10 holds and the superiority of Commitment or Regulatory Flexibility depends on the concrete circumstances. If risks are positively correlated, high environmental damage indicates high variable costs and vice versa. In this case, the emission-free production is likely to be ex-post socially optimal as $\mu_{C_V} > \mu_D$ holds. Results are then similar to the setting in Chapter 3.3. If risks are negatively correlated, high environmental damage indicates low variable costs and vice versa. In the case of high damage and low variable costs, emission-free production is socially optimal. In the case of low damage and high variable costs, in turn, the emission-free production is likely to be welfare reducing. Hence, if risks are negatively correlated, the situation is similar to the setting in Chapter 3.4.

The last simplification of our model we like to stress is the assumption of constant marginal environmental damage. We do not expect our main findings regarding the ranking of the carbon pricing regimes to change if we alleviate this assumption. If the marginal environmental damage was non-constant, the regulator would still choose the Pigouvian tax level after the firms have invested. In contrast to our assumption, the tax level would depend on the number of firms using the emission-free technology, i.e., total emissions. If markets are competitive, the impact of an individual firm on total emissions is negligible, and firms' investment decisions would not change compared to our model.

3.6. Conclusion

The decarbonisation of the industry sector requires large-scale irreversible investments. However, the profitability of such investments is subject to risk, as both, the underlying revenue and the associated costs of switching to an emission-free production process, are unknown and cannot be sufficiently hedged. The European Commission's Hydrogen Strategy and the *Fit for 55* package propose Carbon Contracts for Differences (CCfDs) to support firms facing large-scale investment decisions. Such contracts effectively form a hedging instrument to reduce the firms' risks.

With this research, we contribute to the understanding of how regulators should design this instrument and under which circumstances it is beneficial to offer a CCfD. We analyse the effects of a CCfD in the presence of risks stemming from environmental damage and variable costs on the decisions of a regulator and risk-averse firms facing an irreversible investment decision. Applying an analytical model, we compare three carbon price regimes against the social optimum: *Regulatory Flexibility*, *Commitment*, and offering a *CCfD*.

We conclude that a CCfD can be a welfare-enhancing policy instrument, as it encourages investments when firms' risk aversion would otherwise impede them. Additionally, offering a CCfD is always better than committing early to a carbon price as CCfDs incentivise investments in the same way while keeping the possibility to set the carbon price flexibly if new information,e.g., on the environmental damage, is available. However, if it is likely that the production of the emission-free technology turns out to be socially not optimal, CCfDs have the disadvantage that the regulator is locked in her decision, and she may distort the market clearing. In these situations, *Regulatory Flexibility* can be welfare superior to offering a CCfD. The comparison of *Regulatory Flexibility* and *Commitment* depends on the type of risk involved. With damage risk, *Regulatory Flexibility* is superior to *Commitment* if the level of risk aversion is low and the elasticity of demand is high. With variable cost risk, in contrast, *Regulatory Flexibility* performs worse than *Commitment*. While the regulator can only set the carbon price after the firm's investment under *Regulatory Flexibility*, she can balance additional investment incentives and the consumption level under *Commitment*.

This research focuses on the effects of CCfDs, aiming at mitigating the impact of risk regarding investments in emission-free technologies. Further research analysing CCfDs with more complex features and the interactions between CCfDs and other policy instruments may broaden our understanding of this instrument. To begin with, regulators may combine a CCfD with a subsidy payment to firms. This combination may be justified if the future carbon price is too low to incentivise sufficient emission-free investments, e.g., in the presence of learning effects or other positive externalities. Research could focus on whether combining a CCfD and a subsidy has advantages over offering both instruments separately. Additionally, proposals for the use of CCfDs focus on sectors competing in international markets. Our model assumes complete cost pass-through of the carbon price and, hence, increased revenues for firms investing in abatement. If not all firms on an international market face a (similar) carbon price, this may not hold. It remains open how the design of CCfDs would need to change in such settings to ensure investments' profitability. Future analyses could consider the possibility of introducing carbon border adjustment mechanisms, such that producers from countries without a carbon price at the domestic level cannot offer the goods at a lower price. The question how other hedging instruments offered by private actors compare to CCfDs is also worth analysing in more detail. Moreover, future research could assess the role of shadow costs of public funds by extending our model in this regard. As pointed out in Chapter 3.5, we assume payments under a CCfD to be welfare-neutral. Considering shadow costs of public funds may worsen the welfare ranking of CCfDs compared to pure carbon pricing regimes.

4. Pricing short-term gas transmission capacity: A theoretical approach to understand the diverse effects of the multiplier system

4.1. Introduction

Efficient operation of gas transmission networks is crucial for the gas supply system and overall welfare. Due to the direct effect on network utilisation and the resulting welfare, the applied pricing policy for financing of networks is particularly important. Principles of microeconomics indicate that economic efficiency is maximised when prices reflect short-run marginal costs (Borenstein, 2016). However, the existence of high fixed costs in gas networks necessitates charging tariffs higher than short-run marginal costs so that revenues cover the total network costs.[53] The networks are dimensioned according to maximum (i.e. peak) capacity demand, which in turn largely determines the fixed costs. An important issue when designing the tariff structures then becomes how to charge the network users for the cost of capacity. A common approach for financing networks is to apply capacity tariffs used to distribute the network costs among users depending on their peak capacity demand. As such, in contrast to a pure commodity tariff[54] regime where only the transported volumes are charged, capacity tariffs[55] incentivise the reduction of yearly peak capacity demand and potentially reduce the need for capacity extensions.

Financing of gas networks in the EU occurs via the entry-exit regime. Operated by transmission system operators (TSOs), the EU gas grid consists of numerous regional gas transmission networks (i.e. market areas) which connect producers and neighbouring networks with storage facilities (henceforth storages) and downstream distribution networks. In this context, the entry-exit system requires network users to book entry and exit capacities in explicit auctions whenever transporting gas into or out of a certain market area, paying the corresponding

[53]This is also observed in other natural monopolies such as telecommunication, electricity and railway networks.
[54]Commodity tariffs are also commonly referred to as energy charges or volumetric charges.
[55]Capacity tariffs are also commonly referred to as capacity charges or demand charges.

tariffs.[56] When the entry-exit tariff system was first introduced in the EU with Regulation 2009/715, the offered capacities were limited to yearly capacities. This meant traders were not charged according to the actual transported gas volumes but rather for their expected peak capacity demand, which essentially corresponded to a pure capacity pricing regime. However, in some cases, offering only yearly capacities caused inefficient short-term utilisation of the existing pipelines, where significantly high price spreads between market areas occurred despite the absence of physical congestion (ENTSOG, 2017). This inefficiency was caused by arbitrageurs not being able to exploit short-term regional price spreads without procuring capacity covering a whole year.

In order to reduce the inefficiencies resulting from offering only yearly capacities, the EU Commission introduced the Network Code on Capacity Allocation Mechanisms (NC CAM) with Regulation 2013/984, extending the available capacity products to cover sub-annual durations. The regulation thus required TSOs to offer short-term (ST) transmission capacities, i.e. quarterly, monthly, daily and within-day capacities, while the previously introduced yearly capacities were defined as long-term (LT) capacities. Instead of the necessity to cover the yearly peak demand with a yearly product, capacities could now be booked according to the actual transmission demand. This enabled traders to make capacity bookings correspondingly to the actual transported volumes, similarly to what would occur under a commodity pricing regime. LT and ST capacities generally do not cost the same. According to EU regulations, ST capacities should be priced low enough to incentivise short-term trade but sufficiently high to support enough LT bookings to achieve stable TSO revenues and tariffs. In this context, in the EU, ST products are priced by multiplying the LT tariff with factors called multipliers. Those multipliers are individually specified by the respective national regulatory authorities (NRAs).[57]

By making ST products comparatively more expensive, NRAs can influence the emphasis of capacity vs. commodity pricing in the pricing of transmission capacities in the EU entry-exit tariff structure. This can be best illustrated with two extreme cases: If the multipliers were equal to 1, then the ST capacities would cost the same as LT capacity. In this case, any capacity booking pattern that

[56] The booking of capacities occurs in capacity auctions performed by trading platforms (such as PRISMA, GSA, RBP) in which the reserve prices correspond to the transmission tariffs. In a large share of the EU capacity auctions, demand for capacity remains below the offered capacity (ACER, 2019b). In the remaining cases where demand for capacity exceeds the offered capacity a congestion premium arises.

[57] When NC CAM came into force, multipliers largely varied among countries spanning a wide range from 1 to as high as 5.5 and mostly increased as the run-time of the capacity product decreased. The EU Commission tightened the rules regarding multipliers in their network code on tariff harmonisation (NC TAR) from the EU regulation 2017/460. The regulation limits the range for multipliers for member states to 1–3 from June 2019 onward. Moreover, the EU Agency for the Cooperation of Energy Regulators (ACER) has to decide by April 1st, 2021 whether multipliers are to be further restricted within a range of 1–1.5 starting from April 2023.

includes LT capacities can not be cheaper than booking solely ST capacities.[58] As a result, traders would only book a combination of ST capacities which exactly satisfies their demand profile for transmission capacity. In such a setting, network users behave as being exposed to commodity pricing since they pay for the exact amount of volumes, i.e. the energy they transport. Whereas, if the multipliers were sufficiently high, so that booking LT capacity would be always cheaper than booking ST capacities, then the traders would book only LT capacity. This would essentially result in network users behaving as being exposed to a pure capacity pricing regime, as traders would be required to book enough transmission capacity to cover their yearly peak demand even if their average capacity demand is lower; hence, resulting in them paying for the capacity rather than the energy.

The reality lies somewhere in between these two extreme cases. In a large majority of EU member countries, multipliers are greater than 1 but are still sufficiently low so that both LT and ST bookings are observed (ACER, 2019a). Hence, transmission network users in these countries are implicitly charged a combination of capacity and commodity tariffs. The extent to which aspect dominates over the other, and the ensuing effects on infrastructure and welfare, are determined by the multipliers and the underlying tariff structures—the analysis of which constitutes the focus of this paper.

The issue of how to design tariffs within the EU entry-exit framework has been analysed in the literature, where aspects such as cost recovery, cost distribution and efficiency have been considered. Bermúdez et al. (2016), analysing different methodologies of setting LT tariffs, argues that more cost-reflective methodologies ensure more efficient utilisation of the transmission network. Mosácula et al. (2019), however, points out that approaches which charge full costs at EU interconnectors are unlikely to maximise social welfare. This is also mentioned in Hecking (2015), which suggests to reduce inefficiencies by setting entry and exit tariffs equal to short-run marginal costs for interconnectors within the EU while applying sufficiently high tariffs at the EU outer borders to finance the EU transmission grid. In addition to increasing the efficiency of the gas dispatch, the study estimates that such a tariff regime would also allow to redistribute considerable share of network costs towards suppliers at the EU borders, indicating the relevance of tariff design on the distribution of network costs.

The pricing of LT vs. ST capacities and the topic of multipliers have not been analysed in the academic literature so far.[59] To our knowledge, a tariff framework similar to the current tariff structure of the EU gas transmission capacities is also not observed in any other regulated network neither in the EU nor in other regions, hence the lack of comparable literature. Nevertheless, when multipliers are larger than 1, the EU tariff structure has similarities with the concept of peak-load pricing. In peak-load pricing, higher prices are charged in

[58]It is assumed that no transaction costs exist and enough capacity products are offered.

[59]The topic is qualitatively addressed only in several consulting studies and technical reports (ACER, 2019a, ACER and CEER, 2019, DNV-GL, 2018, EY and REKK, 2018, Rüster et al., 2012, Strategy& and PwC, 2015).

peak periods than in off-peak periods. Similarly, in the EU entry-exit system, when ST capacity is more expensive than LT capacity, traders are incentivised to procure the cheaper LT capacity for meeting base load demand whilst procuring the more expensive ST capacity to meet their peak-load demand. This implicitly results in higher capacity costs for peak periods than for off-peak hours. The founding works of Boiteux (1949) and Steiner (1957) on peak-load pricing have shown that allocating the costs of capacity to peak-load consumers and charging them consequently higher tariffs impacts the networks utilisation and leads to higher long-term efficiency. Further, Gravelle (1976) and Nguyen (1976) indicate that the problem of peak-load pricing remains a valid issue even when storage (with significant costs) is available, which is undeniably the case in the majority of EU gas systems. These findings further underpin the relevance of analysing the effects of the multipliers on network utilisation, efficiency and cost distribution.

In order to improve the understanding of the effects of multipliers and fill the research gap in the literature, we develop a stylised theoretical framework with an analytic solution that depicts the gas procurement, storage, and transmission capacity booking in the EU gas market. The model considers two points in time and two nodes under a setting of perfect competition and perfect foresight. We solve the resulting linear cost minimisation problem analytically using Karush-Kuhn-Tucker (KKT) conditions, providing analyses on the effects of multipliers. The analysed aspects can be grouped into three main categories; the direct impact of multipliers on infrastructure utilisation, effects on prices at the two nodes (referred to as hub prices in the remainder of the text) and welfare implications.

Our model results show that high multipliers indeed reinforce the capacity pricing component and cause bookings to shift from ST capacities to LT capacity, resulting in increased storage utilisation. This leads to a more uniform usage of transport capacities, implying decreased volatility of pipeline transportation. The findings above are expected to be valid for the EU gas system in the majority of situations. Nevertheless, we find that these effects are not universal and depend strongly on whether the traders' capacity demand is elastic or not. We define the elasticity as the shift in capacity demand from the peak period to an off-peak period in response to an increase in the relative price of ST capacity (i.e. the multiplier). This elasticity largely results from gas storages, which provide the traders with inter-temporal flexibility, and give them the possibility of meeting their short-term needs with withdrawals from storages instead of booking ST capacities.

We find that certain proportions of multipliers with respect to the ratio of storage tariffs to transmission tariffs can lead to inelastic capacity demand: Multipliers that are sufficiently low (but still larger than 1) compared to the marginal cost of gas storage—or when no storage capacity exists—can result in a domain with inelastic capacity demand, where a change in multipliers does not affect the volume of booked capacities in the respective time periods. Similarly, we show that sufficiently high multipliers can lead to the same behaviour as in a pure capacity pricing regime, with only LT capacity being booked and the volume of booked capacity being independent of the multiplier level.

Regarding the impact of multipliers on temporal hub prices we identify several effects. We find that maximum regional price spreads increase with higher multipliers, an implication also mentioned by ACER (2019a). However, unlike ACER, who argues that ST capacity tariffs would act as reference prices for the regional spreads, we show that ST tariffs rather form the upper bounds for the spreads. As such, our results imply that the volatility in regional price spreads increases with higher multipliers. Further, we find that increases in multipliers can cause increased temporal volatility in hub prices if storage tariffs are comparably high or if storage capacity is unavailable.

The model results indicate that higher multipliers are associated with higher total system costs and consequently lower total welfare in the short-run. However, for the identified multiplier domain which is representative of the majority of the situations in the EU gas system, our results show that there exists a multiplier level potentially larger than 1, which maximises the total consumer surplus.

Therefore, despite the stylised setting, the implications of our model results are highly relevant for policymakers. Maximising total welfare requires the multiplier to be no greater than 1. However, policymakers, who aim to maximise consumer surplus, may favour a multiplier larger than 1, since transmission tariffs can be lowered by the TSOs, which leads to lower average hub prices. Multipliers higher than 1 also foster the redistribution of the network costs from base load towards peak-load consumers, in line with the principle of peak-load pricing.

The contribution of our paper can be summarised as follows: Academic literature on the effects of short-term transmission capacity multipliers is nonexistent. Hence, being the first of its kind, our paper aims to close this research gap. Thanks to the developed theoretical framework, direct effects and implications are identified within the valid tariff domains. Since our analysis shows that multipliers have significant effects on welfare, distinguishing between ranges of validity also helps support tailor-made policymaking.

4.2. The Model

We develop a theoretical model which depicts the procurement and the subsequent transmission capacity booking in the EU gas market. The model represents the relevant actors in a realistic manner, yet it is simplified enough to have a closed form solution. In this respect, the model considers two points in time (t_1, t_2), and five different groups of players interacting with each other: traders, producers, storage operators, the transmission system operator (TSO), and consumers. The structure of the model and the main assumptions for the considered agents are illustrated in Figure 4.1.

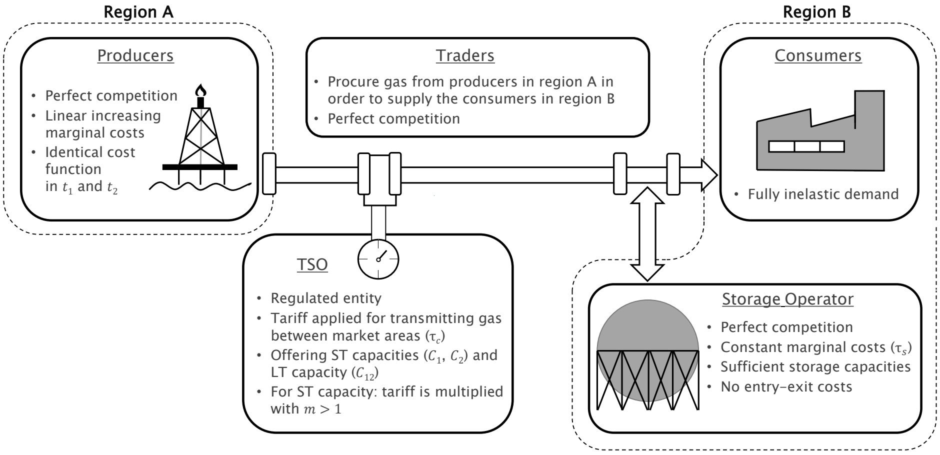


Figure 4.1.: Schematic representation of the model structure and the main assumptions

We assume that the traders are obliged to meet the gas demand of their customers (i.e. consumers) under a perfectly competitive market setting. Accordingly, traders procure gas from the gas producers located at market area A and transport it using the gas transmission network to the consumers which are located at market area B. In order to transport gas over the transmission network, traders need to book sufficient transmission capacities. Furthermore, the traders can store gas in gas storages in t_1 and withdraw it in t_2 to serve the gas demand in t_2. We assume that traders book capacities rationally and efficiently.[60]

We assume producers to face positive and linearly increasing marginal costs[61] and have sufficient capacities. Their aggregated cumulative cost function is linear and remains unchanged in both points in time. The producers are assumed to be under perfect competition and offer their gas at a rate that is equal to their marginal costs. This is in line with the simulations of Schulte and Weiser (2019), which indicate that gas suppliers to Europe behaved competitively in 2016.[62] The aggregated inverse supply function p_t of the producers can be then formulated as follows:

$$p_t(Q_t) = a + b\,Q_t \qquad \forall\, t \in (t_1, t_2) \tag{4.1}$$

where $Q_t > 0$ represents the aggregated gas procurement volumes of the traders.

The storage capacities of the storage operators are located in market area B where the consumers are located. We assume storage operators to face constant positive marginal costs under perfect competition. We further assume that the storage operators have sufficient capacities to meet the demand at all times and

[60]This is a realistic assumption also supported by the empirical analysis of Keller et al. (2019).

[61]Having carried out the analysis also by assuming a supply function with quadratic marginal costs, we find that the main findings regarding the effect of multipliers on gas dispatch remain unchanged. Hence, for the sake of clarity, we assume linear increasing marginal costs for producers in this paper.

[62]With increasing LNG supply and lower prices it can be safely assumed that gas markets have become even more competitive in recent years.

therefore offer their storage capacity at a rate equal to their marginal costs τ_s. This assumption is in line with the situation observed in the EU, where storage operators have been unbundled since the introduction of the third energy package (European Commission, 2010) and have ample storage capacities in the absence of supply disruptions (ACER, 2019a). Furthermore, we assume storage operators to be fully exempt from transmission tariffs when withdrawing or injecting gas in the transmission network.[63]

The consumers have a positive gas demand. The aggregated gas demand of the consumers at t_1 equals d_1. Similarly, the demand in t_2 is equal to d_2. Demand is assumed to be perfectly inelastic. This is a common assumption for stylised short-run gas market models and is also supported by the empirical analysis of Burke and Yang (2016), which finds that short-term elasticities for gas demand are generally low, and for the case of households, do not significantly differ from zero. Demand is assumed to be higher in the second period than in the first period, i.e. $d_2 > d_1 > 0$, representing a winter (d_2) and a summer period (d_1). To be able to examine distributional effects among different consumer groups we assume the aggregated consumer demand (i.e. d_1 and d_2) to be split into two demand groups: first, the demand of the base-load consumers (e.g. industry companies) which equals d_1 in both periods, and second, the demand of the peak-load consumers (e.g. households) which only occurs in t_2 and equals $(d_2 - d_1)$.

The TSO operates a transmission grid which connects the producers in market area A with the storages and consumers in market area B. The TSO is a regulated entity which is allowed to apply a tariff for transmitting gas between the two market areas. As in the case of the EU, the TSO offers LT and ST transmission capacity. The LT capacity product (C_{12}) covers both periods and the ST capacity products cover only a single period (i.e. C_1 in t_1 and C_2 in t_2). Traders need to book sufficient transmission capacity rights such that desired gas volumes can be transported to the costumers and the storages in market area B. Similarly to the EU with the regulation NC CAM, traders in our model are permitted to trade booked capacities in secondary capacity markets. As a consequence, in the given setting of perfect foresight, the sum of bookings of many individual traders would be identical to the booking of a single competitive trader who faces the cumulative demand of these many traders.[64] Since $d_2 > d_1$ and production costs are represented by a quadratic function of production volumes, it is inherently assumed that injection to storages occurs in t_1 and withdrawal occurs in t_2 to meet the higher demand. Zero storage losses are assumed; injection and withdrawal rates

[63]Such an exemption is observed in several EU countries (e.g. Spain, Denmark and Austria) with the goal of inducing positive externalities such as reducing pipeline investment costs and increasing security of supply (ACER, 2019a). In other countries, storages are exempted by at least 50% due to NC TAR regulation; though, most countries apply higher exemptions (ENTSOG, 2019).

[64]Due to the assumptions of perfect competition with perfect foresight, as well as the availability of sufficient transmission capacities and an efficient secondary capacity market, the traders in our model have no incentive to block capacities, as over-booking causes additional costs without additional benefits.

in both periods are the same and equal the stored volumes S. Hence, the supply constraints, where demand in each period is satisfied with corresponding capacity bookings and storage utilisation, can be stated as in Equations 4.2 and 4.3.

$$C_{12} + C_1 \geq d_1 + S \tag{4.2}$$

$$C_{12} + C_2 \geq d_2 - S \tag{4.3}$$

The regulated tariff for a unit of LT capacity equals τ_c (with $\tau_c > 0$) per time period and is fixed for both periods. The total LT tariff which runs over both periods then becomes $2\,\tau_c$. The tariff for the ST capacity is similarly regulated and is set to $m\,\tau_c$. In reality, as regulated entities, TSOs set the entry-exit tariffs (corresponding to the LT tariff τ_c in our model) such that their expected revenues cover their costs, adjusting the tariffs each year as necessary.

In our main analysis, the effects of multipliers on the players' behaviour and welfare implications are derived analytically in a closed form. For that purpose, we keep τ_c fixed and assume τ_c to be sufficiently high such that the TSO covers its costs in a setting without multipliers ($m = 1$). Therefore, the TSO may generate additional surplus if multipliers are larger than 1 ($m > 1$). After having derived the equations describing the behaviour of the players, we analyse the effects of m when the transmission tariff is adjusted. This allows us to derive the effects of m in the more realistic setting where the TSO surplus is independent of m (see Chapter 4.3.4).

The model depicts a setting of perfect competition and consumers' demand is perfectly inelastic in the short-run. Hence, the optimal allocation under perfect competition is equivalent to the solution of the planner's problem of maximising welfare by minimising the total costs ($Cost^{Tot}$). Since the total costs are the sum of production costs ($Cost^{Pro}$), transportation costs ($Cost^{Tra}$), and storage costs ($Cost^{Sto}$), the minimisation problem can be expressed as follows:

$$\min Cost^{Tot} = Cost^{Pro} + Cost^{Tra} + Cost^{Sto} \tag{4.4}$$

The production costs correspond to the integral of the price function $p_t(Q_t)$ with respect to production quantity Q_t:

$$\begin{aligned} Cost^{Pro} &= \int p_t(Q_t)\, dQ_t \\ &= Q_t\,(a + \frac{1}{2}\, b\, Q_t) \end{aligned} \tag{4.5}$$

The aggregated gas procurement Q_t is equal to $Q_1 = d_1 + S$ in t_1 and $Q_2 = d_2 - S$ in t_2. Substituting these into Equation 4.5, total production costs are obtained.

$$Cost^{Pro} = a\,(d_1 + d_2) + \frac{1}{2}\left[b\,(d_1 + S)^2 + b\,(d_2 - S)^2\right] \tag{4.6}$$

The storage costs correspond to the product of the stored gas volume S and the tariff for storing, τ_s:

$$Cost^{Sto} = S\,\tau_s \tag{4.7}$$

The costs for purchasing the capacity rights for transmission is equal to:

$$Cost^{Tra} = \left[m\,(C_1 + C_2) + 2\,C_{12}\right]\tau_c \tag{4.8}$$

Hence, the minimisation problem can be expressed as in Equation 4.9, subject to the constraints that demand needs to be satisfied in both periods and the non-negativity constraints discussed previously.

$$\begin{aligned} \min_{S,\,C_1,\,C_2,\,C_{12}} \quad Cost^{Tot} &= a\,(d_1 + d_2) + \frac{1}{2}\left[b\,(d_1 + S)^2 + b\,(d_2 - S)^2\right] \\ &+ \left[m\,(C_1 + C_2) + 2\,C_{12}\right]\tau_c \\ &+ S\,\tau_s \\ s.t. \quad & C_{12} + C_1 \ \geq\ d_1 + S \\ & C_{12} + C_2 \ \geq\ d_2 - S \\ & C_{12},\ C_1,\ C_2,\ S \ \geq\ 0 \end{aligned} \tag{4.9}$$

Assigning Lagrange multipliers $(\mu_1, \mu_2 ..., \mu_6)$ to the inequality constraints, the Lagrangian of the optimisation problem and the corresponding KKT conditions are obtained. The Lagrangian formulation and the KKT conditions can be found in Chapter C.1.

4.3. Results

4.3.1. Deriving the Effects on Infrastructure Utilisation

In this section, the solutions of the cost minimisation problem illustrated above are presented. We solve this convex optimisation problem by deriving the KKT conditions and finding the feasible KKT points, which provide us with analytic expressions of the analysed variables. Since the problem fulfils Slater's condition, the analysed KKT points are the optimal solutions of the optimisation problem.[65] As the effects of multipliers largely depend on whether they emphasise the commodity or the capacity pricing aspect, we divide our analysis into two subsections. The cases which, by design, correspond to a pure commodity pricing or conversely to pure capacity regime are considered separately from the cases that occur under a mixed-pricing policy—which are more common in reality and comprise more complex effects.

[65]To ensure that no optimal solution is omitted, an extensive analysis of all the possible cases including the non-optimal points are presented in Chapter C.2.

Pure commodity pricing ($m \leq 1$) or pure capacity pricing ($m \geq 2$)

As multipliers determine the relative price of ST capacities with respect to LT capacity, the outcomes of a pure commodity or capacity pricing regime can arise depending on the level of multipliers. For the case of our two-period model, these instances are shown in Proposition 4.3.1.

Proposition 4.3.1. *Multipliers $m \leq 1$ correspond to a pure commodity pricing regime, whereas multipliers $m \geq 2$ correspond to a pure capacity pricing regime.*

Proof. If $m \leq 1$, there exists no demand pattern where booking LT capacity is cheaper than booking ST capacity products. Therefore, the LT product is ignored and only ST capacities are booked. This corresponds to traders being charged for the actual transported volumes. Hence, the behaviour is the same as in a pure commodity pricing regime. If storage tariffs are sufficiently low ($\tau_s < 2b(d_2 - d_1)$), then traders also use storages to meet the demand in the peak period. Else ($\tau_s \geq 2b(d_2 - d_1)$), the demand is met only by booking the ST products at each period, where the transported volumes exactly correspond to the respective demand in each period (d_1 in t_1 and d_2 in t_2). See Chapter C.2 Case 1 (a) for the detailed proof.

If $m \geq 2$, there exists no demand pattern where booking ST capacities is cheaper than booking LT capacity. Hence, only the LT product is booked, inducing the same behaviour seen in a pure capacity pricing regime. Whether gas transmission is aligned between the periods or capacity rights are wasted depends on the ratio of storage tariff to transmission tariff levels: If the relative costs of storage with respect to transmission costs are sufficiently low ($\tau_s \leq 2\tau_c$), storage utilisation aligns transports completely such that the LT capacity is fully utilised. If the storage costs are comparatively high ($\tau_s > 2\tau_c$), the booked LT capacity in off-peak period is underutilised, i.e. some capacity is wasted: Under this condition, if $\tau_s < 2\tau_c + 2b(d_2 - d_1)$, storages align transports partially. In the case that $\tau_s \geq 2\tau_c + 2b(d_2 - d_1)$, storage utilisation is zero. See Chapter C.2 Case 4 (c) for the detailed proof. ■

For $m = 1$, traders' costs are the same as in a pure commodity tariff regime; namely, overall transported volumes determine the traders' transport costs. Further reductions in the multiplier do not change the optimisation rationale of the traders and welfare. For this reason, and since the EU regulation NC TAR 2017 also does not allow for multipliers below 1, the minimum multiplier value considered in the analysis of this paper is $m = 1$.

The multiplier threshold that corresponds to a pure capacity pricing regime equals to LT product duration expressed in terms of number of ST products. As our model has two time periods, this threshold is found to be equal to 2, as shown in Proposition 4.3.1. For such multipliers, we find that capacity wasting occurs if gas transports do not align in t_1 and t_2. Thereby, Proposition 4.3.1 implies that even in a market with perfect foresight, perfect competition, and secondary trading of capacity at no cost, some capacity rights may remain unused with high multipliers if capacity demand is inelastic due to comparatively high storage tariffs or when no storage capacities exist. Increasing multipliers above 2 does not affect the results, as traders do not procure ST capacity, where multipliers are applied. Hence, the highest multiplier considered in this paper is $m = 2$. In the EU, such multipliers, which by design correspond to pure capacity pricing, are ruled out with Regulation NC TAR 2017 as the EU aims to allow for and encourage ST capacity bookings.[66]

Mixed-pricing regime ($1 < m < 2$)

In most EU countries, the range of applied multipliers facilitates traders to consider both long-term and short-term bookings, allowing for an inherent mixed-pricing regime in which capacity and commodity pricing effects are simultaneously present. In our model, this range of multipliers corresponds to $1 < m < 2$.

In the following propositions we present how multipliers influence the capacity booking as well as storage decision and we relate the market outcomes to the regimes of capacity and commodity pricing. We identify specific thresholds for m that affect how changes in m influence the system. We define the lower threshold as $\underline{m}$ and the upper threshold as $\overline{m}$, which then constitute three domains. Despite the inherent mixed-pricing regime, we identify two domains ($m \leq \underline{m}$ and $m \geq \overline{m}$) where the capacity demand is inelastic due to underlying tariff structures. In these domains, the capacity demand in the off-peak and peak periods, and the proportion of LT to ST bookings, are independent of the multiplier. The third domain corresponds to the case with elastic capacity demand ($\underline{m} < m < \overline{m}$) which is representative of the majority of the actual situations observed in the EU gas system.

Proposition 4.3.2. *If $m \geq 1$, but sufficiently small ($m \leq \underline{m} = 1 + \frac{\tau_s}{2\,\tau_c} - \frac{b}{\tau_c}(d_2 - d_1)$) storages are not utilised, LT capacity is booked to cover the demand in t_1, and the remaining demand in the peak period t_2 is met with the ST product. The proportion of ST to LT bookings is independent of m. The capacity booking and storage volumes are:*

$$\begin{aligned} C_1 &= 0 \\ C_2 &= d_2 - d_1 \\ C_{12} &= d_1 \\ S &= 0 \end{aligned} \tag{4.10}$$

[66] The multiplier threshold in the actual EU tariff structure would be equal to 12 between the yearly and monthly products, for instance, or equal to 4 between the yearly and quarterly products. As multipliers are required to be below 3, feasible multipliers are sufficiently low to incentivise ST bookings when storage tariffs are low.

Proof. See Case 5 (a) i. in Supplementary Material C.2 for the proof. ■

Proposition 4.3.2 indicates that multipliers which are sufficiently low with respect to the ratio of storage to transmission tariffs can result in demand in peak periods to be exclusively met by ST capacities rather than storage withdrawals. The reason for that can be clearly seen by rewriting the $m \leq \underline{m}$ condition as $b\,(d_2 - d_1) + m\,\tau_c \leq \tau_c + \frac{\tau_s}{2}$. In this domain, meeting the additional demand in t_2 by procuring the additional volumes in t_2, and correspondingly booking ST capacity, is cheaper than the combined cost of booking LT capacity and storage utilisation. As a result, storages are not utilised and transported volumes in t_1 and t_2 exactly equal the demand d_1 and d_2. Hence, the capacity demand in the two periods remains independent of the multiplier; i.e. capacity demand is inelastic. Given that ratios of base transmission to storage tariffs allow for $m \leq \underline{m}$, network utilisation is the same as if pure commodity pricing ($m \leq 1$) is applied. This domain can appear in reality in the presence of low multipliers if storage tariffs are comparatively high or if no storage capacities exist.

Proposition 4.3.3. *If $m \leq 2$, but is sufficiently large ($m \geq \overline{m} = 1 + \frac{\tau_s}{2\tau_c}$), traders book LT capacity only and transport the same volumes in t_1 and t_2. The proportion of ST to LT bookings is independent of m. The capacity booking and storage volumes are:*

$$\begin{aligned} C_1 &= 0 \\ C_2 &= 0 \\ C_{12} &= \frac{d_2 + d_1}{2} \\ S &= \frac{d_2 - d_1}{2} \end{aligned} \tag{4.11}$$

Proof. See Case 4 (a) in Supplementary Material C.2 for the proof. ■

Proposition 4.3.3 shows that even in situations where m is set to levels, which theoretically allow for ST bookings in the optimum ($m < 2$), ST bookings may not necessarily be part of the optimal solution. This occurs when m is high in comparison to the ratio of storage to transmission tariff such that ST capacities cost more than the combined cost of LT capacity and storage. This can be clearly seen by rewriting the $m \geq \overline{m}$ condition as $m\,\tau_c \geq \tau_c + \frac{\tau_s}{2}$. As a result, the capacity demand is met by booking only LT capacity and using storages. Since transports in both periods align, and consequently there is no potential to shift capacity demand from the peak period to the off-peak period, capacity demand is inelastic. As traders do not procure ST capacity, market outcomes for such multipliers ($m \geq \overline{m}$) are the same as if no ST capacity would be offered; namely, as in a pure capacity pricing regime similar to the one that was in place in the EU before the introduction of NC CAM 2013.

Proposition 4.3.4. *If $1 \leq m \leq 2$ and $\underline{m} < m < \overline{m}$, the traders book LT capacity to cover the base load and ST capacity C_2 to cover the additional demand in the peak period (t_2). Traders utilise gas storages. The proportion of ST to LT bookings depends on m. The capacity booking and storage volumes are:*

$$\begin{aligned} C_1 &= 0 \\ C_2 &= \frac{\tau_s}{2b} - \frac{\tau_c(m-1)}{b} \\ C_{12} &= \frac{d_2+d_1}{2} - \frac{\tau_s}{4b} + \frac{\tau_c(m-1)}{2b} \\ S &= \frac{d_2-d_1}{2} - \frac{\tau_s}{4b} + \frac{\tau_c(m-1)}{2b} \end{aligned} \tag{4.12}$$

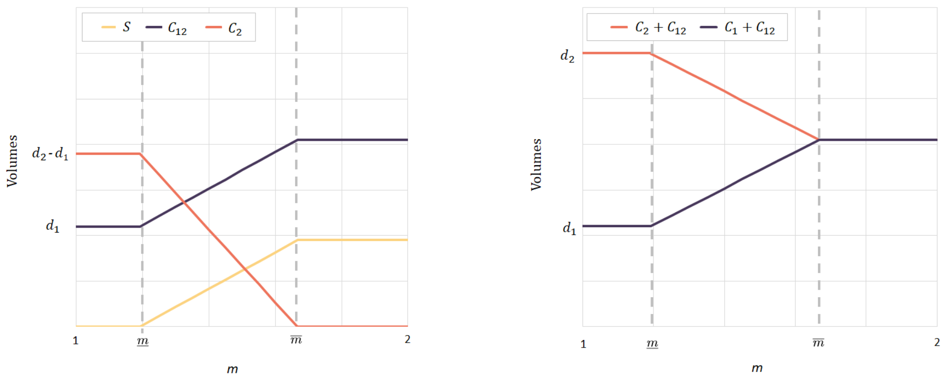
Proof. See Case 5 (a) ii. in Supplementary Material C.2 for the proof.
■

Proposition 4.3.4 shows the results for multipliers, which lie in the domain of moderate multipliers with respect to the ratio of storage to transmission tariffs. The results represent the only solution where the following three aspects occur simultaneously: Both LT and ST capacity are booked, and storages are utilised to satisfy the demand in the peak-period. This corresponds to a situation which can be observed in the EU for most countries. In this domain, the capacity demand is elastic since the capacity demand shifts from peak to off-peak period with increasing multipliers. With increasing m, ST capacity bookings are replaced with LT capacity booking and storage withdrawals. The extent of the effects of an increase in m for the domain $\underline{m} < m < \overline{m}$ can be obtained by taking partial derivatives with respect to m. Thus, an increase in m increases LT bookings by $\frac{\tau_c}{2b}$, decreases ST bookings by $\frac{\tau_c}{b}$, and increases the demand for storage by $\frac{\tau_c}{2b}$.

It can be seen that Propositions 4.3.2 and 4.3.4 include $m = 1$, the multiplier level that induces the same behaviour as in a pure commodity pricing regime (see Proposition 4.3.1). This is because for $m = 1$, traders are indifferent between solely procuring ST capacity, or rather booking LT capacity for the base load and ST capacity for the peak load.[67] The same holds for $m = 2$, the multiplier inducing the behaviour seen in a pure capacity pricing regime (see Proposition 4.3.1). A multiplier of 2 is valid in Propositions 4.3.3 and 4.3.4. This is because for $m = 2$, traders are indifferent between booking solely LT capacity, or rather procuring LT capacity to meet the base load and ST capacity for the peak load.[68] Therefore, the resulting dispatch and the ensuing welfare are not affected by the choices in these cases. This allows us to analyse the effects of the multipliers that induce a pure commodity and capacity regime behaviour by design (i.e. $m = 1$ and $m = 2$, respectively) in the remainder of the analysis without incorporating separate formulas for such multipliers. Thus, for $1 < m < 2$, the identified KKT points in

[67]A proof can be found in Appendix A Case 3a.
[68]A proof can be found in Appendix A Case 5c.

the Propositions 4.3.2, 4.3.3 and 4.3.4 are unique optimal solutions, which allow for a mixed-pricing regime.



(i) Booking and storage volumes

(ii) Transported volumes

Figure 4.2.: Development of the volumes for storage, ST capacity and LT capacity with respect to the multiplier (a); and development of transported volumes at time periods t_1 and t_2 with respect to the multiplier (b)

In Figure 4.2i we illustrate the findings of Propositions 4.3.2, 4.3.3 and 4.3.4 by plotting the traders' booking and storage decision with respect to m.[69] To be able to illustrate the results for all three identified domains, a setting is chosen in which feasible $\underline{m}$ as well as $\overline{m}$ exist (i.e. $\underline{m} > 1$ and $\overline{m} < 2$). This applies to all the figures in this paper, in which the effects are plotted for the respective multiplier domains. However, it should be noted that, depending on tariff levels, feasible $\underline{m}$ as well as $\overline{m}$ may not exist. In that case, storages would be utilised and transports would differ also for $m = 1$ as well as for $m = 2$.

Figure 4.2ii shows the transported volumes, which are equal to the sum of booked capacities in each period (i.e. $C_{12} + C_1$ in t_1 and $C_{12} + C_2$ in t_2). While the overall transported volume remains unaffected by m, the temporal spread of the transports, which can be interpreted as an indicator for transport volatility, decreases with m. In the multiplier range $m > \overline{m}$, the same amount of volumes are transported in both periods.

4.3.2. Deriving the Effects on Prices and Price Spreads

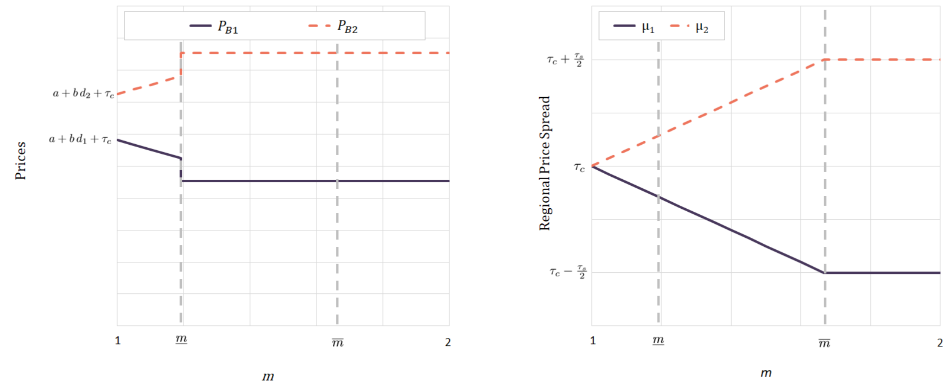
In a next step we derive the prices at each of the nodes. These prices are referred to as hub prices. In the analysed setting of perfect competition, prices correspond to the marginal cost of supply with respect to demand. Therefore, to obtain the prices in the demand region[70], we insert the solutions derived in the

[69]The parameters assumed for the figures in this section are as follows: $d_1 = 11$, $d_2 = 30$, $\tau_c = 6$, $\tau_s = 8$, $a = 4$, $b = 0.15$.

[70]Our analysis does not focus on the prices in production regions. For the sake of completeness, we derive the prices in the production region A in Supplementary Material C.3.

Propositions 4.3.2, 4.3.3, and 4.3.4 in the total cost function shown in Equation 4.9, and differentiate with respect to d_1 and d_2.

$$P_{B1} = \frac{\partial Cost^{Tot}}{\partial d_1} = \begin{cases} a + b\,d_1 + (2-m)\,\tau_c & \text{for } m \leq \underline{m} \\ a + b\left(\dfrac{d_1+d_2}{2}\right) + \tau_c - \dfrac{\tau_s}{2} & \text{for } m > \underline{m} \end{cases}$$
$$P_{B2} = \frac{\partial Cost^{Tot}}{\partial d_2} = \begin{cases} a + b\,d_2 + m\,\tau_c & \text{for } m \leq \underline{m} \\ a + b\left(\dfrac{d_1+d_2}{2}\right) + \tau_c + \dfrac{\tau_s}{2} & \text{for } m > \underline{m} \end{cases} \quad (4.13)$$



(i) Prices

(ii) Regional Price Spreads

Figure 4.3.: Development of the hub prices in region B (a) and the regional price spread between regions A and B (b), at time period t_1 and t_2 with respect to the multiplier

The functions describing the consumer prices in the demand region are plotted in Figure 4.3i. For the domain $m < \overline{m}$, in which the traders do not use storages and their capacity demand is inelastic, the price in peak period (P_{B2}) increases. This occurs as marginal demand is transported using additional ST capacity whose price increases in m. Conversely, the price in off-peak period (P_{B1}) decreases as additional demand is met by a shift from ST to LT capacity in this period. Such a reallocation of network costs from off-peak users towards peak consumers is in line with the concept of peak-load pricing.

In the domain $\underline{m} < m < \overline{m}$, the traders were shown to have elastic capacity demand, meaning that they are able to switch from ST to LT capacities with increasing m by using storages. The prices in this case remain constant over m which may seem counter-intuitive since ST transmission tariffs increase in m. However, this is due to additional demand being met by an increase in LT capacity booking and storage usage while ST capacity bookings remain unchanged. This applies to both d_1 and d_2, resulting in consumer prices (P_{B1} and P_{B2}) to be independent of m. In line with the findings of Nguyen (1976), we also show here that the peak price exceeds the off-peak price by the cost of storage (i.e. $P_{B2} - P_{B1} = \tau_s$). In the domain of $m \geq \overline{m}$, despite the inelastic capacity demand, prices are unaffected by changes in m. This is due to the absence of ST bookings and the utilisation of storages. Furthermore, the temporal price spread here is also set by storages.

Interpreting the temporal price spread as price volatility, it can be said that higher multipliers can cause increased volatility in hub prices unless storages are utilised—which requires enough storage capacities to be available and that storage tariffs are sufficiently low compared to transmission tariffs.

In contrast, we find the average hub price to be constant and independent of the multiplier. The average hub price is equal to the gas procurement price that arises when volumes are bought evenly in both periods, plus the base transmission tariff:

$$\frac{P_{B1} + P_{B2}}{2} = a + b\left(\frac{d_1 + d_2}{2}\right) + \tau_c \tag{4.14}$$

The regional price spreads between the modelled regions A and B correspond to the Lagrange multipliers[71] μ_1 and μ_2, for the time periods t_1 and t_2, respectively. As derived in Case 5 (a) in Supplementary Material C.2, those spreads are presented in Equation 4.15 and are plotted in Figure 4.3ii for the corresponding multiplier domains.

$$\begin{aligned} P_{B1} - P_{A1} = \mu_1 &= \begin{cases} \tau_c\,(2-m) & \text{for } m < \overline{m} \\ \tau_c - \dfrac{\tau_s}{2} & \text{for } m \geq \overline{m} \end{cases} \\ P_{B2} - P_{A2} = \mu_2 &= \begin{cases} m\,\tau_c & \text{for } m < \overline{m} \\ \tau_c + \dfrac{\tau_s}{2} & \text{for } m \geq \overline{m} \end{cases} \end{aligned} \tag{4.15}$$

Results indicate that multipliers cause temporal variation in regional spreads: In the peak period, additional transport demand is met by procuring ST capacity, resulting in a price spread of $m\,\tau_c$. In contrast, additional transport demand in the off-peak period is met by replacing ST capacity with LT capacity, inducing regional spreads of $\tau_c\,(2-m)$. Thus, higher multipliers lead to the widening of the temporal price margin of regional spreads. In sum, the effects in the two periods cancel each other out, such that average regional price spreads remain constant over m.

On the other hand, regional spreads in the domain with pure capacity pricing behaviour ($m \geq \overline{m}$) are found to be independent of the multiplier. As the same volumes are transported in both periods (due to only LT product being booked with storage utilisation), the regional spreads in this case are defined by the storage tariff and are constant. Nevertheless, since the majority of real situations in the

[71] Alternatively, regional price spreads can be derived by subtracting the prices in regions A and B.

EU are expected to correspond to mixed-pricing regimes, our results indicate that higher multipliers are likely to cause increased volatility in regional price spreads.

4.3.3. Deriving the Effects on Surpluses and Welfare

Having illustrated the impacts of multipliers on prices and price spreads, we now proceed with the analysis of the effects on the surplus of consumers, gas producers, the TSO and the traders as well as on the resulting welfare.

Consumer surplus
To allow for a clear illustration of welfare effects we assume the consumer surplus (CS) of base-load and peak-load consumers[72] to be zero for the range of multipliers which result in the highest costs for those consumers. As a result, consumer surplus is obtained as a function of the multiplier, corresponding to the difference between this threshold and the respective consumer costs. The respective consumer surpluses can be expressed as follows:

$$\textit{Base-load } CS = 0$$

$$\textit{Peak-load } CS = \begin{cases} \frac{1}{2}\,(d_2 - d_1)\left(\tau_s - 2\,\tau_c\,(m-1) - b\,(d_2 - d_1)\right) & \text{for } m \leq \underline{m} \\ 0 & \text{for } m > \underline{m} \end{cases} \tag{4.16}$$

In Figure 4.4, which plots the derived surplus and welfare functions of the respective agents in the model, the development of consumer surplus in the identified multiplier domains can be seen. Base-load consumers do not earn a surplus with increasing m since their overall costs are not affected by m due to the average prices being constant and their demand being inelastic. For peak-load consumers, in contrast, total costs depend on m as more gas is bought in t_2 than in t_1. Therefore, when prices in t_2 increase and prices in t_1 decrease with the same magnitude, despite the average price remaining constant, overall consumer costs increase. Hence, when P_{B2} is highest (i.e. $m > \underline{m}$) consumers do not earn any surplus. Consumer surplus is greatest, when P_{B2} is lowest (i.e. $m = 1$).

[72] Remember, consumers are assumed to be divided into two groups: Base-load consumers with a flat demand equal to d_1 in both periods and peak-load consumers who consume $d_2 - d_1$ in t_2.

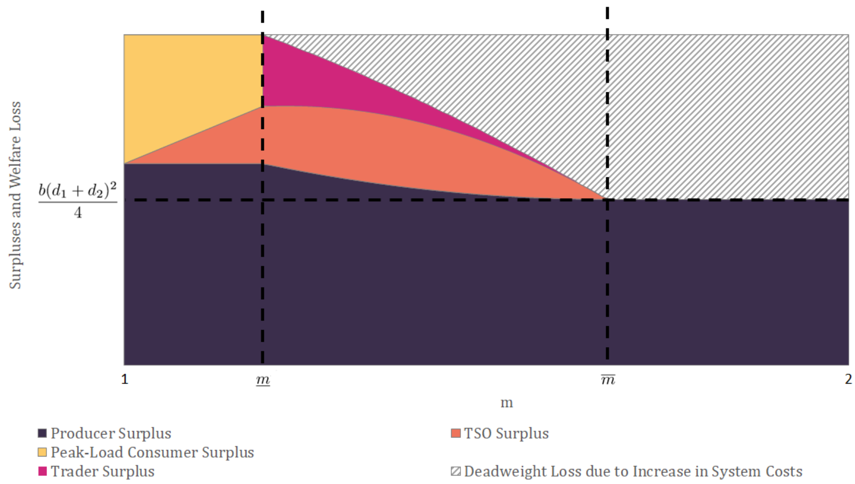


Figure 4.4.: Producer, trader, consumer and TSO surpluses, and deadweight loss with respect to m

Producer Surplus

Producers earn a surplus by selling their gas for a price which is higher than their marginal costs. Producer surplus occurs since marginal costs increase in procured volumes in the model setting, which is representative of the real cost structures for the producers. The resulting surplus thus equals:

$$Producer\ Surplus = \begin{cases} \dfrac{b\left(d_1^2+d_2^2\right)}{2} & \text{for } m \leq \underline{m} \\ \dfrac{b\left(d_1+d_2\right)^2}{4} + \dfrac{\left(2\tau_c\left(m-1\right)-\tau_s\right)^2}{16b} & \text{for } \underline{m} < m < \overline{m} \\ \dfrac{b\left(d_1+d_2\right)^2}{4} & \text{for } m \geq \overline{m} \end{cases} \tag{4.17}$$

Producer surplus is highest when $m < \underline{m}$ and lowest for $m \geq \overline{m}$. For multipliers lying in the interval $\underline{m} < m < \overline{m}$, producer surplus decreases with m. This is because profits depend exponentially on sold volumes per period, and as such, producer surplus decreases as sold volumes in t_1 and t_2 converge to the same value.

TSO surplus

The TSO receives revenues from the capacity products booked by the traders. We assume the TSO's revenues to be sufficient to cover costs in a setting without multipliers (i.e. $m = 1$) and any increase in the multiplier level can therefore result in surplus revenues. The resulting surplus can then be expressed as follows:

$$TSO\ surplus = \begin{cases} \tau_c\,(d_2 - d_1)\,(m-1) & \text{for } m \leq \underline{m} \\ \dfrac{\tau_c\,\tau_s\,(m-1)}{2b} - \dfrac{\tau_c^2\,(m-1)^2}{b} & \text{for } \underline{m} < m < \overline{m} \\ 0 & \text{for } m \geq \overline{m} \end{cases} \tag{4.18}$$

When $m = 1$ or when solely LT capacities are booked, i.e. $m \geq \overline{m}$, the TSO does not earn a surplus. Between those thresholds, the TSO surplus follows a concave form and reaches its maximum at $m = 1 + \frac{\tau_s}{\tau_c}$, as can be seen in Figure 4.4. The path of the surplus function is based on the combination of two effects: Firstly, the TSO's income increases with increasing m directly due to ST capacity becoming more expensive—an effect that exists for all $m > 1$. Secondly, as traders increasingly shift their bookings from ST to LT capacity with increasing m, the additional revenue generated by the TSO due to more expensive ST capacities is reduced. This effect emerges when m reaches $\underline{m}$, as the storages become utilised and switches from ST to LT booking start to take place. For $m < 1 + \frac{\tau_s}{\tau_c}$, the first effect is more dominant; while for larger values of m, the second effect dominates.

Storage operator surplus
Storage operators do not earn any surplus under perfect competition as they are assumed to have constant marginal costs.

Trader surplus
Surplus of the traders equals the difference of consumer prices and costs of gas provision (i.e. sum of procurement, transport and storage) which is equal to:

$$Trader\ surplus = \begin{cases} \dfrac{\big(\tau_s - 2\,\tau_c\,(m-1)\big)^2}{8\,b} & \text{for } \underline{m} < m < \overline{m} \\ 0 & \text{otherwise} \end{cases} \tag{4.19}$$

Traders start making surplus when the multipliers cross the $\underline{m}$ threshold. This is because storages become part of the optimal solution. The utilisation of storages creates markups of $\frac{\tau_s}{2}$ in the peak period (t_2) and markdowns of $\frac{\tau_s}{2}$ in the off-peak period (t_1). Since sold volumes in t_2 are higher, a profit is generated. However, as storage utilisation increases with increasing m, this results in higher storage costs and subsequently diminished profits. Traders also bear the additional ST capacity costs arising from increased multipliers, which further reduce the trader surplus.

Welfare
Having derived the individual surplus functions of all the relevant agents of the model, we now derive the total welfare function. Total welfare corresponds to the

sum of consumer, producer, TSO, and trader surplus. This equals:

$$Welfare = \begin{cases} \dfrac{(d_2 - d_1)\,\tau_s}{2} + b\,\big(d_1\, d_2\big) & \text{for } m \leq \underline{m} \\ \dfrac{b\,(d_1 + d_2)^2}{4} + \dfrac{(\tau_s - 2(m-1)\,\tau_c)(3\tau_s + 2(m-1)\,\tau_c)}{16\,b} & \text{for } \underline{m} < m < \overline{m} \\ \dfrac{b\,(d_1 + d_2)^2}{4} & \text{for } m \geq \underline{m} \end{cases} \tag{4.20}$$

Welfare is maximal when the gas dispatch is not distorted by transmission tariffs. In our model with perfectly inelastic consumer demand, efficient outcomes with maximal welfare are achieved for $m < \underline{m}$ in the case where $\underline{m} \geq 1$ (plotted in Figure 4.4), or for $m = 1$ if $\underline{m}$ does not exist in the feasible multiplier domain (plotted in Figure C.2 in Supplementary Material C.4).[73] Note that the welfare being maximal when $m = 1$ is not surprising and can be considered as a trivial solution since the transport costs are not affected by multipliers.

As soon as $m > \underline{m}$, higher multipliers reduce welfare by causing additional costs, which occurs as a result of two opposing effects: On the one hand, since the total production cost function is quadratic, total costs of gas production decrease as gas is produced more evenly. On the other hand, total costs of storing gas increase. However, as the increase in storage costs is higher than the decrease in production costs, welfare declines with increasing m. Welfare becomes independent of the multiplier when the multiplier reaches the threshold $\overline{m}$ as gas production in t_1 and t_2 fully converges.

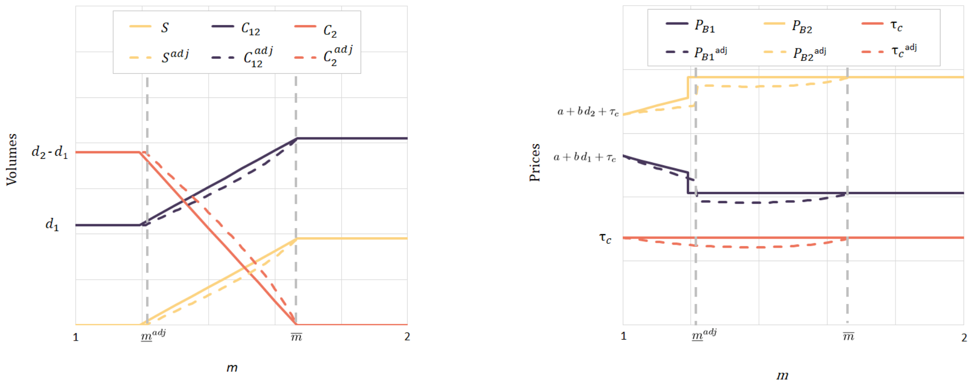
4.3.4. The regulated TSO: Transmission Tariff Adjustment

We have shown that the TSO makes a surplus as long as $m > 1$ and the traders book ST capacity when $m < \overline{m}$. In reality, being natural monopolies, TSOs are regulated entities and are not allowed to exceed certain revenue caps. Hence, in the case of a potential surplus due to multipliers, the TSO would have to lower its transmission tariffs (i.e. entry/exit tariffs) accordingly for the next year in order to remain at the regulated revenue cap. In this model extension, we consider this aspect by introducing the adjusted transmission tariff τ_c^{adj} which is set such that the TSO surplus is zero for all m. Since τ_c^{adj} is only a parameter for the agents of our model and does not change the nature of the problem; the optimisation rationale of the agents remains the same as in our main model.

We find that the results with adjusted transmission tariff τ_c^{adj} are similar to the model results with fixed τ_c. All the general findings regarding the effect

[73] According to economic theory, when consumers' demand is elastic, variable transmission tariffs to cover fixed network costs reduce welfare since they reduce consumers' demand. Such variable costs arise in the entry-exit system independent of the level of multipliers. To achieve more efficient outcomes in the presence of elastic demand, other tariff regimes (e.g. fixed grid fees) may be more appropriate (Borenstein, 2016).

of m on volumes and prices and price spreads remain intact. The lowered τ_c^{adj} slightly increases $\underline{m}$, the multiplier threshold which is sufficient to incentivise the use of storages. The upper threshold $\overline{m}$ remains unchanged. We define this adjusted threshold as $\underline{m}^{adj}$. Plotting the capacity and storage volumes resulting from adjusted tariffs in Figure 4.5i, we see that adjusting the transmission tariff also slightly increases ST capacity bookings, decreases LT bookings, and as a consequence, results in lower utilisation of storages for $\underline{m}^{adj} < m < \overline{m}$. New hub prices as a result of adjusted tariffs are plotted in Figure 4.5ii. The average regional price spread still equals the transmission tariff. However, since τ_c^{adj} is lower than τ_c for $1 < m < \overline{m}$, tariff adjustment leads to lower average regional price spreads for $m > 1$. The price spreads are lowest for $m = 1 + \frac{\tau_s}{\tau_c^{adj}}$. Similarly, the lowered transmission tariff translates directly to lower gas consumer prices, hence the average prices are also lowest at $m = 1 + \frac{\tau_s}{\tau_c^{adj}}$.



(i) Volumes

(ii) Prices and average regional price spread (τ_c^{adj})

Figure 4.5.: Volumes and prices when τ_c is adjusted such that the TSO does not earn a surplus

The surpluses and welfare effects are plotted in Figure 4.6. When transmission tariffs are adjusted, the TSO does not earn a surplus anymore. The surpluses of traders and gas producer surplus are impacted very slightly. These effects result from the changes in the production pattern and storage volumes and not from a shift of the TSO's surplus. Instead, the tariff adjustment redistributes all of the surplus formerly earned by the TSO to the consumers. Base-load consumers, who did not earn any surplus when the tariff was fixed, earn a surplus with adjusted tariffs. In the domain $m < \underline{m}^{adj}$, the surplus of base-load consumers increases in m. In the domain $\underline{m}^{adj} < m < \overline{m}$, surpluses of both base-load and peak-load consumers increase in m for sufficiently low multiplier levels ($m < m^{CS,max}$) due to lower consumer prices resulting from decreased LT tariffs. This implies that if feasible $\underline{m}^{adj}$ does not exist due to tariff structures, a multiplier level equal to $m^{CS,max} = 1 + \frac{\tau_s}{\tau_c^{adj}}$ maximises the total consumer surplus (such a case is plotted in Supplementary Material C.4). For $m > m^{CS,max}$, consumer surplus decreases with m due to increasing system costs. In the domain $m > \overline{m}$, consumer surplus is zero, which was also the case with fixed tariffs.

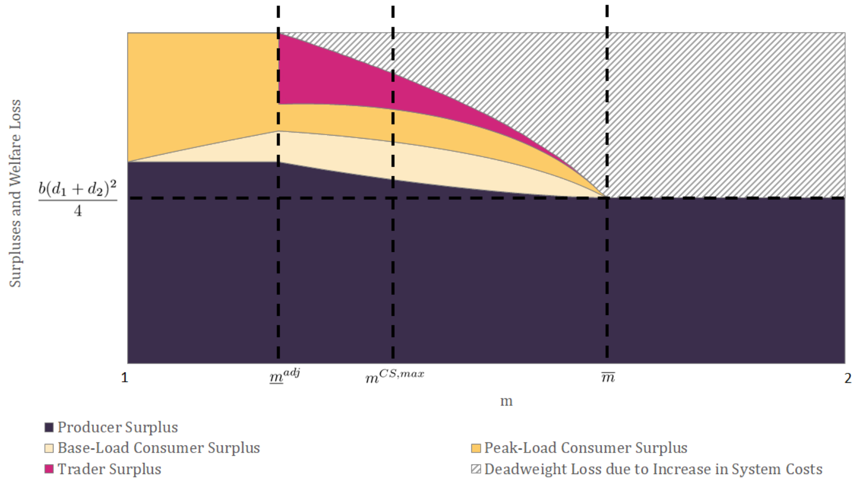


Figure 4.6.: Producer, trader, consumer and TSO surpluses, and deadweight loss with respect to m when τ_c is adjusted such that the TSO does not earn a surplus

4.4. Discussion

4.4.1. Effects on Infrastructure Utilisation

Multipliers, by making ST products comparably more expensive, can cause a switch from ST capacities to LT capacities, decrease the volatility of pipeline transports and consequently lead to more uniform capacity utilisation. These aspects associated with higher multipliers have also been mentioned in several studies related to the EU tariff structures (ACER and CEER, 2019, DNV-GL, 2018, EY and REKK, 2018, Rüster et al., 2012, Strategy& and PwC, 2015). We also find that gas storages can have increased utilisation rates with higher multipliers. However, these effects are not universal and strongly depend on the underlying tariff structures, i.e. occurring only when multipliers are neither too high nor too low with respect to the ratio of storage to transmission tariffs such that the capacity demand of the traders is elastic, meaning that the traders can switch between LT and ST products.

The proportions of these tariffs and multipliers constitute the multiplier thresholds (i.e. $\underline{m}$ and $\overline{m}$), which define domains with varying effects of multipliers. We find that multipliers equal to 1 or lower than the threshold $\underline{m}$ result in users to behave as if in a pure commodity pricing regime, while multipliers larger than $\overline{m}$ induce the same behaviour as observed in a pure capacity tariff regime. When multipliers are in between $\underline{m}$ and $\overline{m}$ an inherent mixed regime of capacity and commodity pricing occurs.

The multiplier domains identified by the theoretical model can also be observed in the EU gas markets. Depending on the circumstances, multipliers in the EU can lie in each of the domains identified by the model, their magnitude corresponding to values smaller than $\underline{m}$, higher than $\overline{m}$ or to those that lie in between.

The domain $m < \underline{m}$, for instance, represents a situation where storages are not used. This would occur when marginal storage costs are sufficiently high compared to m. Further, cross-border transports in each period match the corresponding demand. An utmost example, in this regard, would be the case of Finland where there are no gas storages and all of the gas was imported only from a single source until recently[74]; namely, Russia (Jääskeläinen et al., 2018). This implies infinitely large storage tariffs ($\tau_s \to \infty$) for Finland, irrespective of the existing multiplier levels in the country. Hence, any multiplier lies below the lower threshold $\underline{m}$.

Situations corresponding to the domain $m > \overline{m}$, on the other hand, occur when the transported volumes are constant and storages fill the gap between the demand and the imports instead. This would be observed when transmission capacity tariffs are sufficiently high with respect to the multiplier. Such instances can arise for pipelines that are consistently operated at their full capacities as this indirectly corresponds to transmission tariffs being infinitely high for marginal capacity demand ($\tau_c \to \infty$). Hence, any $m > 1$ would already be larger than the upper threshold $\overline{m}$.

In the majority of situations including connections between market areas, both pipelines and storages are utilised and neither of the two operate at their full capacity. These situations correspond to the $\underline{m} < m < \overline{m}$ domain where an inherent mixed regime of capacity and commodity pricing occurs, and as a result, the transmission capacity demand of traders is elastic. This is also valid for countries that apply multipliers equal to 1, where both LT and ST capacities are booked and storages are utilised (this implies feasible $\underline{m}$ does not exist).[75]

Even though we have implied here the possibility of directly observing those domains and their effects in the EU gas transmission system for various country pairs and pipelines, it is likely that a mixture of these effects would be prevalent in numerous regions. This is because all the analysed domains arise simultaneously within the EU and on its outer borders, and gas is often transported through several countries. On average, the aggregate effect on volumes, prices and, surpluses would likely be a combination of all of those domains for the EU.

[74] As of 1 January 2020, Finland is connected with Estonia via the Balticconnector pipeline (European Commission, 2020).

[75] A corresponding example is the case of Germany during the period 2012–2015 before the introduction of the BEATE regulation. More information can be found in the resolutions BK7-10-001 and BK9-14/608 of the German regulatory agency Bundesnetzagentur.

4.4.2. Effects on Hub Prices

Regarding hub price levels in gas importing regions, model results have several implications: Temporal price spread increases with increasing m if storage utilisation is zero due to comparably high storage tariffs or unavailability of storage capacity (i.e. the domain $m \leq \underline{m}$). In such cases, higher multipliers can cause increased volatility in hub prices. In the case storages are utilised (i.e. the domain $m \geq \underline{m}$), then the storages dampen the effect on temporal price spreads.

Our analysis indicates that increasing multipliers can result in higher regional price spreads, since the upper limit of the spread is shown to be equal to the price of ST capacities ($m\,\tau_c$).[76] ACER refers to such a price spread ($m\,\tau_c$) as the "reference" regional spread (ACER, 2019a), implying that price spreads increase with increasing multipliers on average. In our model, in contrast, increases in spreads are only limited to temporal variations (i.e. increased volatility in spreads), while the average regional price spread remains equal to the transmission tariff (τ_c). This is because the marginal demand is satisfied by LT capacity. In reality, uncertainty as well as frictions in the secondary market for capacity may require the booking of ST capacity to satisfy marginal demand in some situations. As a result, average price spreads are likely to be between LT and ST capacity tariffs.

Whether multipliers increase or decrease regional price spreads also depends on the effect of multipliers on the LT tariff. In our model extension in Chapter 4.3.4, which takes into account transmission tariff adjustments by the TSO, we have shown that increases in m allow the TSO to reduce the tariff (τ_c^{adj}) if multipliers are sufficiently low ($m < \frac{\tau_s}{\tau_c^{adj}}$), an aspect also mentioned in several consulting studies (Rüster et al., 2012, Strategy& and PwC, 2015). Therefore, increases in m can both decrease average hub prices and average regional price spreads, which were shown to depend on the transmission tariff. This is an aspect, which studies such as ACER (2019a) and EY and REKK (2018) apparently do not consider when stating that increases in multipliers are likely to increase regional price spreads. By reducing LT tariffs, sufficiently high multipliers may also help support tariff stability by mitigating the tariff increase which is expected to occur when historical LT bookings expire (ACER, 2019a).[77] However, if policymakers set multipliers too high such that they discourage traders from booking ST capacities, we have shown that increasing m elevates the transmission tariff and prices.

[76]Applies to uncongested pipelines.

[77]For instance, during the period 2016–2018, about 80% of the total capacity used by traders stemmed from existing LT bookings which were undertaken before ST capacities were introduced (ACER, 2019a), the majority having been booked upfront covering multiple years. As those old bookings start expiring during 2020–2030, the prevalent situation of overbooked capacities and the sunk costs associated with them will start disappearing such that the cost of new bookings will represent the actual opportunity costs (EY and REKK, 2018).

4.4.3. Effects on Surpluses and Welfare

Model results show that the lowest total system costs and correspondingly the highest total welfare are associated with lower multipliers. This is because higher multipliers cause the gas dispatch to deviate from an ideal dispatch based on short-run marginal costs. Nevertheless, the notion that an increase in m always results in higher system costs and lower welfare does not universally apply, but is highly dependent on which domain the system lies in (i.e. the ratio of storage to transmission tariffs with respect to multipliers).

For the identified domain without storage utilisation ($m < \underline{m}$), an increase in m does not cause additional system costs and no consequent welfare losses, as the transported volumes are fixed and independent of m. Similarly, producer surplus remains constant due to fixed volumes. Because storage utilisation is zero in this domain, traders do not make any surplus as they cannot exploit the intertemporal arbitrage potential. Consumer surplus, on the other hand, decreases with increasing m and is passed on to the TSO as a surplus unless the transmission tariffs (τ_c) are adjusted. In the case where the tariffs are adjusted such that the TSO does not make surplus (i.e. no additional TSO revenue than the regulated amount), higher multipliers cause consumer surplus to be redistributed from peak-load consumers (i.e. households) to base-load consumers (i.e. industry). This finding is in line with the implications stated by Strategy& and PwC (2015) and DNV-GL (2018).

We have also shown that sufficiently high multipliers ($m > \overline{m}$) are associated with higher total system costs and lower total welfare. In this setting, the surpluses of the consumers, traders and the TSO are all zero while only the producers make a constant surplus.

In the domain where storages are utilised and both ST and LT products are booked ($\underline{m} < m < \overline{m}$)—a case which is likely to be present in the majority of EU countries—increasing m results in increased system costs and decreased welfare. Trader surplus exists in this domain. However, it decreases exponentially with increasing m as gains by intertemporal arbitrage are reduced due to higher storage utilisation and the respective convergence in gas prices in the production region. The same effect causes the producer surplus to decrease as well. This also offers an explanation why gas traders such as Uniper SE and Gazprom Export and gas producers such as Shell Energy request low multipliers in their statements during the multiplier consultations (BNetzA, 2019).

TSO makes surplus for $m > 1$ if the transmission tariff is not adjusted. For multipliers that are sufficiently low ($m < 1 + \frac{\tau_s}{\tau_c}$), the TSO surplus increases initially with increasing m due to the additional revenue from ST products. As TSOs may be able to retain at least some of this surplus, they have an incentive to request higher multipliers than traders and producers do. Something which can be observed in the consultation statements of TSOs such as Open Grid Europe, Bayernets, ONTRAS (BNetzA, 2019). When the transmission tariff is adjusted for zero TSO surplus, then the surplus is passed on to the consumers due to lower hub prices.

The results also indicate that for the domain where the capacity demand of traders is elastic and which is representative of the majority of the situations observed in the EU, there exists a multiplier larger than 1 that maximises consumer surplus (i.e. $m = 1 + \frac{\tau_s}{\tau_c^{adj}}$).

This presents us with an interesting trade-off: Minimising total system costs and maximising total welfare in the short-run requires setting the multiplier equal to 1. However, a policymaker willing to maximise consumer surplus would aim for a multiplier greater than 1 but sufficiently low. Furthermore, higher multipliers may enhance security of supply due to increased storage utilisation and potentially resulting in storage investments. Since higher multipliers are more in line with peak-load pricing, and thus help decrease the peak-load capacity demand, the policymaker may also prefer higher multipliers to reduce the need for capacity expansion and to increase long-term efficiency.

We assume demand to be perfectly inelastic, although one could argue that the gas demand from power generation has a certain elasticity due to fuel switching between gas and coal plants. This could have the following effects. As we have shown, multipliers larger than 1 decrease average prices and thereby would increase demand and consumer surplus if overall demand was elastic. On the other hand, if only peak demand was elastic, peak prices may increase which would decrease demand and consumer surplus.

We should note that the assumption of perfectly efficient secondary markets is relevant when interpreting our model results regarding welfare. The importance of developed and liquid secondary capacity markets for efficient explicit auction mechanisms is highlighted in the literature (Kristiansen, 2007, Oren et al., 1985, Peŕez-Arriaga and Olmos, 2005). Secondary markets allow traders to exchange booked capacities, enabling them to adjust their commercial positions (Peŕez-Arriaga and Olmos, 2005) and balance their marginal benefits (Oren et al., 1985). Therefore, an imperfect secondary market can hinder the exchange of some booked LT capacities and can lead to instances of contractual congestion[78], even if sufficient technical capacity is available to meet the demand. In such a situation, some traders waste their capacity rights, whereas other traders, whilst having positive capacity demand, are not able to book capacities—a phenomenon that consequently results in underutilised pipelines and inefficient dispatch. As such, Hallack and Vazquez (2013) argues within the context of the EU entry-exit tariff system that secondary markets help relieve contractual congestion. We have shown that the ratio of LT bookings increases with increasing multipliers. Therefore, in the case where secondary markets for gas transmission capacities in the EU are not efficient, higher multipliers could cause additional welfare losses due to more frequent instances of contractual congestion, a view shared also in

[78] Contractual congestion means a situation where the level of firm capacity demand exceeds the technical capacity of a pipeline.

several technical reports (Rüster et al., 2012, Strategy& and PwC, 2015). Hence, in order to minimise those additional welfare losses, policymakers should further promote efficient secondary markets.

4.5. Conclusion

In this paper, we take a theoretical perspective on the effects of multipliers on gas infrastructure, hub prices and welfare. The model developed for this purpose depicts a setting of perfect competition and is solved analytically by minimising total costs using KKT conditions. The effects of multipliers are then derived from the various solutions to the problem.

Our model results indicate that higher multipliers can cause a switch from short-term (ST) transmission capacity bookings to long-term (LT) bookings, lead to more uniform pipeline transports, and increase gas storage utilisation. In the majority of countries and situations these findings are expected to hold. However, the effects are not universal and are found to depend on the traders' elasticity of capacity demand. Depending on the proportion of multipliers with respect to storage and transmission tariff levels, situations with inelastic capacity demand can arise. It is possible when multipliers are sufficiently low with respect to the tariffs, gas storages are not utilised in the context of capacity bookings. On the other hand, multipliers that are considerably high can cause only LT capacities to be booked.

Regarding the effects of multipliers on hub prices, we find that higher multipliers cause maximum regional price spreads to increase, indicating that they can result in increased volatility in regional price spreads. However, on average, we show that hub prices and regional price spreads can decrease with increasing multipliers, as long as multipliers remain sufficiently low. These effects occur since higher multipliers allow the TSO to lower the transmission tariffs.

Model results show that higher multipliers are associated with higher total system costs and consequently lower total welfare in the short-run. Despite that, for the identified multiplier domain, which is representative of the majority of the situations in the EU gas system, our results indicate that the multiplier maximising total consumer surplus is larger than 1.

Our findings have various policy implications: Setting the multipliers equal to 1 minimises total costs of gas dispatch and thereby maximises total welfare. However, if the aim of the policymakers is to maximise consumer surplus, then opting for multipliers that are greater than 1 but are still sufficiently low can help in achieving the desired outcome. Moreover, a multiplier greater than 1 would lead to redistributing the consumer surplus from peak-load consumers to base-load consumers, if that is desired. In that sense, higher multipliers can also help reduce peak load and therefore result in potential welfare gains in the long-term due to a decreased need for new capacity investments. Since we have shown that higher multipliers cause increased storage utilisation, it could be argued that setting multipliers sufficiently high can also contribute to security of supply by incentivising additional storage investments. Multipliers that are considerably high, however, increase regional price spreads and undermine market integration; and if sufficiently high, can cause only LT capacities to be booked, potentially impeding efficient gas dispatch.

We have shown that optimal level and thresholds for multipliers depend on the level of transmission and storage tariffs. Therefore, it is important to consider the existing tariff structures when setting multipliers. As the current EU tariff landscape has significant variation in tariff structures and levels, this implies a one-size-fits-all approach with a single uniform EU multiplier may not lead to optimal outcomes for individual countries. We therefore find it appropriate that EU regulation specifies the allowed multiplier levels in ranges and not in absolute values. Nevertheless, whether the specified range covers the optimal levels or is too restrictive remains to be researched.

In future work, the model can be applied in a real-world setting by incorporating more time periods and a realistic network structure representative of the EU gas transmission system. The extended model can be used to quantify the effects of multipliers with numerical simulations. This would allow to analyse the effects of regional variations in multiplier levels throughout the EU. An interesting aspect in this case would be to evaluate whether optimal multipliers for individual countries are also optimal for the overall EU system, or whether they cause negative externalities on other countries. Another possibility would be to extend the model by including stochasticity regarding capacity demand in order to represent the realistic situation of imperfect information and uncertainty.

5. Internal and external effects of pricing short-term gas transmission capacity via multipliers

5.1. Introduction

When a region decides on network pricing, different circumstances lead to different optimal tariff settings. In this context, two questions arise in particular: First, how does the optimal tariff setting vary among different regions? And second, because networks connect multiple regions, do the individual regional optima contribute to the joint optimum or do they cause negative externalities such that only a superordinate regulator can achieve the joint optimum?

These questions also arise in the case of the gas transmission network of the European Union (EU), which connect different regional networks called market areas. To finance the networks in the individual market areas the transmission system operators (TSO) charge transmission tariffs. Regulation (EC) 2009/715 introduced a tariff regime that obligates gas traders to book entry and exit capacity when transporting gas from one market area into another.[79] In this context, traders are offered capacity products with varying run-times: long-term (LT) yearly products, and the short-term (ST), quarterly, monthly, daily, and intra-daily products. Regulation (EC) 2009/715 allows each national regulator to define their relative price of ST versus LT capacities within specified ranges. The relative prices of the ST capacities are defined by factors called multipliers, i.e, the ST capacity prices are equal to the LT capacity price multiplied by the corresponding multipliers. The levels of those multipliers are found to affect the proportion of ST to LT capacity booking and consequently impact the infrastructure utilisation, prices, and welfare distribution (Çam and Lencz, 2021).

The effects of multipliers in the EU gas system are expected to become more amplified in the coming decades. A major contributor in this regard will be the expiration of old long-term bookings.[80] For instance, between 2016 and 2019,

[79] Capacities are booked in capacity auctions performed on trading platforms (such as PRISMA, GSA, RBP) in which the reserve prices correspond to the transmission tariffs. In a large share of the capacity auctions in the EU, demand for capacity remains below the offered capacity (ACER, 2019c). In the remaining cases where demand for capacity exceeds the offered capacity, a congestion premium occurs.

[80] A large share of current transmission capacity is booked by previous LT bookings at a time when ST capacity products did not exist. Those long-term bookings covered usually multiple years upfront.

about 80% of the total capacity used by traders stemmed from existing long-term bookings which were undertaken before the current system of LT and ST capacities were introduced (ACER, 2020b). For some connections between market areas these old long-term bookings exceeded the demand for capacity, inducing marginal transmission costs of zero. As those old bookings are about to expire over the period 2020–2035, the prevalent situation of overbooked capacities, and the sunk costs associated with them, will start disappearing.[81] In the future, the cost of new bookings will represent the actual opportunity costs, a development that is also mentioned in a study commissioned by the EU on gas market design (EY and REKK, 2018).

Our paper is strongly motivated by Çam and Lencz (2021), which has analysed the effects of multipliers on gas infrastructure utilisation, prices, and welfare using a theoretical model within a stylised setting. Applying the stylised theoretical model with two time periods and two regions, where pipeline and storage capacities were assumed to be unlimited, Çam and Lencz (2021) showed that a multiplier value of 1 leads to highest total welfare and multipliers greater than 1 cause welfare loss. The paper found that higher multipliers can nevertheless maximise the consumer surplus depending on the cost of gas transport and storage. This indicates that the consumer-surplus-maximising multiplier levels can differ between individual regions. In this respect, it is plausible to assume that EU would rather aim to maximise the consumer surplus instead of total welfare, since a substantial share of the surpluses generated by producers, storage operators, and traders arise outside the EU. Hence, we refer to consumer-surplus-maximising multipliers as optimal multipliers.

In a more complex setting with multiple time periods, multiple regions and limited infrastructure capacities—such as in the case of the EU gas transmission system—there are additional aspects that would influence the optimal multiplier levels. For instance, the temporal profile of gas demand in a region could substantially influence the proportion of LT to ST bookings. In countries that have relatively flat demand profiles throughout the year, gas imports and bookings would be at similar levels during winter and summer, allowing for a very high share of LT bookings. In this case, effects of multipliers can be limited if LT bookings are preferred irrespective of the multiplier levels. In contrast, in regions with highly seasonal demand but limited storage capacities, booking ST capacities could be preferred. With sufficiently high multipliers, booking ST capacities to cover the peak winter demand could eventually become more expensive than booking only LT capacity. In this case, traders could choose to book only LT capacity while letting some capacity during the summer months remain unused. Multipliers could therefore exacerbate this type of booking patterns in such regions. Hence, due to having different features as mentioned above, individual regions can be

[81] ACER (2020b) states that more than a third of such old long-term capacity bookings in place at the end of 2019 will have expired by the end of 2023, while more than 60% of them will no longer be in place by 2028. Old long-term contracts will almost completely expire by the end of 2035.

affected differently by multipliers and can have varying optimal multiplier levels. In order to determine the individually optimal multiplier levels, it is necessary to represent these regional features and analyse the internal effects of multipliers in a more realistic model setting.

In addition to inducing internal effects, multiplier levels in a region can cause externalities in other regions due to the fact that gas is transported through different regions. It is commonly acknowledged that tariff adjustments in a country can cause external effects in another country within the EU gas network, depending on their location along the gas transport chain. For instance, Cervigni et al. (2019) points out that national regulators can impact the sharing of transport costs between the consumers of individual countries through their selection of entry and exit tariff levels. It is argued that a transit country can transfer the cost of transmission investments, which largely benefit its own citizens, to a downstream country's consumers via its choice of entry-exit tariffs at the interconectors. Similarly, Petrov et al. (2019) mentions that the tariff adjustments in Germany (in the context of the REGENT regulation) can cause significant costs in the neighbouring market areas of Czechia and Italy when the costs of the network tariff change are passed on to the gas consumers in these regions. Since multipliers influence the relative tariff levels of ST capacities, it is therefore natural to think that they can also cause external effects. Therefore, it is not clear whether a multiplier level that is optimal for a region would also be optimal for the whole system. If not, then the question arises whether the individual multipliers should rather be set by a superordinate regulator. These questions can be answered by analysing the external effects of multipliers in a more realistic model setting that considers the spatial characteristics of the gas network.

In order to identify the internal and external effects of multipliers in different regions in the EU, and to provide insight into optimal multiplier levels, we use for our analysis the numerical simulation model, TIGER.[82] The TIGER model optimises the gas dispatch in Europe under perfect foresight and perfect competition. We extend the model by including the costs of capacity booking and specifying the necessary restrictions. The model has a monthly temporal resolution, where yearly, quarterly, and monthly capacity products are offered. Six regional clusters of countries are considered: Central Europe, British Isles, South East Europe, Italy, Iberia, and Baltics. The aggregation of countries takes into account the geographical location of individual countries, existence of interconnecting pipelines, and at what stage a country lies in the gas transport chain (i.e. transit, downstream or peripheral). We simulate the gas dispatch for the gas year of 2017–2018 and analyse and quantify the effects of the multipliers on infrastructure utilisation, prices, and welfare distribution.

We identify significant regional effects with regards to multipliers. Our analysis shows that in regions characterised by relatively flat gas demand profiles (such as Spain and Portugal), multipliers do not have notable effects, as LT capacities

[82] A detailed formulation of the model can be found in Lochner (2012).

are preferred irrespective of the multiplier levels. In contrast, in regions that have a highly volatile demand but limited supply flexibility via storages (e.g. Britain), multipliers can have a strong impact on the base and peak prices, as they determine the marginal supply costs. Therefore, when specifying multipliers in such regions, regulators would also have to consider the strong distributional effect on the allocation of consumer surplus between the base and peak consumers.

We find that adjusting multipliers in a region can cause external effects in other regions. Consumer surplus gains in transit regions (e.g. Central Europe) due to multipliers are passed on to regions that lie downstream (e.g. Italy). We show that downstream regions can influence the transit regions indirectly by affecting the storage utilisation in the transit regions. Peripheral regions (e.g. South East Europe), which receive their gas directly from the production regions, can also influence other regions by affecting procurement prices in the production regions. Because of those external effects of multipliers, we find that individually optimal multipliers do not lead to the maximum total EU consumer surplus. Despite that, when comparing the gains in consumer surplus from applying multipliers, individually optimal multipliers result in about 12% higher consumer surplus gains in the EU compared to an optimal uniform EU-wide multiplier level. Hence, the current EU regulation of specifying allowed multipliers in ranges instead of absolute values is appropriate and can increase the EU consumer surplus. However, we show that the surplus gains achieved by individually optimal multipliers are about 9% lower than the maximum achievable EU consumer surplus gains by multipliers. This indicates that letting national regulators set the multipliers may not lead to an EU optimum.

Our paper is related to two streams of literature. The first relevant literature stream includes the analysis or modelling of capacity bookings in the European gas markets. Keller et al. (2019) analyses historical capacity bookings in German gas market areas. Using historical data from the PRISMA capacity booking platform for the year 2016, the paper shows that network users make efficient booking decisions and choose transport alternatives with the lowest tariffs. Grimm et al. (2019) presents a mathematical framework depicting the entry-exit gas markets. The paper shows that, under perfect competition, the booking and nomination decisions can be analysed in a single level and that this aggregated market level has a unique equilibrium. Dueñas et al. (2015) develops a combined gas-electricity model, which simulates the gas procurement and capacity booking of a gas-fired generation plant under residual demand uncertainty. The analysis shows that the capacity booking behaviour of the individual generator is significantly affected by how risk-averse it is.

The second relevant stream of literature analyses gas markets using numerical simulations based on cost minimisation models. It is common within this literature stream to analyse the effects of various developments on the gas infrastructure, identify possible bottlenecks and simulate potential effects on prices. In this context, previous versions of the TIGER model are applied to address various questions (Dieckhöner, 2012, Dieckhöner et al., 2013, Lochner, 2011a,b, 2012).

Dieckhöner et al. (2013) for instance simulates the European gas dispatch under different scenarios and analyses the level of market integration and potential congestions. Using a similar model, Hauser et al. (2019) investigates whether increasing natural gas demand in the power sector could cause congestions in the German gas grid. Eser et al. (2019) combines a Monte-Carlo simulation model for annual gas sourcing with a cost minimisation model that optimises the detailed hourly gas dispatch.

The contribution of our paper with regards to the above-mentioned literature can be summarised as follows: The capacity booking system and the effects of multipliers have not been yet analysed in the literature using numerical simulation models of gas dispatch. Thus, by integrating capacity booking into a cost minimisation model and simulating the European gas dispatch, we show that the level of multipliers can significantly impact infrastructure utilisation, prices, and welfare distribution. We identify and differentiate between internal and external effects of multipliers over a range of regional clusters, and provide insight into those effects that influence the optimal multiplier levels in the EU.

5.2. Identifying the main drivers

When a region adjusts its multipliers, it can affect the gas dispatch, regional prices and welfare within that region. This is shown by Çam and Lencz (2021) using a stylised model containing one demand region and two periods. The paper also finds that optimal multiplier levels for maximising consumer surplus can vary depending on the storage and transport costs. In addition, demand structures among regions vary, which can also play an important role on the effects of multipliers. It is therefore natural to assume that different regions could be affected differently from multipliers and would have varying optimal levels of multipliers. However, it is not clear if individually optimal multipliers would also be optimal for the whole system, since multipliers can additionally cause external effects. This would imply that the adjustment of multipliers in one region can affect market results in other regions. In this chapter, building upon the theoretical findings of Çam and Lencz (2021), we extend the discussion on internal effects of multipliers by highlighting several aspects which were not considered in that paper. We then present some intuition on the potential external effects of multipliers.

5.2.1. Internal effects of multipliers

A multiplier value of 1 results in a pricing regime similar to commodity pricing. In this case, traders, who transport gas from one market area into another, would book a combination of ST capacities that would perfectly satisfy their demand profile and pay for the exact amount of volumes they transported.[83] Higher multipliers incentivise traders to avoid ST capacities, encouraging them to book yearly (LT) capacities and flatten their winter and summer transports by increasingly storing gas in the demand regions. When multipliers reach a certain threshold, traders book solely LT capacity and behave as being exposed to a capacity pricing regime, irrespective of the costs of LT capacity and storage. Applying the finding from Çam and Lencz (2021) to the twelve-period model used in our current analysis, such multipliers are found to be 4 for quarterly and 12 for monthly capacity (see Supplementary Material D.1.1 for proof).

Çam and Lencz (2021) shows that, due to the relative costs of transmission and storage, in the majority of the situations already lower multipliers can induce a capacity pricing regime. This means that traders would book only LT capacity to cover their yearly peak demand, resulting in them paying for the capacity rather than the energy.[84] When only LT capacity is booked, increasing the multipliers does not affect market results. This is because LT tariffs are not affected since TSO revenues remain unchanged.

According to Çam and Lencz (2021), multipliers also affect gas prices, which in turn impact overall consumer surplus as well as its distribution among base and peak consumers. In this case, the minimum demand level is assigned to base consumers, which is constant throughout the considered time periods. Any demand that is above this minimum level is then defined as peak demand and is attributed to peak consumers.

When storage capacity is abundant, gas prices are affected by the LT transmission and storage tariffs. When multipliers are increased, TSOs can charge higher tariffs for ST capacity, allowing them to reduce the price for LT capacity. Thereby, gas prices decrease such that peak and base consumers profit. However, this effect is counteracted by bookings shifting from ST towards LT capacity. When supply flexibility from storages is restricted, Çam and Lencz (2021) finds that peak prices are determined by the price for short-term capacity. Hence, with increasing multipliers, peak prices increase. Off-peak prices on the other hand are found to decrease, reinforcing the distributional effect between base and peak consumers.

The above-mentioned findings are derived from the analysis presented in Çam and Lencz (2021), which uses a stylised theoretical model with two regions and two time periods. However, additional internal effects with respect to multipliers are to be expected in a more complex setting with multiple regions, multiple time periods, and more than one type of ST capacity product—such as in the case of the EU. It is to be expected that in regions with relatively flat demand profiles

[83] Multiplier levels below 1 would neither change the optimisation rationale of the traders nor the market results (see Çam and Lencz (2021) for a more detailed discussion). For this reason, and since the EU regulation NC TAR 2017 does not allow for multipliers below 1, the minimum multiplier value considered in this paper is equal to 1.

[84] For a more detailed discussion of capacity pricing and commodity pricing aspects of multipliers, please see Çam and Lencz (2021).

comparably less ST capacities would be booked, making the effect of multipliers limited. In contrast, in regions with volatile demand structures, multipliers would have a much higher impact on the proportion of bookings and, consequently, on the prices and welfare.

An additional effect that would be observed in a more realistic setting would be related to the costs of gas storages. Çam and Lencz (2021) assumes constant storage costs for the stylised model. In reality, gas storages have varying operating costs depending on their physical characteristics (Neumann and Zachmann, 2009). With higher multipliers, as more of the storage capacities are used, the more expensive storage types would be utilised. This means that marginal cost of storage would increase, causing higher temporal spreads in regional prices. While increased spreads would not affect the overall costs for base consumers, peak consumers would end up paying more.

The fact that storage capacities as well as the injection/withdrawal rates are limited in reality, which were assumed to be unlimited in Çam and Lencz (2021), can result in multipliers causing additional effects. When supply flexibility from storage capacities is exhausted, the seasonal spread in regional prices is not defined by the cost of storage, but by the cost of importing gas in the short term, which increases the temporal spread in prices even further. In such a setting, booking solely LT capacity while letting some seasonal capacity remain unused[85] can be optimal when multipliers reach a certain threshold (see Supplementary Material D.1.2 for proof)—an effect which cannot be observed in the simplified two-period model with unlimited capacities. Overall, as outlined above, additional internal effects due to multipliers would be observed in a more complex setting.

5.2.2. External effects of multipliers

When regions adjust their multiplier levels they may also affect other regions. To what extent a multiplier adjustment would have an external effect largely depends on how a region is located along the gas transport chain. In this context, a region can be classified into one of the four region types, as schematically shown in Figure 5.1: production, transit, downstream, and peripheral. Gas is transported from a production region (e.g. Russia) through a transit region (e.g. Central Europe) to downstream regions (e.g. Italy). Countries which do not lie downstream of a transit region but receive their gas directly from the production region can be referred to as peripheral regions (e.g. Baltic countries). While a transit region imports and re-exports substantial amount of gas volumes, downstream and peripheral regions import but do not re-export significant volumes.

[85]Letting some booked capacity remain unused is also referred to as capacity wasting.

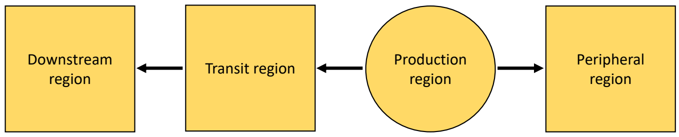


Figure 5.1.: Schematic representation of the types of regions

When traders transport gas through several borders, tariffs are accumulated, which is commonly referred to as tariff pancaking (EY and REKK, 2018). Due to pancaking, downstream regions are generally affected by the tariff structures and the ensuing effects over the whole transport chain. Therefore, price and welfare effects caused by changes in multiplier levels in transit regions would also likely be passed on to the connected downstream regions. Additionally, traders who want their gas to be shipped from a transit region to a downstream region have to procure capacity for exiting the transit region. Increasing multipliers in the transit region would therefore incentivise traders to book long-term and to flatten transports from the transit region to downstream regions. As a result, at what levels the multipliers are set in the transit regions can create direct external effects on the downstream regions. In contrast, any changes in multiplier levels in the downstream or peripheral regions would not have direct external effects on other regions, as changes in tariffs are not passed through to other regions. Nevertheless, it is possible that multiplier levels in any region can also cause external effects in other regions indirectly. By influencing the seasonal gas procurement patterns, multipliers can affect the temporal spreads in the regions where gas is imported from, as also shown in Çam and Lencz (2021). This would in turn influence the price levels in other regions which import gas from the same region.

Due to the above-mentioned internal and external effects, it is likely that different regions in the EU could be affected differently from multipliers, hence having varying optimal levels of multipliers. Then the question arises whether the individually optimal multipliers would also be optimal for the whole EU, since countries individually specifying multipliers could cause externalities in other countries. In this paper, we aim to address these questions with the help of a gas dispatch optimisation model.

5.3. Methodology

5.3.1. Model

To analyse the effects of multipliers in the EU we apply and extend the TIGER model developed at the Institute of Energy Economics (EWI) at the University of Cologne.[86] TIGER simulates the gas dispatch in Europe in a setting with perfect competition and perfect foresight. The model is formulated as a linear

[86] For a comprehensive formulation of the model see Lochner (2012).

optimisation problem with the objective function of minimising total system costs. It models the producers, consumers, traders and storage operators and includes the production capacities, demand regions, pipeline network, gas storages and LNG terminals.

The TIGER model is extended by including the costs of capacity booking in the objective function and specifying the necessary restrictions. A complete notation of the model extension is presented in Table 5.1.

Table 5.1.: Notation used in the TIGER model extension

Sets	$t \in T$	Points in time
	$i, j \in N$	Nodes in the pipeline network
	$p \in P$	Capacity products (defined by duration, start and end date)
Parameters	m_p	Tariff multiplier per capacity product
	$\tau_{i,j}$	Base entry/exit tariff
Variables	$C^{Tra}_{t,i,j,p}$	TSO revenue (Gas transport costs)
	$CB_{t,i,j,p}$	Booked capacities per product type
	$TR^{CB}_{t,i,j,p}$	Volumes transported per product type
	$TR_{t,i,j}$	Total volumes transported
	$CB^{Map}_{i,j,p}$	Capacity booking mapping parameter

The objective function corresponds to minimisation of total costs (C^{Tot}). Total costs are equal to the sum of production costs (C^{Pro}), transport costs (C^{Tra}), storage costs (C^{Sto}) and costs associated with LNG imports and regasification (C^{LNG}).

$$\min C^{Tot} = C^{Pro} + C^{Tra} + C^{Sto} + C^{LNG} \tag{5.1}$$

Gas transport costs at time t from node i to j for a particular capacity product p equal the level of booked capacities $CB_{t,i,j,p}$ multiplied with the base entry-exit tariff $\tau_{i,j}$ and the corresponding product multiplier m_p. Like in the EU, traders have to procure entry and exit capacity when transporting gas between market areas where entry-exit tariffs are applied.[87] Furthermore, we assume

[87] In the EU gas markets, traders are able to trade booked capacities in secondary markets. We assume in our analysis these secondary markets to be perfect. Therefore, under the model assumption of perfect foresight, the total booked capacities of individual traders would be identical to the booked capacities of a single competitive trader who faces the cumulative demand of all these traders. For a detailed discussion of secondary markets see Çam and Lencz (2021).

storage operators to be fully exempt from transmission tariffs when withdrawing or injecting gas in the transmission network.[88]

$$C^{Tra}_{t,i,j,p} = CB_{t,i,j,p} \cdot \tau_{i,j} \cdot m_p \tag{5.2}$$

TSOs are regulated entities and are allowed certain revenue caps. If adjusting the multipliers causes the revenues of a TSO to change, then the TSO would adjust the entry-exit tariffs accordingly to reach the same revenue cap. This fact is considered in our analysis. As each TSO's revenue should be independent from the multipliers applied, the base entry-exit tariff $\tau_{i,j}$ has to be adjusted such that a TSO's revenue (C^{Tra}) for each entry-exit point remains constant. This results in a quadratic function that cannot be solved in a linear model. Therefore, an iterative approach is applied to solve the model. In the first run, the $\tau_{i,j}$ is kept constant, resulting in increased TSO revenue for high multipliers. In the next iteration $\tau_{i,j}$ is adjusted in order to reach the intended TSO revenue for each multiplier level. As the adjusted tariff levels may result in an adjusted booking behaviour, the procedure is repeated until the revenues of all TSOs equal the intended individual levels.[89]

Booked capacities at each entry-exit pipeline are required to be greater than or equal to the transported volumes associated with the particular capacity product (Equation 5.3). Each capacity product (e.g. quarterly capacity for October, November and December) is valid only in its dedicated time period. (e.g. $t = 1, 2, 3$). Therefore, for the model with monthly resolution, one yearly, four quarterly and twelve monthly capacity products are offered for each entry-exit point.

$$CB_{t,i,j,p} \geq TR^{CB}_{t,i,j,p} \tag{5.3}$$

To ensure that each capacity booking is booked with the same level of capacity for the whole period it is valid in, a mapping equation is introduced as in Equation 5.4. This equation forces the booked capacities ($CB_{t,i,j,p}$) to be equal to the same value for each t it is valid in.

$$CB_{t,i,j,p} = cb^{Map}_{i,j,p} \tag{5.4}$$

[88] Storages are commonly exempted from transmission tariffs in the EU to a varying extent with the goal of inducing positive externalities such as reducing pipeline investment costs and increasing security of supply (ACER, 2019a). For example, several EU countries grant full exemption (e.g. Spain and Denmark). Storages are exempted by at least 50% due to NC TAR regulation in other countries; however, most countries apply higher exemptions (ENTSOG, 2019).

[89] Due to the convexity of the problem the converged solution is a global optimum.

Finally, the physically transported volumes on a pipeline must be equal to the sum of flows per capacity products.

$$TR_{t,i,j} = \sum_{p} TR^{CB}_{t,i,j,p} \tag{5.5}$$

5.3.2. Assumptions and data

For the purposes of this paper, the TIGER model is adjusted with regards to its spatial resolution where six regions are considered in order to be able to identify robust regional effects. The regional aggregation takes into account the geographical location of individual countries, existence of pipelines between them and whether a country is transit, downstream or peripheral. A transit country imports gas from a production region and re-exports significant volumes of gas to a downstream region. A downstream country imports from the transit region but does not re-export significant volumes. A peripheral country imports directly from the production region, but does not import significant volumes from a transit region and also does not re-export. Hence, despite the lower spatial resolution, the aggregation aims to represent the inter-regional gas flow patterns in a realistic manner. The spatial structure of the model as well as the considered regions and the countries they include can be seen in Figure 5.2.

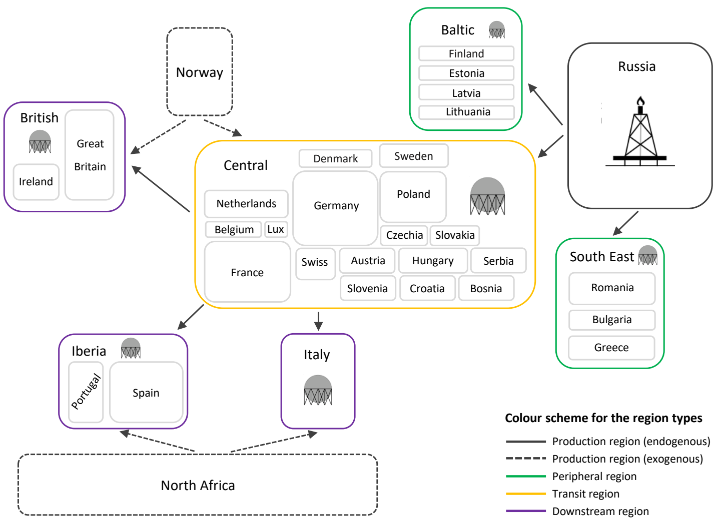


Figure 5.2.: Schematic representation of the spatial model structure

The transit Central region receives gas from the Norwegian and Russian production regions and can transport gas to southern downstream regions such as Italy and Iberia. Those regions also receive gas over North Africa. The downstream British region is connected to Norway and the Central region. The peripheral Baltic and the South East regions receive pipeline gas only from the Russian production region. Furthermore, all demand regions can import gas through their LNG regasification terminals. All demand regions have gas storage as well.

The model covers the historical gas year of 2017–2018, which starts on 1. October 2017 and ends on 30. September 2018. The gas year of 2017–2018 is chosen due to being the most recent gas year with publicly available data at the time of our analysis.[90] The model has a monthly temporal resolution. Correspondingly, yearly, quarterly and monthly capacity products are offered in the model. We assume that traders book their capacity in the analysed year. Historical capacity bookings are not considered, which allows us to assess the effects of multipliers more generally.[91]

The existing pipeline network, storages and LNG import capacities of 2018 are considered. The pipelines connecting individual regions are assigned their historical capacities based on TSO information and ENTSOG data for pipelines (ENTSOG, 2019). Within regions, pipeline capacities are assumed to be not restricted.[92]

Storage data, such as maximum storage volume as well as maximum injection and withdrawal rates for all storages in Europe, is based on Gas Infrastructure Europe (GIE, 2018) as well as storage operators' data. Similarly, data for LNG import terminals are obtained from ENTSOG and GIE LNG map (GIE, 2019). Thereby, LNG import, regasification and storage capacities are considered. The costs for storing gas are based on several studies (Enervis, 2012, Le Fevre, 2013, Redpoint, 2012) and consider the cost variation among different types of storages. We assume linear increasing marginal costs for storages, implementing it into the model as a step-wise linear function. Tariffs for the entry-exit zones are historical values observed in 2018 and are acquired from ACER (2019a).

Gas demand is assumed to be perfectly inelastic and is specified as an exogenous parameter. Historical country-level consumption data for the analysed period is used.[93] The Russian production region is the only flexible gas producer in the model. The Russian supply function to Europe is assumed to be linear increasing and is integrated into the model as a step-wise linear function.[94] Annual production capacities for other producers are assumed to be equal to

[90] The methodology is nevertheless not only applicable to different gas years but can also consider multiple consecutive years. Optimising multiple consecutive years would not change the rationale of the model since long-term capacity booking decisions are made on a yearly scale.

[91] This situation will be more prevalent from the year 2035 onward when historical long-term capacity bookings are almost completely expired (ACER, 2020b).

[92] The majority of the interconnection points in the EU are physically not congested, making this assumption plausible. According to ACER (2020a), physical congestion was likely to have happened in 2019 only in the 7 interconnection points among the 239 interconnection points considered in the study.

[93] Consumption data is sourced from EUROSTAT and websites of TSOs.

[94] The cost function is calibrated with respect to historical import volumes and prices and implicitly considers the transmission costs to Ukraine and Belarus. See Supplementary Material D.2 for the reference case and model validation.

their historical production levels observed in 2018 (BP, 2019) and are specified as exogenous parameters.

The model considers a simplified LNG supply structure due to several reasons. The previously explained iterative approach to have constant TSO revenues requires yearly import and export levels to be unaffected by changes in multipliers, since otherwise TSO revenues would not converge. If LNG provision would be modelled as in the case of Russian supply, the level of LNG and Russian supply would be affected by multiplier levels. This would in turn result in yearly import and export levels to vary and prevent the model results to converge. Therefore, LNG imports are modelled in the following manner: While yearly LNG imports are fixed to historical levels, LNG imports are allowed to be shifted within the year. For example, if high multipliers incentivise flatter pipeline import profiles, then LNG imports can be shifted to months with high gas demand. Such shifts of LNG imports are associated with costs. Hence, the stronger the deviation from the historical import profile, the higher the associated costs.

5.4. Results

In this section, we investigate the internal and external effects of multipliers. For this purpose, we apply the model presented in Chapter 5.3 and optimise the gas dispatch with different multiplier levels. We chose ten multiplier pairs in the range between commodity and capacity pricing (i.e., $m1$, $m2$, ..., $m10$). We set the starting multiplier pair ($m1$) to 1 for monthly and quarterly bookings. This setting induces commodity pricing (see Chapter 5.2.1). The $m2$ level is specified manually as 1.03 for the quarterly product and as 1.07 for the monthly product. The remaining quarterly and monthly multiplier pairs ($m3$ to $m10$) used in this analysis are derived with an exponential function.[95] Applying the exponential function for the highest multiplier pair ($m10$) gives multipliers of 5.42 for the quarterly product and 12.04 for the monthly product. These values exceed the threshold multipliers of 4 for the quarterly capacity product and 12 for the monthly capacity product which induces capacity pricing (see Chapter 5.2.1). The chosen multiplier setting allows to analyse a realistic range of the currently applied multiplier levels in the EU while also including the extreme levels that per definition induce commodity or capacity pricing. However, the exact values and the function form applied to derive the multiplier pairs are chosen arbitrarily.

As default multipliers, we use multipliers that are representative of the EU average and coincide with the multipliers applied in Germany (i.e., $m4$) (ENTSOG, 2018). Note, that from 01.01.2019 onward the EU regulation 2017/459 limits quarterly and monthly multipliers to 1.5. Therefore, within the set of analysed multipliers (i.e., $m1$-$m10$), the $m5$ level represents the highest multipliers which comply with EU regulation 2017/459.

[95] The formula used for deriving the multiplier pairs is as follows: $m_n = (m_{n-1})1.88 + 1$ for $n \geq 3$, where n is the multiplier pair number $n \in \{1, 2, ..., 10\}$.

Table 5.2.: The chosen multiplier levels for the analysis

	Quarterly	Monthly
$m1$ (commodity pricing)	1.00	1.00
$m2$	1.03	1.07
$m3$	1.05	1.13
$m4$ (default)	1.10	1.25
$m5$	1.19	1.47
$m6$	1.35	1.88
$m7$	1.66	2.66
$m8$	2.25	4.12
$m9$	3.35	6.87
$m10$ (capacity pricing)	5.42	12.04

In order for the results to have explanatory power, the model is first validated comparing the simulated prices, import volumes, and storage utilisation with the historical values observed over the considered time period. For this purpose, uniform multipliers equal to the default BEATE levels are assumed for the whole EU. Since many countries in the EU have multipliers similar to the BEATE levels, this is a realistic approximation. Results for model validation are presented in Supplementary Material D.2.

If a region individually adjusts its multipliers, it induces internal effects in the region itself. However, as highlighted in Chapter 5.2, it is possible for it to cause external effects on other regions. In order to identify those internal and external effects in this section we first consider a case where regions individually and independently adjust their own multiplier levels.

5.4.1. Internal effects

In a first step we investigate the internal effects of multipliers. For this purpose, we vary the multipliers in each of the six regions individually while keeping the multipliers in the other regions constant.[96] The internal effects in each region on capacity bookings, infrastructure utilisation, prices and consumer surplus are analysed.

Capacity bookings

The change in the volumes of booked capacities with respect to varying multipliers in the considered regions is plotted in Figure 5.3. The absolute height of the bar charts represent the total booked capacities, corresponding to the sum of yearly,

[96] Multipliers are fixed to the default $m4$ level as this represents the average multipliers in the EU according to ENTSOG (2018).

quarterly and monthly bookings. With multipliers at the level of m1, the procurement of ST capacity (e.g., monthly) to transport a constant load throughout the year does not result in additional costs compared to the procurement of LT capacity. Hence, traders are indifferent between the different capacity products, when transporting a constant load traders. For a monthly fluctuating load, it is optimal to procure only the monthly capacity required. As a result, only monthly capacity is procured to transport all volumes. However, the use of LT bookings for constant transport volumes that are constant throughout the year would have resulted in the same cost and dispatch. As regions individually increase their multipliers, the proportion of ST bookings in these regions decreases, while the proportion of yearly bookings increases. This is as expected, since higher multipliers make ST capacities proportionally more expensive and incentivise the booking of LT capacities instead. It is also observed that when multipliers reach high enough levels, such as the $m6$ level in Central, they indirectly induce a capacity pricing regime and cause only LT capacities to be booked. The individual level of multipliers that induce capacity pricing differ among the regions. For example, while a higher multiplier level of $m9$ causes capacity pricing in South East, a lower level of $m5$ is enough to cause capacity pricing in the Baltic region and Italy. These findings are in line with the theoretical findings of Çam and Lencz (2021).

Note that in South East and the British region, traders waste LT capacity when multipliers reach $m8$ and $m9$, respectively. This is because in those regions traders cannot fully flatten their monthly imports due to limited storage capacities, resulting in some LT capacity to remain unused, i.e. to be wasted (shown with dashed lines in the figure). Hence, unlike the theoretical model used in Çam and Lencz (2021) with two time periods and unlimited storage capacities, capacity wasting can occur in a realistic setting with multiple time periods and limited storage capacities.

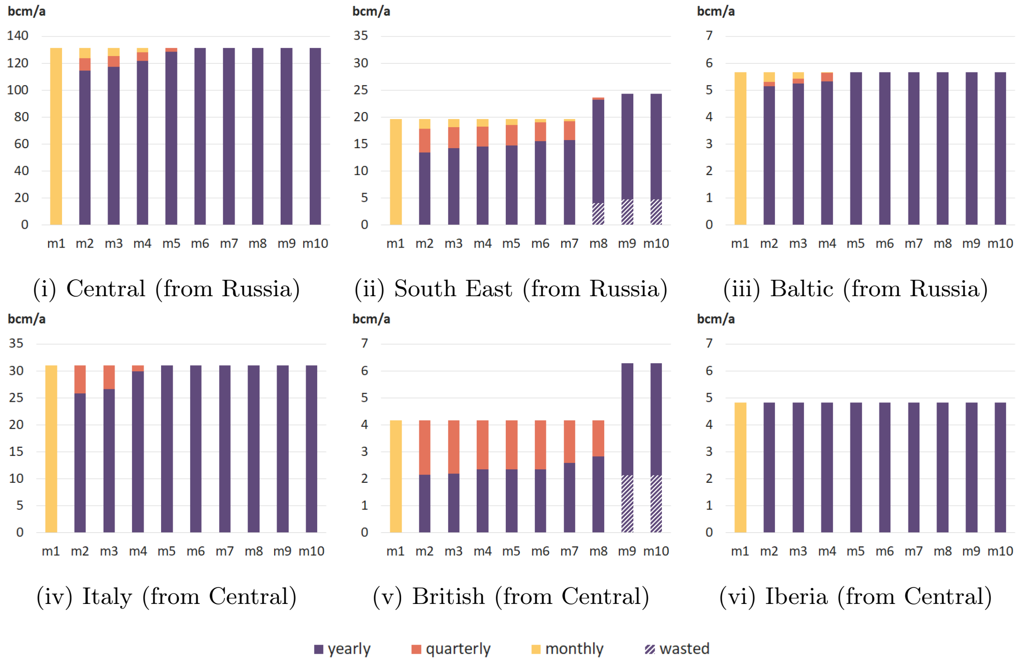


Figure 5.3.: Capacity bookings by run-time and wasted capacity in each region when adjusting their multipliers

In Iberia, as soon as multipliers reach $m2$, only yearly capacity is booked. This is due to two reasons. On the one hand, Iberia is a downstream region, connected to the transit region Central. Hence, it is still subject to the default multipliers ($m4$) set in Central. On the other hand, the seasonal demand profile is relatively flat (i.e. low winter-summer demand spread) such that even very low multipliers are sufficient to fully flatten the transports between Central to Iberia. Therefore, it can be deduced that the structure of the demand profile in a region can greatly influence how multipliers affect capacity booking.

Infrastructure utilisation

In Figure 5.4, the yearly stored gas volumes and the monthly peak import volumes per region are plotted against varying multiplier levels. The monthly peak import in a region corresponds to the highest monthly volumes imported by that region in the considered year. In all the analysed regions except Iberia, a general trend can be observed: As the multipliers increase, the transported peak volumes decrease. In parallel with this, the stored volumes increase. These findings are in line with Çam and Lencz (2021) and occur due to higher multipliers strengthening the capacity pricing aspect. In Iberia, infrastructure utilisation is not affected by multipliers since capacity booking is independent of multiplier levels, as shown previously.

Figure 5.4.: Relative change in import volumes in the peak-demand month and yearly storage volumes in each region when adjusting their multipliers

Prices

In a competitive market, regional prices are determined by marginal costs of gas provision. Çam and Lencz (2021) shows that average marginal costs of gas provision are equal to the costs of gas procurement plus the costs for long-term (i.e. yearly) import transmission capacity. Hence, when multipliers affect yearly import (entry-exit) transmission capacity tariffs they also influence the average prices in regions.

Model results on the effects of multipliers on prices are plotted in Figure 5.5 for each individual region. In all regions where both LT and ST products are booked (see Figure 5.3), increasing multipliers up to a sufficient level causes the average prices to decline. This is because increasing the multipliers allows TSOs to reduce the tariff for their LT product.

In South East and the British region, however, the average price levels remain constant after they reach their minimum, which is caused by the capacity wasting that occurs in these regions with high multipliers. In Iberia, as only LT capacities are booked irrespective of multiplier levels, no price effects are observed.

As can be seen in Figure 5.5, multipliers not only have an impact on the average price levels, but also affect the temporal price volatility i.e. the standard deviation of the prices. When flexibility from storage and LNG imports is not fully utilised, the maximum price spread is defined by the marginal costs of such flexibility in the respective region. We have shown previously (see Figure 5.4) that multipliers increase the volumes stored in storages. As more expensive storage capacities start being used, the regional prices in peak months increase because marginal costs of storage increase. Since the differences in marginal storage costs are limited, the effect on temporal spreads is less pronounced for regions where storage capacities are not fully utilised (i.e. Central, Italy, Baltic).

In contrast, in British and South East regions, flexibility from storage capacities as well as LNG is fully utilised when the multiplier level reaches $m4$ and $m6$, respectively. In these cases, the maximum price spread is determined by the marginal costs for ST (i.e. monthly) capacity. As increasing multipliers result in higher prices for monthly capacity bookings, the maximum price spread increases. This process stops as soon as booking yearly capacity—which is not subject to multipliers—gets cheaper than booking monthly capacity. For the British and South East regions this is the case when multipliers reach $m9$ and $m8$, respectively.

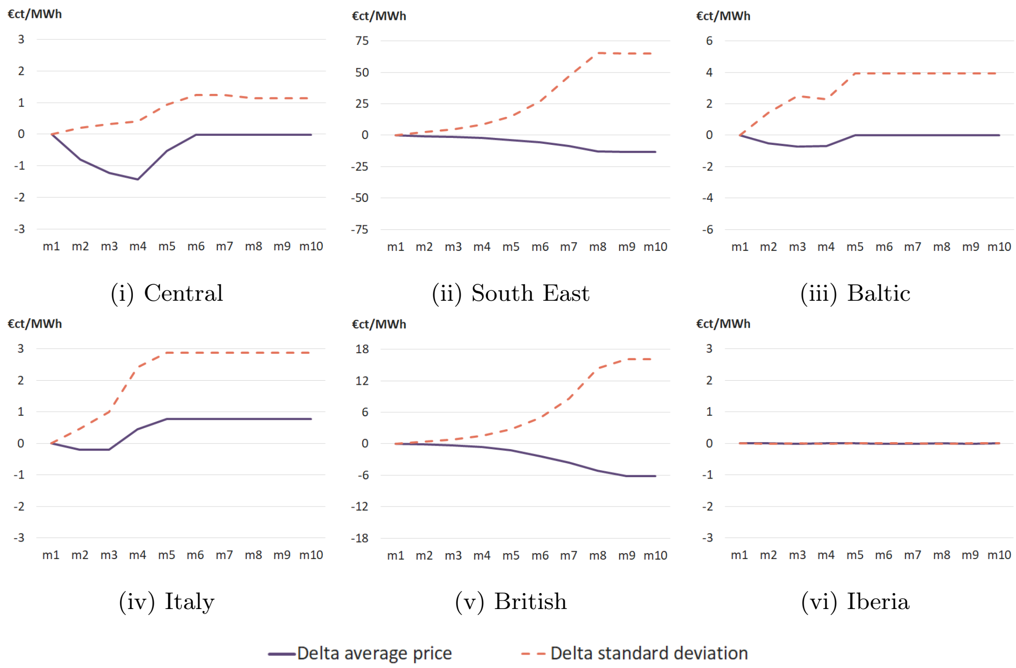


Figure 5.5.: Absolute change in the average price (i.e. delta LT tariff) with respect to m1 level and the absolute change in the standard deviation in each region when adjusting their multipliers individually

Multipliers also affect the regional price spreads, as the average regional price spread corresponds to the yearly transmission tariff. Therefore, multipliers that minimise the average price also minimise the average regional spread with respect to the region exporting gas. Furthermore, we find that higher multipliers increase the volatility in regional price spreads, thus, confirming the findings of Çam and Lencz (2021). We identify two effects which drive the volatility in regional spreads.

The price volatility in a region that increases its multipliers rises. At the same time, the increase in multipliers tends to decrease the temporal volatility in the exporting region. As a result, these two effects combined together amplify the volatility of the price spread between those two regions. A detailed analysis of the regional price spreads can be found in Supplementary Material D.3.

Consumer surplus

We have shown that multipliers affect the average price levels as well as the peak prices. As such, they directly affect the consumer surplus in the individual regions and how it is distributed between different types of consumers with varying demand patterns (i.e. base vs peak). In Figure 5.6, the change in consumer surplus in each region with respect to multipliers is plotted. The consumer surplus is defined relative to the $m1$ level. Since the gas demand is inelastic, consumer surplus corresponds to the change in prices multiplied with the demand. Further, we distinguish between base consumer surplus and the peak consumer surplus. Base consumer surplus corresponds to change in average prices multiplied by the base demand. Base demand is assumed to be constant throughout the year and equals the overall minimum monthly demand of a region. Any demand above this base level is then defined as peak demand. Thus, peak consumer surplus corresponds to the peak demand multiplied by the change in the corresponding prices.

Consumer surplus and its distribution between base and peak consumers are affected differently in each region with increasing multipliers, depending on which of the following three effects dominates:

- **Effect 1**: The first effect is the change in average prices due to tariff adjustment, which affects the overall consumer surplus. In this case, both base and peak consumers benefit if the tariffs are reduced or both consumer types lose if the tariffs are increased.
- **Effect 2**: The second effect is the increased spreads between off-peak and peak prices caused by higher storage utilisation. With higher storage utilisation, more expensive storages are used, which increase the spread between peak and off-peak prices. In this case, base consumers are not affected, while peak consumers lose.
- **Effect 3**: In case that flexibility from storage and LNG imports is exhausted, there exists a third effect: The prices in the peak periods are determined by the price of ST capacity, resulting in increased peak prices. Therefore, as multipliers increase, peak prices also increase, causing the peak consumer surplus to decrease.

In Central, the reduction in the average price causes both the base and peak consumer surplus to increase and reach a maximum at the multiplier level of $m4$ (Effect 1). Nevertheless, both peak and base consumer surplus decrease with higher multipliers as the LT tariff is increased due to the shift to LT capacity. Peak consumer surplus decreases additionally because of higher storage utilisation (Effect 2).

In the South East and Baltic regions, base consumers also increasingly benefit from the average price reduction with higher multipliers (Effect 1) while the peak consumers lose due to higher peak prices caused by increased storage utilisation (Effect 2). In South East, flexibility from storages is exhausted at $m6$ and from then onward Effect 3 dominates, causing a large decrease in the peak consumer surplus and reducing the overall consumer surplus substantially. In both the South East and Baltic regions, low multipliers ($m1$) maximise the overall consumer surplus, which is due to the relatively small size of those regions in terms of gas demand as well as their position as peripheral regions. When the two regions increase their imports in summer and decrease them in winter because of higher multipliers, prices in Russia are affected (lowering effect on winter prices and raising effect on summer prices). However, the transit Central region mitigates the effect on Russian prices almost fully when it exploits the lowered temporal Russian price spread. The mitigating effect is more pronounced since imports of the transit Central regions are five times higher than the sum of both peripheral regions' imports. Hence, Effect 2, which reduces peak consumer surplus, is reinforced such that optimal multipliers in the peripheral regions Baltic and South East are found to be low.

In Italy, the decrease in average prices causes a slight increase in the total consumer surplus, which reaches a maximum at the multiplier level of $m2$. Due to the peak price effect caused by higher storage utilisation (Effect 2), peak consumer surplus decline is steeper than the decline in base consumer surplus. Effect 2 is reinforced by Italy's relative position as a downstream region from Central. As Italy flattens its import profile from Central, gas storage is shifted from Central to Italy, reducing the summer-winter price spread in Central. In response, Central adjusts its import behaviour and imports more gas during winter. This mitigates the effect on the temporal price spread in Central, which further causes increased storage utilisation in Italy.

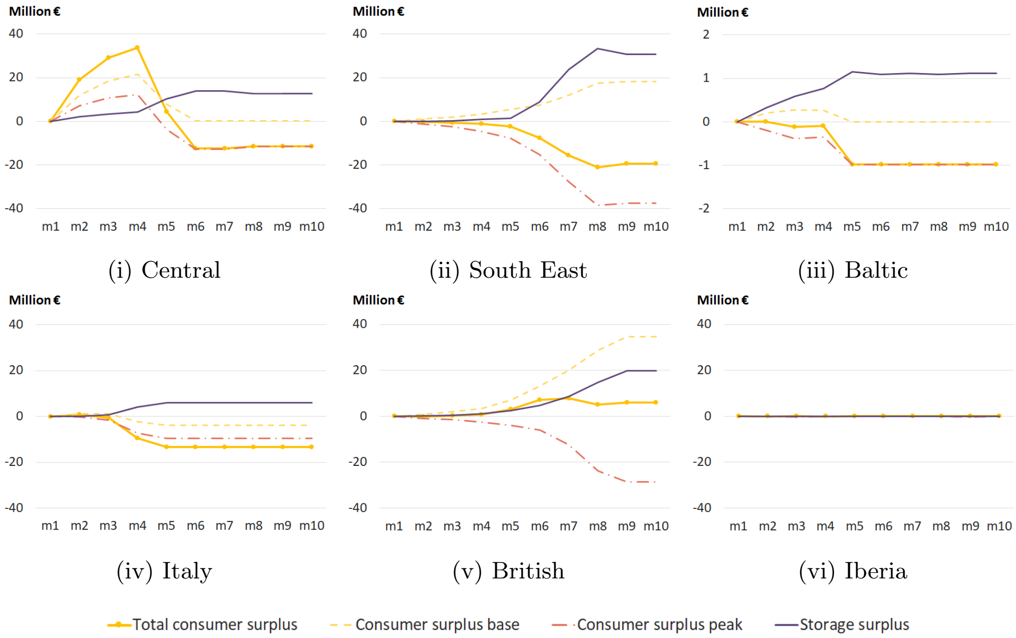


Figure 5.6.: Consumer and storage operator surplus in each region when adjusting their multipliers individually

In the British region, the effects are similar to those observed in South East. However, in contrast to South East, import tariffs can be reduced to a larger extent, such that Effect 1 dominates and total consumer surplus is maximised at $m7$. This is because, irrespective of multipliers, imports occur predominantly in winter. As the TSO revenue is kept constant, LT tariffs can be reduced significantly, limiting the increases in ST tariffs. In Iberia, the consumer surplus is unaffected since only LT capacity is booked irrespective of the multiplier level.

5.4.2. External effects

As highlighted in Chapter 5.2, if a region individually adjusts its multipliers, it is possible for it to also cause external effects on other regions. Those external effects can be direct or indirect, and depend on whether the regions that adjust their multipliers are transit, downstream or peripheral.

Transit region adjusts its multipliers

In this case, the transit Central region is allowed to vary its multipliers while all the other regions have unchanged multipliers equal to the default ($m4$) levels. Adjusting multipliers in the Central region has direct effects on the regions that are connected and lie downstream such as Iberia, Italy and the British region.

Figure 5.7 shows the changes in consumer surplus and storage surplus in these regions with respect to multiplier levels in the Central region.

The first direct external effect arises from the change in average prices in Central which is passed on to the downstream regions (arising from Effect 1 in Central). This external effect can be clearly observed in Iberia, where minimum average prices in Central for $m4$ also lead to lowest prices (i.e. highest consumer surplus) in Iberia.

For Italy and the British region, changes in multipliers also impact the booking behaviour and the gas dispatch for transports from Central, which induces additional external effects in the downstream peripheral regions. These effects depend on which of the previously discussed three effects ensue and dominate.

In Italy, the consumer surplus of peak consumers falls significantly with increasing multipliers. This is because higher multipliers for exporting gas from Central to Italy incentivise the flattening of transports from Central to Italy. The required utilisation of more expensive storages in Italy increases the peak prices in Italy, reducing the peak consumer surplus (Effect 2). In combination, the sum of the two external effects (Effect 1 and Effect 2) is highest for $m3$.

Similarly, when transporting gas from Central to the British region, traders are also incentivised to flatten transports with higher multipliers. In the case of the British region, as flexibility from storage and LNG is limited, a full flattening of transports is not possible. Hence, in peak periods the cost of ST capacity determines the prices, causing significant decline in the peak consumer surplus (Effect 3). Similar to the individual adjustment case, base consumer surplus increases due to tariff reduction (Effect 1). Overall, the highest positive external effect from Central on the British region arises for $m3$ due to combination of Effect 1 and Effect 3.

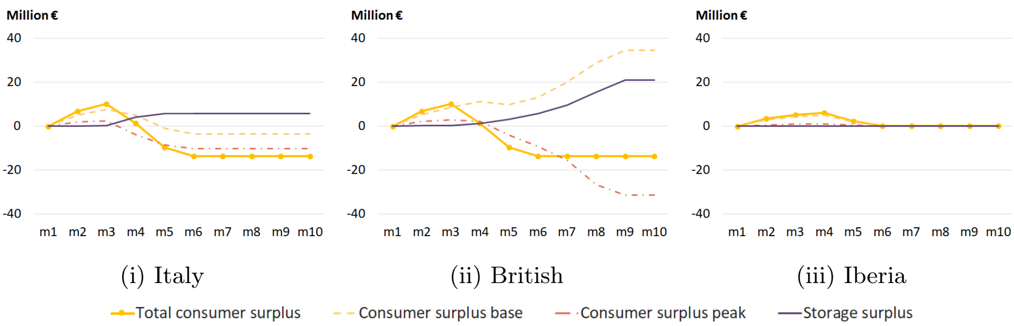


Figure 5.7.: Changes in the consumer and storage operator surplus in the regions which lie downstream of Central when Central adjusts its multipliers: (a) Italy, (b) British, (c) Iberia

Adjusting multipliers in the transit Central region also induces indirect external effects on the peripheral regions which are not directly connected with it such as the South East and the Baltic regions. Figure 5.8 shows the development of consumer and storage surplus in South East and Baltic with respect to changing multipliers in Central. Increasing the multipliers in Central causes the spread between peak and off-peak procurement prices in the Russian production region to decrease, i.e. off-peak prices increase and peak prices decrease. As a result, in the South East and Baltic regions, peak consumer surplus increases.[97]

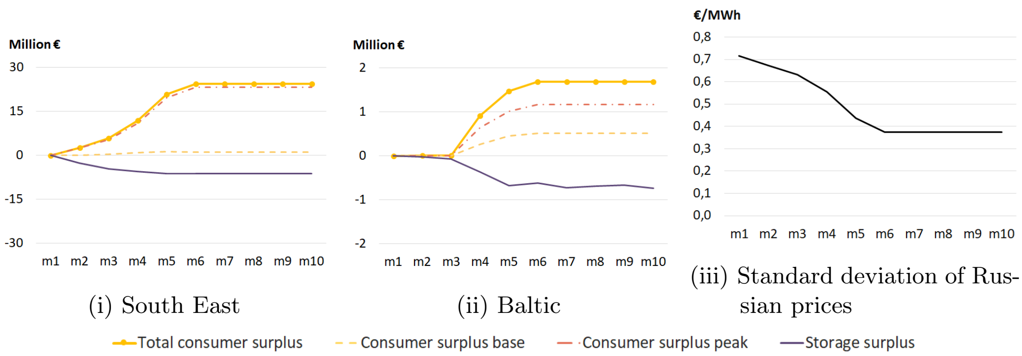


Figure 5.8.: Changes in the consumer and storage operator surplus in the regions which are not directly connected to Central when Central adjusts its multipliers: (a) South East, (b) Baltic, and (c) the corresponding development of the standard deviation of Russian prices

Downstream or peripheral region adjusts its multipliers

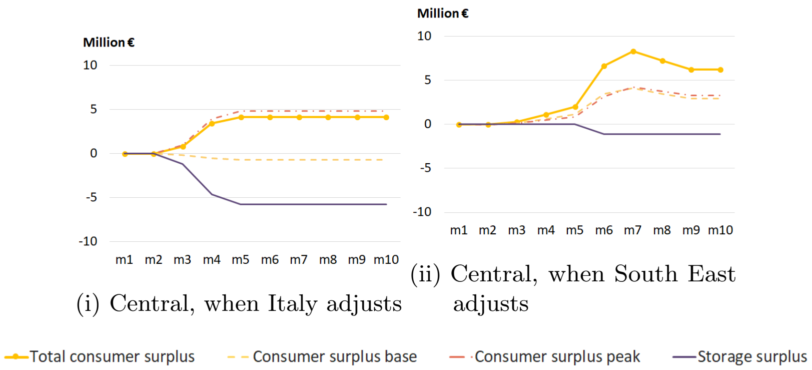
When downstream or peripheral regions adjust their multipliers, they can also cause external effects on other regions. Figure 5.9 shows the changes in storage and consumer surplus in Central with respect to the multiplier levels in Italy and South East, respectively. In the case of Italy, multipliers in Italy are varied while other regions have the default multiplier level. Similarly, in the case of South East, only the multipliers in South East are varied while other regions have the default multiplier level. In both cases, we observe significant impact on the Central region.

In the case of adjustments in Italy, higher storage utilisation in Italy due to increased multipliers results in storages in Central to be utilised less. As a result, peak prices in Central decrease and peak consumer surplus increases consecutively.

The overall impact from changes in the multipliers in South East on the consumer surplus in Central arises from a combination of two specific effects: Increasing the multipliers in South East causes the spread between peak and

[97]Due to cheaper procurement prices during the peak period, more ST products are booked in South East and Baltic regions to transport Russian gas to cover the peak demand. The increased share of ST bookings allows the TSOs to slightly reduce their transport tariffs, such that the overall prices in the South East and Baltic regions sightly decrease, benefiting both the peak consumers and the base consumers. Here, this effect can be more easily seen in the case of the Baltic region.

off-peak procurement prices in the Russian production region to decrease, i.e. off-peak prices increase and peak prices decrease. At the same time, due to cheaper procurement prices during the peak period, more ST products are booked in Central to transport Russian gas to cover the peak demand. Increased amount of ST bookings allows the TSO to reduce the transport tariffs. Consequently, overall prices in Central decrease, benefiting both the peak consumers and the base consumers.



(i) Central, when Italy adjusts

(ii) Central, when South East adjusts

Figure 5.9.: Changes in the consumer and storage operator surplus in Central when (a) Italy adjusts its multipliers, (b) when South East adjusts its multipliers

The external effects of multiplier adjustments in Italy and South East on other regions except Central are found to be very small. Any multiplier adjustment in the British region is found to have negligible impact on other regions because a large share of gas consumption is produced within the region or imported by LNG. Baltic region is found to cause similar external effects as the other peripheral region South East, albeit at a much smaller scale, because the imported volumes are comparably low. Iberia, having shown that no internal effects ensue with respect to multipliers, does not cause any external effects either. Those cases are not shown in this section explicitly but can be found in Supplementary Material D.4, where the external effects of multiplier adjustments of all the regions are presented.

5.4.3. Overall distributional effects

We have shown that multipliers can cause both significant internal and external effects in various regions in the EU by influencing the price levels and the consumer surplus. Higher multipliers were also shown to cause increased storage utilisation (storage surplus), resulting in flattened import profiles from the Russian production region. These effects would also have an impact on the producer surplus and the trader surplus. As such, multipliers would influence the welfare and its distribution in the EU and in the production regions.

In order to clearly show the overall distributional effects of multipliers in the EU and in the production regions, we assume in a first step that the multipliers are specified in the EU by a superordinate regulator and every region has the same uniform multiplier level. In Figure 5.10, the changes in surplus of the consumers, producers, traders and storage operators as well as the change in overall welfare with increasing multipliers are plotted. All the values are defined and plotted in relation to the case where multipliers are equal to 1 ($m1$). Hence, at $m1$ the change in surpluses and welfare are zero. It can be seen that the overall consumer surplus increases significantly with higher multiplier levels and reaches a maximum of about 82 million EUR at $m4$. Peak-load consumers receive a much smaller share (31% at $m4$) of this additional consumer surplus compared to base-load consumers (69% at $m4$).

Producer surplus decreases substantially with increasing multipliers. The reason for that is the rise in yearly bookings and a corresponding decrease in purchased volumes from Russia in the peak periods. The producer surplus decreases as the purchased volumes in the peak and off-peak periods converge. At the consumer-surplus-maximising multiplier level of $m4$, Russian producers incur a loss of 69 million EUR compared to the $m1$ level.

Storage operators have surplus gains with higher multipliers due to increased storage utilisation, as more of the expensive storages are used that set the price of storage. At $m4$, the storage operator surplus equals 5 million EUR. When multipliers reach $m6$ and storages are fully utilised in the British and South East region, storage operators can charge bottleneck prices, increasing the storage operator surplus up to 77 million EUR for multiplier levels of $m9$ and $m10$, almost 15 times greater than the surplus observed with $m4$.

Trader surplus equals the revenue from selling gas to consumers minus the costs of gas provision, i.e., the costs for gas procurement, transport and storage. When the uniform multipliers increase to $m4$ levels, traders make less profit (-43 million EUR) as consumer prices decrease while at the same time booking costs remain constant. For higher multipliers, trader surplus increases again. This happens mainly due to increased consumer price levels. In addition to the consumer price effect, traders profit from lower gas procurement costs but bear higher costs for storing natural gas. Those two effects largely cancel each other out.

Welfare is defined as the sum of all surpluses and is highest for $m1$. Higher multipliers increase the distorting effect of transmission tariffs, causing the gas dispatch to further deviate from an optimal dispatch that is based on short-run marginal costs, as was also shown in Çam and Lencz (2021). Higher multipliers reduce welfare by causing additional costs, which occur as a result of two opposing effects. On the one hand, total costs of gas production decrease as gas is produced more evenly. On the other hand, total costs of storing gas increase. However, as the increase in storage costs is higher than the decrease in production costs, welfare declines with increasing multipliers. For multipliers higher than $m6$, welfare becomes mostly independent from increases in multipliers, as traders start to behave as being subject to capacity pricing in an increasing number of regions as shown previously, such that increases in multipliers do not affect procurement or storage volumes.

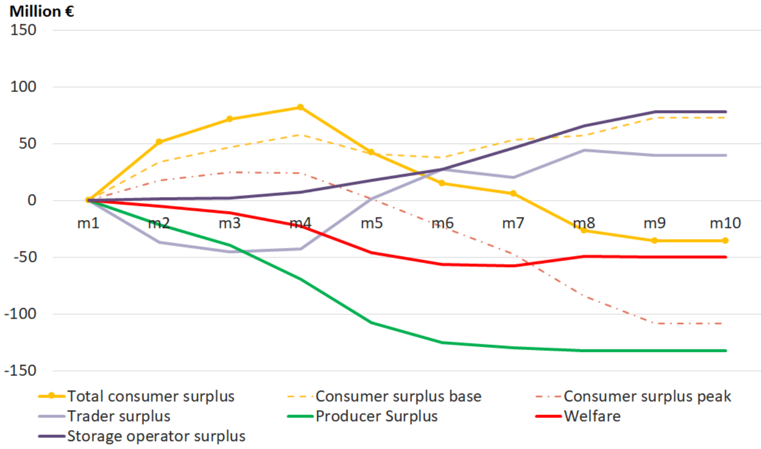


Figure 5.10.: Changes in the consumer, producer, trader, and storage surplus and welfare with respect to multipliers in the EU

5.4.4. Comparing different optimal multiplier levels

A major research question of this paper is whether multipliers in the EU should be set by a superordinate regulator or whether individually optimal multipliers can lead to a joint (i.e. EU-wide) optimum. In this part of our analysis we aim to answer those questions. To do so, we compare consumer surpluses for three cases: (1) EU-wide uniform optimal multiplier level, (2) individually optimal multipliers that maximise the consumer surpluses in the individual regions, and (3) multipliers for individual regions that lead to a joint optimum. Optimal multipliers in this context correspond to multipliers that maximise the consumer surplus.

From Chapter 5.4.3 we know that the EU-wide uniform multiplier level resulting in the highest consumer surplus is $m4$. Furthermore, we have shown previously in Chapter 5.4.1 that the individually optimal multiplier levels vary among the analysed regions. For Central, the optimal level was found to be $m4$ while for Italy $m2$ was shown to be optimal. In South East and Baltic regions, optimal multipliers should be as low as possible; namely equal to $m1$. In contrast, in the British region, multipliers as high as $m7$ were found to be optimal. In Iberia no effects with respect to multipliers were observed.

To find the multiplier levels resulting in the EU-wide joint optimum, we vary the multiplier levels of the four regions that were found to cause significant external effects (i.e. Central, South East, Baltic and Italy) in combination. With 4 regions and 10 multiplier levels, this corresponds to 10^4, namely, 10000 combinations.

Multiplier level in the British region is set to its individually optimal level of $m7$, while Iberia is set to the default level of $m4$. We find that individually optimal multipliers for Central and Italy also lead to the joint optimum. In contrast, the jointly optimal multiplier level for the peripheral regions, South East and Baltic, differ from their individually optimal levels and are found to be $m6$ and $m5$, respectively. The optimal multiplier levels in the three cases are summarised in Table 5.3.

Table 5.3.: Multiplier levels maximising consumer surplus

Region	Uniform multipliers	Individual optimum	Joint optimum
Central	$m4$	$m4$	$m4$
South East	$m4$	$m1$	$m6$
Baltic	$m4$	$m1$	$m5$
Italy	$m4$	$m2$	$m2$
British	$m4$	$m7$	$m7$
Iberia	$m4$	$m4$	$m4$

Figure 5.11 shows the corresponding change in consumer surplus for the optimal multiplier levels in the three cases. The delta consumer surplus is calculated relative to the consumer surplus resulting from uniform multipliers in all regions equal to $m1$. It can be seen that the uniform optimal multiplier level of $m4$ increases consumer surplus substantially compared to a uniform multiplier level of $m1$. The overall gains in consumer surplus amount to 82 million EUR. The optimal uniform multiplier level of $m4$ is also the individually optimal multiplier of the Central region. Since Central was shown to cause the highest internal and external effects, the uniform $m4$ level results in a significant increase in the EU-wide consumer surplus.

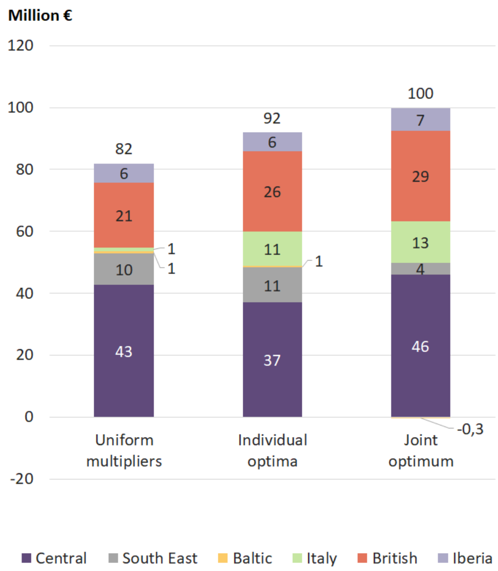


Figure 5.11.: Changes in regional consumer surplus with respect to how the multipliers are specified

When regions specify their individually optimal multipliers, total consumer surplus in the EU increases by 10 million EUR compared to the maximum consumer surplus achieved with uniform multipliers. Hence, the internal increase in consumer surplus by setting multipliers individually outweighs the negative external effects. However, consumers in Central are worse off. This occurs mainly because Italy sets lower multipliers, shifting storage utilisation from Italy to Central. As more expensive storages are utilised in Central, peak prices increase, reducing peak consumer surplus in Central.

In the case that regional regulators specify the multipliers in order to maximise the joint EU-wide consumer surplus, total consumer surplus increases by another 8 million EUR. The effect is limited, because for the majority of regions the individually and jointly optimal multiplier levels coincide. For Central, this occurs as downstream regions profit from lower average prices in Central such that both external and internal effects due to multipliers are highest for $m4$. For Italy, the positive internal effect on consumer surplus outweighs the negative impacts on the consumer surplus in Central. For British and Iberia, multipliers are found to have negligible external effects such that the individual and joint optima also coincide. Whereas, in South East and Baltic regions, jointly optimal multipliers ($m6$ and $m5$) diverge from the individually optimal multiplier level $m1$. Hence, the positive external effects from setting multipliers relatively high in South East and Baltic outweigh the negative internal effects. As outlined previously, this occurs because high multipliers in peripheral regions reduce the temporal price spread in the Russian production region, from which the other gas importing regions profit.

5.5. Discussion

5.5.1. Overall effects

Our analysis has shown several adverse impacts that multipliers can have on the overall gas dispatch. A multiplier of 1 is shown to be the optimal multiplier level that maximises overall welfare. This is not surprising, since higher multipliers reinforce the capacity pricing aspect and cause the gas dispatch to further deviate from an ideal dispatch that would be based on short-term marginal costs. Therefore, increasing multipliers more than necessary would also increase the inefficiency in gas dispatch and cause welfare losses as our analysis has shown. Furthermore, higher multipliers are shown to increase volatility of prices and regional price spreads. Hence, unnecessarily high multipliers may be detrimental to the integration of the EU gas market.

Despite the above-mentioned inefficiencies associated with multipliers, multipliers that are sufficiently high can nevertheless be favoured by the regulators for several reasons. We have shown that multipliers determine how gas transmission capacity is booked, in turn affecting how gas infrastructure is utilised. Overall, higher multipliers were shown to decrease the peak transport volumes and increase the volumes stored in gas storages. In this respect, it can be argued that higher multipliers may strengthen the security of supply of the system by reducing the volatility of gas import volumes and promoting storage. Furthermore, the ensuing flatter gas import profiles may also reduce the need for future capacity extensions, potentially resulting in higher long-term efficiency.

Regulators can also favour higher multipliers due to their distributional effect. Multipliers that are sufficiently high can maximise consumer surplus by allowing transport tariffs to be reduced. Setting the multipliers for the purpose of maximising consumer surplus penalises the traders and the producers while benefiting the storage operators. The producers in this case are the Russian gas production companies and the traders would be the various EU and non-EU energy and trading companies. Storage operators are predominantly EU companies with some storages owned by non-EU firms (e.g. Gazprom). Therefore, from an EU perspective, setting the multipliers to maximise consumer surplus would likely be optimal as it would largely benefit the consumers in the EU while penalising the non-EU producers.

5.5.2. Regional effects

National regulators can set the multipliers accordingly to maximise the consumer surplus. However, we have shown that the effects of multipliers vary significantly among regions. According to our analysis, the issue of choosing optimal multipliers becomes less important in regions with a relatively flat demand profile such as Iberia (Spain and Portugal), since in these regions exclusively LT capacities are booked in the model. In reality, due to decision-making under uncertainty—especially with respect to highly uncertain and volatile LNG prices—ST capacities are observed and imports from continental Europe via pipeline are less flat. The fact that overall LNG imports may be affected by multipliers may also contribute to the observation of ST bookings.

In regions with limited storage flexibility such as in the British region (United Kingdom and Ireland) and South East Europe (Romania, Bulgaria and Greece), we find that higher multipliers can cause substantial increases in the temporal price spread, benefiting base consumers while penalising peak consumers. When specifying multipliers, regulators in these regions would also have to take into account this strong distributional effect on the allocation of consumer surplus between the base and peak consumers.

In South East Europe and the British region, we have shown that wasting of booked capacities can occur with sufficiently high multipliers. This means that a portion of the booked capacities remain unused because traders cannot fully flatten their monthly import profile due to limited storage capacities. In our model, this occurs only with very high multiplier levels that lie out of the range suggested by the EU. In reality, due to decision-making under uncertainty, the capacity wasting effect of multipliers could occur even in regions with sufficient storage flexibility and with lower multipliers, being much more prevalent than what our model with perfect foresight projects. Therefore, regulators may opt for lower multipliers if it is desired to reduce the wasting of booked capacities.

Our analysis indicates significant variation in the individually optimal multiplier levels for maximising the consumer surplus in the respective regions. We have shown that these multiplier levels are influenced by three main effects. The first effect is the reduction of the overall regional price due to TSOs being able to reduce the transport tariffs. The second effect is the increase in peak prices due to higher storage costs caused by increased storage utilisation. And the third effect is the increase in peak prices when storage flexibility is limited as the prices in this case are determined by the cost of ST capacities. For the Central region considered in the model, which is an aggregation of numerous transit countries in Central and West Europe, we find that the first effect dominates. Whereas, in Italy, a downstream region with abundant storage capacities that imports gas from the transit Central region, the second effect plays an important role. In the downstream British region as well as the peripheral South East and Baltic regions with limited storage flexibility, the third effect is found to be the dominant effect. Thus, our analysis indicates that multipliers can reinforce different effects in different regions.

5.5.3. External effects and the EU optimum

National regulators can set the multipliers accordingly to maximise the consumer surplus. However, our results confirm that adjusting multiplier levels in a region does not only cause effects in that region itself but can also induce external effects in other regions. We have shown that consumer surplus gains in transit regions are directly passed on to regions that lie downstream of the transit regions (i.e. import gas from the transit region). In contrast, a direct transfer of consumer surplus gains in the downstream and peripheral regions to transit regions does not occur. Nevertheless, our results show that multiplier adjustments in the peripheral and downstream regions can still influence the transit regions in more indirect ways, such as via affecting the procurement prices in the production region or affecting the storage utilisation in the transit region itself, respectively. Consequently, setting multipliers to maximise the consumer surplus in the individual regions, i.e. setting individually optimal multipliers, does not maximise the total EU consumer surplus.

We find that individually optimal multipliers nevertheless result in a significantly higher EU consumer surplus compared to an optimal EU uniform multiplier level that applies in every region. In our analysis, the maximum EU potential consumer surplus gains via a uniform multiplier level is 82 million EUR per year while the individually optimal multipliers increase this value by 12% to 92 million EUR. In this sense, we find it appropriate that EU regulation provides an allowed range of multipliers and not absolute values. Yet, we show that this allowed range can be too restricting for some regions. While the individually optimal multipliers in the majority of regions considered in the model lie lower than the maximum allowed multipliers in the EU, the British region is found to have a much higher optimal multiplier. Hence, our results imply that the current range of allowed multipliers can be too restricting for this region, limiting the potential consumer surplus gains.

When multipliers are set in individual regions with the purpose of maximising the total EU consumer surplus, the surplus gains increase by 9% to 100 million EUR. This indicates that letting national regulators set the multiplier levels—as is the case with the current EU regulation—may not lead to an EU optimum. In the EU optimum case, we have shown that the consumers in the transit and downstream regions benefit while those in the peripheral regions are worse off compared to the individually optimal case. As such, national regulators in the peripheral regions would have little incentive to choose EU-optimal multipliers. Therefore, incentivising those regions would require some of the EU consumer surplus gains to be redistributed to peripheral regions.

The maximum consumer surplus gains in the EU of almost 100 millions EUR estimated by our model are relatively low when compared to overall EU gas market costs. The yearly EU internal gas market purchases alone are estimated to be 100 billion EUR in total (ACER, 2020b). However, contemplating those gains via multipliers with the total costs associated with the entry into the EU and entry-exit between EU market areas is more meaningful. In our model such costs amount to 4.6 billion EUR. Hence, multipliers that maximise overall consumer surplus shift approximately 2.2% of the transmission costs from the consumers to the producers and traders compared to the situation without multipliers.[98]

In our analysis, we group several market areas into individual regions and ignore the transmission costs within the regions that occur in reality. Because of that, real-world transmission costs would be higher than those in our model. Cervigni et al. (2019) estimate the total costs associated with the entry into the EU and entry-exit between EU market areas to be 5.7 billion EUR for the year of 2017. These transmission costs are 24% higher than the corresponding costs in our model, supporting the notion that the overall effects of multipliers on the consumer surplus would be higher in reality due to additional transmission costs within the regions. Another aspect which would further reinforce the effects of multipliers in reality is the presence of uncertainty. Compared to in our model with perfect foresight, traders in reality would be more inclined to book short-term capacities when there is short-term uncertainty with respect to their capacity demand. Since multipliers increase the prices of short-term capacities, the distributional effects of multipliers could be more pronounced in this case. We assume in our analysis all storages to be fully exempt from transmission tariffs. While the majority of countries in the EU either fully exempt storages from transport tariffs or apply very large discounts up to 90%, there are also countries where tariff discounts for storages are not as high. In these regions, the effects of multipliers on storage utilisation would be less pronounced and comparably more short-term products would be booked. This would allow long-term tariffs to be further decreased, increasing potential consumer surplus gains via setting multipliers optimally.

Despite the above-mentioned aspects, potential consumer surplus gains via optimal multipliers could in some cases be smaller in reality due to existing long-term bookings. In our analysis, we ignore the historical long-term capacity bookings that are already in place. In regions with particularly high proportion of historical long-term bookings, multipliers would have overall less impact due to less demand for short-term capacities. This would especially be the case where historically booked capacities exceed the demand for capacity such that traders face zero marginal costs for transmission. Nevertheless, since the historical capacity bookings will almost completely expire until 2035, it will eventually become less of a factor.

5.6. Conclusion

In the European Union's gas transmission system, the relative prices of short-term transmission capacities are specified via multipliers. Multipliers can have varying internal effects in different regions, resulting in consumer-surplus-maximising multipliers to differ between the regions. Moreover, even if individual regions

[98] Trader surplus decreases even further as traders also bear the costs from increased storage utilisation. Producer surplus also decreases further due to reduced profits from selling less gas in peak periods.

specify their own optimal multipliers, it is not obvious if it would lead to an EU optimum. This is because multiplier levels in one region can cause external effects in other regions. In order to address these issues, this paper analyses the effects of multipliers on regional prices, infrastructure utilisation, and welfare. A numerical simulation model is used to simulate the European gas dispatch and quantify the effects of multipliers in a spatial setting with six different representative regional clusters in Europe.

Overall, our results show that sufficiently high multipliers can help maximise consumer surplus by allowing transport tariffs to be reduced. Hence, optimal multiplier levels that maximise consumer surplus on a regional level or in the whole EU do exist. Nevertheless, we show that multiplier effects and consequently optimal multiplier levels depend strongly on regional characteristics. In regions with relatively flat demand profiles, i.e. with low winter-summer variation in demand, such as Portugal and Spain, only long-term capacities are booked under the model assumption of perfect foresight, irrespective of the multiplier level. In reality, under the presence of uncertainty, ST bookings are also observed. Nevertheless, our results indicate setting multipliers optimally is comparably less of an issue in such regions with flat demand profiles. In contrast, we show that in regions with limited supply flexibility via storages, such as Britain and South East Europe, higher multipliers significantly reduce the consumer surplus of peak consumers while base consumers profit. In such regions, the effects on the internal redistribution of consumer surplus between peak and base consumers should also be taken into account when specifying the multipliers.

Our analysis indicates that multiplier levels in a region can cause external effects in other regions. In transit regions, which import and re-export significant gas volumes (e.g. Central Europe) consumer surplus gains are passed on to regions that lie downstream (e.g. Italy). We show that multipliers in downstream regions can influence the transit regions indirectly due to adjusted import structure, affecting the storage utilisation in the transit region. Peripheral regions (e.g. South East Europe) can influence other regions also by affecting the temporal price spreads in the procurement prices in the production regions (e.g. Russia). Because of those external effects caused by multipliers, individually optimal multipliers do not necessarily lead to the EU optimum.

Allowing the regions to set their multipliers individually, nevertheless, results in a much more optimal outcome with 92 million EUR consumer surplus gains annually, 12% higher than what can be achieved with a uniform multiplier level applied in all regions. In this respect, it is appropriate that the current EU regulation specifies allowed multipliers in ranges and not in absolute values, as it can allow for consumer surplus gains in the EU. Nevertheless, our results indicate that letting national regulators set the multipliers may not lead to an EU optimum since the consumer surplus gains with individually optimal multipliers is found to be 9% lower than the maximum achievable consumer surplus.

In our analysis we considered a simplified spatial structure with aggregated regions for the purpose of isolating and identifying effects. In reality, due to high number of individual transit countries interconnected with each other, multiplier levels in a transit region can have a more amplified impact on the downstream regions and the whole system due to the pancaking effect. Additionally, we assumed perfect foresight when simulating the gas dispatch and the capacity booking, which results in the capacities in our model to be booked optimally as necessary. In reality, because of uncertainty and forecast errors, not all booked capacities are optimal and wasting of booked capacities is a common occurrence. We have shown that higher multipliers can result in capacity wasting. In this context, regulators may have to take into account these aspects as well when specifying the multipliers.

In future work, the modelling framework could be extended to include stochasticity in order to consider the influence of imperfect information and uncertainty on the capacity booking behaviour and their impact on the effects of multipliers. Significant changes in the gas demand structure are expected to occur in the next decades. As the share of intermittent renewables in electricity generation increases as part of the energy transition to meet the climate targets, volatile residual load will be increasingly met by flexible gas-fired generation. This will correspond to increased demand for short-term transmission capacity, especially for daily and intra-daily capacities. Therefore, it would also be worthwhile to extend the analysis by including a more granular temporal resolution and modelling daily and intra-daily capacity bookings.

A. Supplementary Material for Chapter 2

A.1. Properties of curtailment due to limited transmission capacity

Curtailment (K) can arise due to limited transmission capacity (K_i^t) and due to global potential supply exceeding global demand excluding the curtailment due to limited transmission capacity (K^d). In this section the properties of curtailment due to limited transmission capacity (K_i^t) are assessed.

A.1.1. Functional form of curtailment

Curtailment from limited transmission capacity (K_i^t) arises when for at least one period r the potential supply exceeds the nodal demand plus transmission capacity (i.e., $V_i \cdot avail_{i,r} > d_i + t$). Otherwise, there is no curtailment arising from limited transmission capacity. As the beta distribution describes the deterministic distribution of the availabilities, K_i^t can be expressed as follows:

$$K_i^t = \int_0^1 \max\left\{0, \underbrace{(x - \frac{(d_i+t)}{V_i})V_i}_{\text{level of curtailment}} \cdot \underbrace{\frac{1}{B(\alpha_i,\beta_i)} x^{\alpha_i-1}(1-x)^{\beta_i-1}}_{\text{availability density}}\right\} dx \tag{A.1}$$

For all $x < \frac{d_i+t}{V_i}$ the second element of the $max\{\}$ function is negative, such that the $max\{\}$ function returns 0. Hence, the curtailment from limited transmission capacity can be rewritten as follows:

$$K_i^t = \begin{cases} 0 & \text{if } V_i \leq d_i + t \\ \int_{\frac{d_i+t}{V_i}}^1 (x - \frac{(d_i+t)}{V_i})V_i \frac{1}{B(\alpha_i,\beta_i)} x^{\alpha_i-1}(1-x)^{\beta_i-1} dx & \text{if } V_i > d_i + t \end{cases} \tag{A.2}$$

A.1.2. Marginal curtailment given all capacity is allocated to node h

In this section I derive the marginal curtailment at node h when marginally increasing the VRE penetration v given that all capacity is allocated to node h

(i.e., $K_h^t|_{V_h=v}$). To do so I substitute V_h with v in a first step:

$$K_h^t\Big|_{V_h=v} = \begin{cases} 0 & \text{if } v \leq d_h{+}t \\ \int_{\frac{d_h+t}{v}}^{1} (x - \dfrac{(d_h{+}t)}{v}) v \frac{1}{B(\alpha_h,\beta_h)} x^{\alpha_h-1}(1{-}x)^{\beta_h-1}dx & \text{if } v > d_h{+}t \end{cases} \tag{A.3}$$

In a second step I derive $\frac{\partial K_h^t}{\partial v}|_{V_h=v}$: For $v \leq d_h + t$ the derivative of $\frac{\partial K_h^t}{\partial v}|_{V_h=v}$ is zero.

To calculate the derivative of $\frac{\partial K_h^t}{\partial v}|_{V_h=v}$ for $v > d_h + t$ I first rewrite K_h^t, such that the finite integral is solved:

$$K_h^t\Big|_{V_h=v} = \frac{\alpha_h}{\alpha_h{+}\beta_h} v + \frac{(d_h{+}t) B_{\frac{d_h+t}{v}}(\alpha_h, \beta_h) - v B_{\frac{d_h+t}{v}}(1{+}\alpha_h, \beta_h)}{B(\alpha_h, \beta_h)} - (d_h{+}t) \tag{A.4}$$

This expression is then used to calculate the derivative of $\frac{\partial K_h^t}{\partial v}|_{V_h=v}$ for $v > d_h + t$:

$$\frac{\partial K_h^t}{\partial v} = \frac{\alpha_h}{\alpha_h + \beta_h} - \frac{B_{\frac{d_h+t}{v}}(1 + \alpha_h, \beta_h)}{B(\alpha_h, \beta_h)} \tag{A.5}$$

Combining the results for the two sections, $\frac{\partial K_h^t}{\partial v}|_{V_h=v}$ can be expressed by:

$$\frac{\partial K_h^t}{\partial v}\Big|_{V_h=v} = \begin{cases} 0 & \text{for: } v \leq d_h + t \\ \frac{\alpha_h}{\alpha_h+\beta_h} - \frac{B_{\frac{d_h+t}{v}}(1+\alpha_h,\beta_h)}{B(\alpha_h,\beta_h)} & \text{for: } v > d_h + t \end{cases} \tag{A.6}$$

For $v > d_h + t$ and valid parameter values (i.e., $\alpha_h, \beta_h, d_h + t > 0$) the marginal curtailment $\frac{\partial K_h^t}{\partial v}\Big|_{V_h=v}$ is strictly positive. This is because $\frac{\alpha_h}{\alpha_h+\beta_h} > \frac{B_{\frac{d_h+t}{v}}(1+\alpha_h,\beta_h)}{B(\alpha_h,\beta_h)}$ for the valid parameter values. Further, the marginal curtailment ($\frac{\partial K_h^t}{\partial v}\Big|_{V_h=v}$) strictly increases in v. This is because $B_{\frac{d_h+t}{v}}(1 + \alpha_h, \beta_h)$ strictly decreases in the VRE penetration (v) for valid parameter configurations, while all other elements of Equation A.6 are independent of v.

A.1.3. Effect of various parameters on marginal curtailment

In this section I access the effect of the transmission capacity (t), the demand distribution (d_h), and the average availability on marginal curtailment evaluated at $V_h = v$. To do so I take the derivative of $\frac{\partial K_h^t}{\partial v}|_{V_h=v}$ with respect to respective parameter.

Effect of transmission capacity

The derivative of $\frac{\partial K_h^t}{\partial v}\Big|_{V_h=v}$ with respect to the transmission capacity (t) is given by:

$$\frac{\partial^2 K_h^t}{\partial v\, \partial t}\bigg|_{V_h=v} = \begin{cases} 0 & \text{for: } v \leq d_h + t \\ -\dfrac{(\frac{d_h+t}{v})^{\alpha_h}(1-\frac{d_h+t}{v})^{-1+\beta_h}}{vB(\alpha_h, \beta_h)} & \text{for: } v > d_h + t \end{cases} \tag{A.7}$$

For $v > d_h + t$ and valid parameter values (i.e., $\alpha_h, \beta_h, d_h + t > 0$) the term $\frac{\partial^2 K_h^t}{\partial v\, \partial t}\Big|_{V_h=v}$ is strictly negative. This implies that marginal curtailment when marginally increasing v evaluated at $V_h = v$ increases at a lower rate with increasing t.

Effect of demand distribution

The global demand d_{h+l} is split into the nodal demands d_h and d_l. However, only d_h affects the marginal curtailment evaluated at $V_h = v$ (see Equation A.6).

Hence, I analyse how increasing d_h affects the marginal curtailment by taking the derivative of $\frac{\partial K_h^t}{\partial v}\Big|_{V_h=v}$ with respect to the demand at node h:

$$\frac{\partial^2 K_h^t}{\partial v\, \partial d_h}\bigg|_{V_h=v} = \begin{cases} 0 & \text{for: } v \leq d_h + t \\ -\dfrac{(\frac{d_h+t}{v})^{\alpha_h}(1-\frac{d_h+t}{v})^{-1+\beta_h}}{vB(\alpha_h, \beta_h)} & \text{for: } v > d_h + t \end{cases} \tag{A.8}$$

For $v > d_h + t$ and valid parameter values (i.e., $\alpha_h, \beta_h, d_h + t > 0$) the term $\frac{\partial^2 K_h^t}{\partial v\, \partial d_h}\Big|_{V_h=v}$ is strictly negative. This implies that marginal curtailment when marginally increasing v evaluated at $V_h = v$ increases at a lower rate with increasing d_h.

Effect of average availability

The availability is described by the parameters α_h and β_h. As stated in Chapter 2.2, increasing α_h primarily increases the average availability, while the variance remains rather constant. To assess the effect of changes in the average availability, I analyse the effects arising from changes in α_h. However, only α_h affects the marginal curtailment evaluated at $V_h = v$ (see Equation A.6).

Hence, I analyse how increasing α_h affects the marginal curtailment by taking the derivative of $\frac{\partial K_h^t}{\partial v}\Big|_{V_h=v}$ with respect to α_h:

$$\frac{\partial^2 K_h^t}{\partial v \partial \alpha_h}\Bigg|_{V_h=v} = \begin{cases} 0 & \text{for: } v \leq d_h + t \\ \frac{\beta_h}{(\alpha_h+\beta_h)^2} \\ -\frac{B_{\frac{d_h+t}{v}}(\alpha_h+1,\beta_h)\left(\log\left(\frac{d_h+t}{v}\right)+\psi^{(0)}(\alpha_h+\beta_h)-\psi^{(0)}(\alpha_h)\right)}{B(\alpha_h,\beta_h)} & \text{for: } v > d_h + t \\ +\frac{\left(\frac{d+t}{v}\right)^{\alpha_h} {}_3F_2\left(\alpha_h+1,\alpha_h+1,1-\beta_h;\alpha_h+2,\alpha_h+2;\frac{d+t}{v}\right)}{\frac{v(1+\alpha_h)^2}{d+t}B(\alpha_h,\beta_h)} \end{cases} \tag{A.9}$$

While the effect of t and d_h on marginal curtailment induce the same change in marginal usable supply, this is not the case for changes in α_h. This is because changes in α_h also affect the level of marginal potential supply ($US_h = PS_h - K_h$). Hence, to access the effect of α_h on the spatial allocation of marginal capacity I also derive the effect on marginal usable supply.

In a first step I derive the effect of α_h on the marginal potential supply:

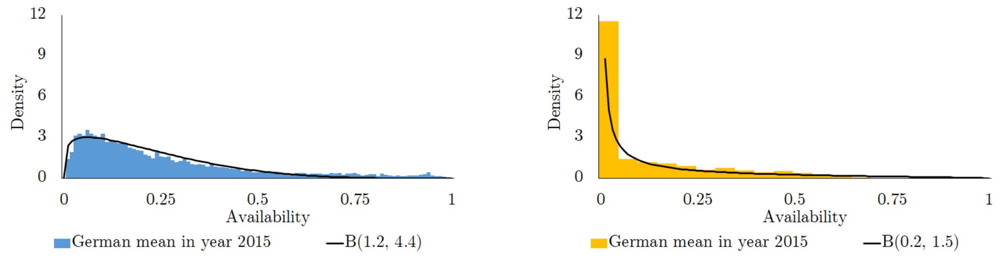
$$\begin{aligned} PS_h\Bigg|_{V_h=v} &= v \cdot \frac{\alpha_h}{\alpha_h+\beta_h} \\ \frac{\partial PS_h}{\partial v}\Bigg|_{V_h=v} &= \frac{\alpha_h}{\alpha_h+\beta_h} \\ \frac{\partial^2 PS_h}{\partial v \partial \alpha_h}\Bigg|_{V_h=v} &= \frac{\beta_h}{(\alpha_h+\beta_h)^2} \end{aligned} \tag{A.10}$$

In a second step I derive $\frac{\partial^2 US_h}{\partial v \partial \alpha_h}\Big|_{V_h=v}$ by subtracting $\frac{\partial^2 K_h^t}{\partial v \partial \alpha_h}\Big|_{V_h=v}$ from $\frac{\partial^2 PS_h}{\partial v \partial \alpha_h}\Big|_{V_h=v}$:

$$\frac{\partial^2 US_h}{\partial v \partial \alpha_h}\Bigg|_{V_h=v} = \begin{cases} \frac{\beta_h}{(\alpha_h+\beta_h)^2} & \text{for: } v \leq d_h + t \\ \frac{B_{\frac{d_h+t}{v}}(\alpha_h+1,\beta_h)\left(\log\left(\frac{d_h+t}{v}\right)+\psi^{(0)}(\alpha_h+\beta_h)-\psi^{(0)}(\alpha_h)\right)}{B(\alpha_h,\beta_h)} & \text{for: } v > d_h + t \\ -\frac{\left(\frac{d+t}{v}\right)^{\alpha_h} {}_3F_2\left(\alpha_h+1,\alpha_h+1,1-\beta_h;\alpha_h+2,\alpha_h+2;\frac{d+t}{v}\right)}{\frac{v(1+\alpha_h)^2}{d+t}B(\alpha_h,\beta_h)} \end{cases} \tag{A.11}$$

For $v > d_h + t$ and valid parameter values (i.e., $\alpha_h, \beta_h, d_h + t > 0$) the term $\frac{\partial^2 K_h^t}{\partial v\, \partial \alpha_h}\Big|_{V_h=v}$ can be positive or negative. This implies that marginal usable supply when marginally increasing v evaluated at $V_h = v$ can either increase at a lower rate or higher rate with increasing α_h.

A.2. Historical availabilities for wind and solar in Germany and corresponding Beta distribution

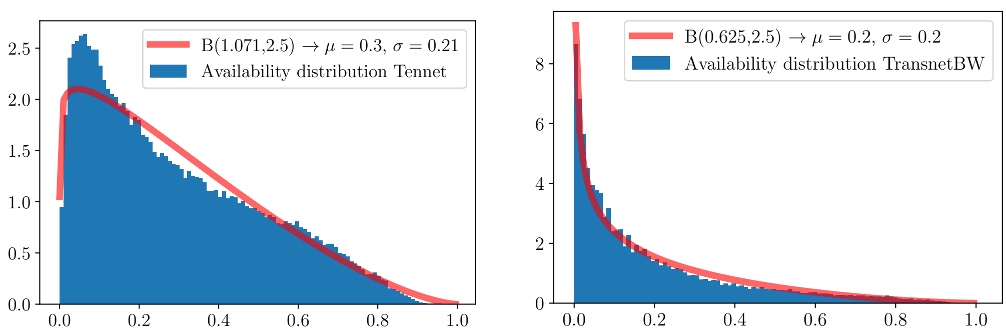


(i) Wind Onshore (ii) Solar

Figure A.1.: Comparison of historical availabilities for wind and solar in Germany with the corresponding Beta distribution.

A.3. Applied and historical densities for wind power

Figure A.2 shows the density of potential supply assumed for the high and low-availability node in all figures of Section 3 and 4 with constant availability distribution parameters (i.e., Figure 2.2, 2.3, 2.4, 2.5, 2.8, 2.9, 2.10). Additionally, the Figure A.2 displays the density of estimated historical wind power availabilities for the year 2015-2022 in the German market areas TransnetBW and Tennet. The estimation is based on data from the Bundesnetzagentur's electricity market information platform (BNetzA, 2022). TransnetBW is located in the south of Germany, and most wind power plants in the Tennet market area are located in the north. This implies that the availability density parameters in the analysis at hand resemble the availabilities for wind in the north (h) and south (l) of Germany.



(i) High-availability node (ii) Low-availability node

Figure A.2.: Historical availability densities for the years 2015-2022 and in this analysis applied densities.

A.4. Description and explanation of formulas used to compute Figures 2.2-2.12

The Figures 2.2-2.12 are the result of numerical optimisations. In each figure, the optimisation is obtained for 201 different VRE penetration levels (v) ranging from 0 to $6d_{h+l}$, such that the interval between two analysed VRE penetration levels equals $\frac{6d_{h+l}}{200}$.

The optimisation problem solved numerically resembles the optimisation problems I describe in Chapter 2.3 and 2.4. For the case of nodal pricing (Chapter 2.3) I use Equation 2.8 as objective function. The objective function contains the terms PS_i and K, which both depend on the decision variable V_h.

To calculate the potential VRE supply (PS_i) I apply Equation 2.4. However, instead of using the beta density function as availability profile, implicitly assuming availability realisations going towards infinity ($n \to \infty$), I use a finite number of availability realisations ($avail_{i,r}$). Thereby r represents the index for different realisations. Considering 10 million periods balances the trade-off between being as close as possible to the underlying beta distribution and allowing for computational feasibility (i.e., $n = 1 \cdot 10^7$). For $n = 1 \cdot 10^7$ the 99%-confidence interval of μ_i is given by $\mu_i \pm 0.0026\sigma_i$.

Like described in Equation 2.7 curtailment (K) can be separated into K_i^t and K^d. K_i^t is calculated by applying Equation A.1. However, instead of using the exact beta density function (implicitly assuming an $n \to \infty$) I consider 10 million periods which leads to the following function:

$$K_i^t = \frac{1}{n} \sum_r \max \left\{ 0, V_i avail_{i,r} - (d_i + t) \right\} \tag{A.12}$$

K^d occurs, when the global potential supply minus curtailment from limited transmission capacity exceeds the global demand. The level of K^d is calculated using the following function:

$$K^d = \frac{1}{n} \sum_r max \left\{ 0, \sum_i \underbrace{V_i avail_{i,r}}_{PS_{i,r}} - \underbrace{\sum_i \max \left\{ 0, V_i avail_{i,r} - (d_i + t) \right\}}_{K_{h,r}^t + K_{l,r}^t} - d_{h+l} \right\} \tag{A.13}$$

Thereby I consider each of the 10 million periods (i.e., $n = 1 \cdot 10^7$).

For the case of uniform pricing, I use Equation 2.19 as objective function (Chapter 2.4). The objective function contains the terms potential supply (PS_i) and commercial curtailment ($\overline{K^c}$). The potential supply is calculated analogue as presented above. The commercial commercial ($\overline{K^c}$) I calculate using the following

Equation:

$$\overline{K^c} = \frac{1}{n}\sum_{r}^{n} \max\{0, \underbrace{\sum_i V_i avail_{i,r}}_{PS_{h,r}+PS_{l,r}} - d_{h+l}\} \tag{A.14}$$

Like in the case of nodal pricing I consider 10 million periods (i.e., $n = 1 \cdot 10^7$).

Next to solving the problem numerically, I derive several results. First, I calculate the marginal capacity share at node h, which represents how nodal capacity changes with an increase of v. This is done by calculating $\frac{\Delta V_h}{\Delta v}$ with $\Delta v = \frac{6d_{h+l}}{200}$ using the optima calculated for the 201 VRE penetration levels. Second, I calculate the average capacity share at node h (i.e., $\frac{V_h}{v}$). Third, I show the marginal usable supply and node h and l (i.e., $\frac{\Delta US_h}{\Delta v}$ and $\frac{\Delta US_l}{\Delta v}$). The numbers are calculated by using the following Equation:

$$\frac{\Delta US_i}{\Delta v} = \frac{US|_{v=\overline{v}+\epsilon, V_i=(V_i^*|_{v=\overline{v}})+\epsilon} - US|_{v=\overline{v}, V_i=V_i^*}}{(\overline{v}+\epsilon)-\overline{v}} \tag{A.15}$$

The equation contains ϵ to identify the effect on an incremental increase in the VRE penetration (v). In the numerical simulation I assume $\epsilon = 0.01$. Fourth, the global VRE share is provided, defined by $\frac{\sum_i PS_i - K_i}{d_{h+l}}$. The reduction in global VRE share arising from the inefficient allocation under uniform (while keeping all parameters constant) is provided as well in the figures in Chapter 2.4. Figure 2.8 additionally shows the marginal saleable supply at node h and l under uniform pricing (i.e., $\frac{\Delta SS_h}{\Delta v}$ and $\frac{\Delta SS_l}{\Delta v}$). The numbers are calculated by using the following Equation:

$$\frac{\Delta SS_i}{\Delta v} = \frac{SS|_{v=\overline{v}+\epsilon, V_i=(V_i^*|_{v=\overline{v}})+\epsilon} - SS|_{v=\overline{v}, V_i=V_i^*}}{(\overline{v}+\epsilon)-\overline{v}} \tag{A.16}$$

Like in Equation A.15 I assume $\epsilon = 0.01$.

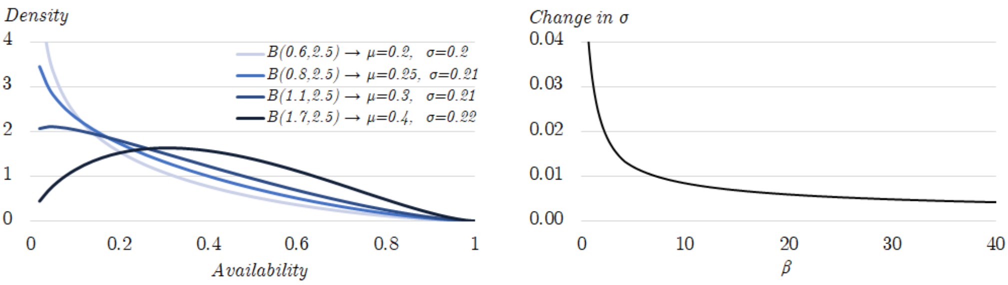
To do the calculations and to generate the figures I applied Python 3.9 using the Packages Scipy 1.9.3 (generating the beta distribution and optimising the problem), Numpy 1.24.0 and Pandas 1.5.2 (data preparation) as well as Matplotlib 3.6 (plotting the figures).

A.5. Effect of changing μ_i with the means of α_i on σ_i

The Beta distribution $B(\alpha, \beta)$is defined by the parameter α and β. These parameters define the average and the standard deviation. When changing the average with the means of changing α. The standard deviation remains rather constant.

Figure A.3a) shows the effect of varying μ_i in the interval [0.2,0.5] with the means of changing α_i for the case of $\beta = 4.4$. The density function and the resulting standard deviation are displayed. One can see that the density functions moves, while standard deviation remains rather constant, varying only between

0.2 and 0.22.
Figure A.3b) displays the maximum change in the standard deviation when varying α_i, such that μ_i is varied in the interval [0.2,0.5] for different level of β. One can see, that the maximum change in the variance for $\beta \in [0.5, 40]$ does not exceed 0.04. For $\beta > 2.2$ the maximum change does not exceed 0.02.

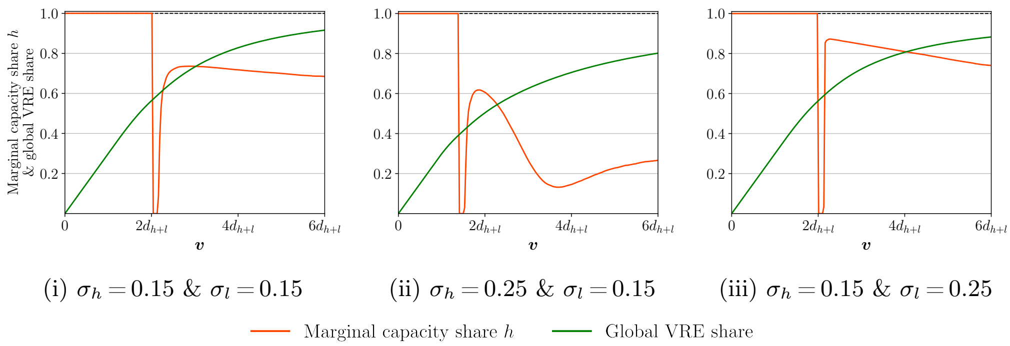


(i) Effect on density (ii) Effect on σ when varying $\mu \in [0.2, 0.5]$

Figure A.3.: Effect on σ when changing μ with the means of α.

A.6. Effects of the variance on the spatial allocation ranges when transmission capacity is high

Figure A.4 displays the insights from Finding NP 6 numerically for the case of high transmission capacity (i.e., $t = \frac{3}{4}d_i$). Assumptions regarding the demand and the availability profiles are identical to Figure 2.6.
When the variance is increased at node h (compare Figure A.4a and b), the *high-potential deployment range* is shortened from $2.0d_{h+l}$ to $1.4d_{h+l}$. Additionally, the figure confirms that increasing σ_i^2 lowers the nodal capacity share in the *split capacity range* independent of the VRE penetration level.



(i) $\sigma_h = 0.15$ & $\sigma_l = 0.15$ (ii) $\sigma_h = 0.25$ & $\sigma_l = 0.15$ (iii) $\sigma_h = 0.15$ & $\sigma_l = 0.25$

Figure A.4.: Effect of the variance in the availability profile on the spatial allocation ranges under nodal pricing.

B. Supplementary Material for Chapter 3

B.1. Proof of Proposition 1

For the proof of Proposition 3.2.1, we compare the socially optimal outcome to the three carbon pricing regimes. In the following, we derive the outcomes of these regimes.

Regulatory flexibility

In a setting with *Regulatory flexibility*, the regulator sets the carbon price after the firms have invested in the emission-free technology. The regulator faces the optimisation problem:

$$\max_{p} \mathcal{W} = \int_{p}^{\infty} Q(z)dz + (p-d)Q(p) + \int_{0}^{\overline{\chi}} (d - c_v - c_i z)dz \tag{B.1}$$

We derive the optimal solution by deriving the first-order conditions:

$$\frac{\partial \mathcal{W}}{\partial p} = -Q(p) + Q(p) + Q'(p)(p-d) = 0 \longrightarrow p^{Flex} = d \tag{B.2}$$

As in the social optimum, the carbon price equals the damage of one additional unit of the good. In t_3, the investments are already set, and, hence, the social planner and the regulator face identical problems. The carbon price does not influence the emission-free production capacity but only determines the optimal level of consumption and, in consequence, pollution.

In t_2, the firms choose to invest in the emission-free technology, as long as the associated profits are positive. Firms anticipate the carbon price that arises in the subsequent stage. The profit of the marginal firm investing in the emission-free technology is zero and, hence, the emission-free production capacity is defined by

$$\begin{aligned} \pi(\overline{\chi}) &= p^{Flex} - c_v - c_i\overline{\chi} = 0 \\ \longrightarrow \overline{\chi}^{Flex} &= \frac{p^{Flex} - c_v}{c_i} \end{aligned} \tag{B.3}$$

The optimal emission-free production capacity is at the socially optimal level, as the carbon price set in t_3 equals the marginal damage ($p^{Flex} = d$), i.e. $\overline{\chi}^{Flex} = {}^{d-c_v}/_{c_i}$.

Commitment

When the regulator commits to a carbon price, she faces no decision in t_3. In t_2, the firms choose to invest in the emission-free technology if the associated profits are positive, such that the marginal firm investing is defined by:

$$\begin{aligned} \pi(\overline{\chi}) &= p - c_v - c_i\overline{\chi} = 0 \\ \longrightarrow \overline{\chi}^{Com} &= \frac{p - c_v}{c_i} \end{aligned} \tag{B.4}$$

In t_1, the regulator chooses the carbon price that maximises the social welfare function while anticipating the reaction function of firms to the announced price.

$$\begin{aligned} \max_p \mathcal{W} &= \int_p^{\infty} Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}(p)} (d - c_v - c_i z)dz \\ \frac{\partial \mathcal{W}}{\partial p} &= -Q(p) + Q(p) + Q'(p)(p-d) + \overline{\chi}'(p)(d - c_v - c_i\overline{\chi}) = 0 \end{aligned} \tag{B.5}$$

Inserting the optimal investment level $\overline{\chi}^{Com}$ from (B.4), the expression yields:

$$Q'(p)(p-d) = \overline{\chi}'(p)(p-d) \longrightarrow \quad p^{Com} = d \tag{B.6}$$

As under *Regulatory flexibility*, the solution yields the social optimum. In the absence of risk, there is no difference for the regulator in setting the carbon price in t_1 or t_3.

CCfD

When the regulator offers a CCfD, she sets the carbon price in t_3 after the firms invested in the emission-free technology. The solution yields the same result as under *Regulatory flexibility*, as the regulator can only control the size of the market at this stage.

$$\begin{aligned} \max_p \mathcal{W} &= \int_p^{\infty} Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}} (d - c_v - c_i z)dz \\ \longrightarrow \ p^{CCfD} &= d \end{aligned} \tag{B.7}$$

In t_2, the firms choose to invest in the emission-free technology according to their profit function, which depends on the strike price of the CCfD. The carbon price is irrelevant to the firms.

$$\begin{aligned} \pi(\overline{\chi}) &= p_s - c_v - c_i\overline{\chi} = 0 \\ \longrightarrow \overline{\chi}^{CCfD} &= \frac{p_s - c_v}{c_i} \end{aligned} \tag{B.8}$$

The result is the socially optimal emission-free production capacity that balances the marginal costs and the benefit of abatement, i.e., savings from reduced payment of the strike price. In t_1, the regulator chooses the strike price that she offers to the firms. She faces the following optimisation problem:

$$\begin{aligned} \max_{p_s} \mathcal{W} &= \int_p^\infty Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}(p_s)} (d - c_v - c_i z)dz \\ \frac{\partial \mathcal{W}}{\partial p_s} &= [d - c_v - c_i \overline{\chi}(p_s)]\overline{\chi}'(p_s) = 0 \end{aligned} \tag{B.9}$$

Inserting the optimal investment level $\overline{\chi}^{CCfD}$ from (B.8), the expression yields $p_s^{CCfD} = d$. Hence, the strike price equals marginal damage, and the strike price and carbon price have the same level in the absence of risk. Firms and consumers receive the same signal regarding the benefit from investments or the damage from consumption, respectively. Both prices are at the socially optimal level.

Welfare ranking

As all three carbon pricing regimes result in the socially optimal carbon price and the socially optimal emission-free production capacity, it is straightforward that the respective welfare is equal to the social optimum.

B.2. Proof of Proposition 2

For the proof of Proposition 3.3.1, we derive the optimal solutions in the respective carbon pricing regimes and under the assumption of a social planner.

Social optimum

In the social optimum, the social planner sets the carbon price p in t_3 after the actual environmental damage revealed. She optimises:

$$\max_{p} \mathcal{W} = \int_p^\infty Q(z)dz + (p-\hat{d})Q(p) + \int_0^{\overline{\chi}} (\hat{d} - c_v - c_i z)dz \tag{B.10}$$

Given the first-order conditions, the optimal solution is equal to:

$$\frac{\partial \mathcal{W}}{\partial p} = -Q(p) + Q(p) + Q'(p)(p-\hat{d}) = 0 \longrightarrow p^{Opt} = \hat{d} \tag{B.11}$$

The investments are due before the actual damage reveals. Hence, the social planner must choose the emission-free production capacity in the presence of risk. The social planner optimises the expected welfare with respect to the emission-free production capacity $\overline{\chi}$.

$$\max_{\overline{\chi}} E[\mathcal{W}] = E\left[\int_p^\infty Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}} (d - c_v - c_i z)dz\right] \tag{B.12}$$

Given the expected damage, the optimal solution is equal to:

$$\frac{\partial E[\mathcal{W}]}{\partial \overline{\chi}} = E[d] - c_v - c_i\overline{\chi} = 0 \longrightarrow \overline{\chi}^{Opt} = \frac{E[d] - c_v}{c_i} = \frac{\mu_D - c_v}{c_i} \tag{B.13}$$

Regulatory flexibility

Under *Regulatory flexibility*, similar to the assumption of a social planner, the regulator sets the carbon price after the actual damage revealed. As shown in B.1, in this case, the regulator and the social planner have the same objective function. Hence, in *Flex*, the regulator optimises (B.10), which yields $p^{Flex} = \hat{d}$.

In t_2, the firms choose to invest in the emission-free technology, as long as the associated profits are positive. They anticipate the subsequent carbon price:

$$\begin{aligned} EU(\pi(\overline{\chi})) &= E[p^{Flex}] - c_v - c_i\overline{\chi} - \lambda\sigma_{p^{Flex}} = 0 \\ \longrightarrow \overline{\chi}^{Flex} &= \frac{p^{Flex} - c_v - \lambda\sigma_{p^{Flex}}}{c_i} = \frac{\mu_D - c_v - \lambda\sigma_D}{c_i} \end{aligned} \tag{B.14}$$

where the last step stems from replacing the statistical moments of the carbon price in *Flex* with the ones of the environmental damage, i.e., $E[p^{Flex}] = \mu_D$ and $\sigma_{p^{Flex}} = \sigma_D$. The emission-free production capacity decreases with the volatility of the environmental damage and firms' risk aversion, as $\frac{\partial \overline{\chi}^{Flex}}{\partial \lambda} = -\frac{\sigma_D}{c_i}$ and $\frac{\partial \overline{\chi}^{Flex}}{\partial \sigma_D} = -\frac{\lambda}{c_i}$ are both smaller than zero.

Commitment

When the regulator commits to a carbon price, she faces no decision in t_3. In t_2, the firms make their investment decision given the announced carbon price level. In this setting, all parameters are known, such that firms face no risk:

$$\begin{aligned} \pi(\overline{\chi}) &= p - c_v - c_i\overline{\chi} = 0 \\ \longrightarrow \overline{\chi}^{Com} &= \frac{p - c_v}{c_i} \end{aligned} \tag{B.15}$$

In t_1, the regulator sets the carbon price maximising expected welfare and accounting for the firms' reaction function to the announced price:

$$\begin{aligned} \max_p E[\mathcal{W}] &= E\left[\int_p^\infty Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}(p)} (d - c_v - c_i z)dz\right] \\ \frac{\partial E[\mathcal{W}]}{\partial p} &= -Q(p) + Q(p) + Q'(p)(p-\mu_D) + \overline{\chi}'(p)(\mu_D - c_v - c_i\overline{\chi}) = 0 \end{aligned} \tag{B.16}$$

Inserting the resulting emission-free production capacity $\overline{\chi}^{Com}$ from (B.15), the expression yields:

$$Q'(p)(p-\mu_D) = \overline{\chi}'(p)(p-\mu_D) \longrightarrow \quad p^{Com} = \mu_D \tag{B.17}$$

CCfD

When the regulator offers a CCfD, she sets the carbon price in t_3 after the firms made their investment decision. Hence, she optimises (B.10), and the solution is identical with the one of the social planner and under *Regulatory flexibility*, i.e., $p^{CCfD} = \hat{d}$.

In t_2, the firms choose to invest accounting for the strike price of the CCfD. The carbon price is irrelevant to firms. Hence, the maximisation problem is identical to (B.8), and the solution is equal to:

$$\overline{\chi}^{CCfD} = \frac{p_s - c_v}{c_i} \tag{B.18}$$

In t_1, the regulator chooses the strike price that maximises the expected social welfare:

$$\begin{aligned} \max_{p_s} E[\mathcal{W}] &= E\left[\int_p^\infty Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}(p_s)} (d - c_v - c_i z)dz\right] \\ \frac{\partial E[\mathcal{W}]}{\partial p_s} &= [\mu_D - c_v - c_i\overline{\chi}(p_s)] = 0 \end{aligned} \tag{B.19}$$

Inserting the optimal investment level $\overline{\chi}^{CCfD}$ from (B.18), the first-order condition yields $p_s^{CCfD} = \mu_D$. Hence, the strike price equals the expected marginal damage. Inserting p_s^{CCfD} into (B.18) shows that the investment level is socially optimal and equals the solution under *Commitment*.

Welfare ranking

As shown before, the carbon price and the emission-free production capacity are identical in the social optimum and in the *CCfD* regime. Thus, welfare in the *CCfD* regime and in the social optimum is identical, i.e., $E[\mathcal{W}_{\sigma_D}^{Opt}] = E[\mathcal{W}_{\sigma_D}^{CCfD}]$.

The emission-free production capacity under *Regulatory flexibility* is lower than the under the *CCfD* regime, as:

$$\overline{\chi}^{CCfD} - \overline{\chi}^{Flex} = \frac{\mu_D - c_v}{c_i} - \frac{\mu_D - c_v - \lambda\sigma_D}{c_i} = \frac{\lambda\sigma_D}{c_i} \geq 0 \tag{B.20}$$

Expected welfare increases with the number of firms investing in the emission-free technology, as long as $\overline{\chi} \leq \overline{\chi}^{CCfD} = \frac{\mu_D - c_v}{c_i}$, since $\frac{\partial E[\mathcal{W}]}{\partial \overline{\chi}} = \mu_D - c_v - c_i\overline{\chi}$ which is a positive number for all $\overline{\chi} < \frac{\mu_D - c_v}{c_i}$. Hence, welfare under regulatory flexibility is lower than socially optimal, i.e., $E[\mathcal{W}_{\sigma_D}^{CCfD}] \geq E[\mathcal{W}_{\sigma_D}^{Flex}]$.

The difference in welfare between the policy regimes of *Commitment* and *CCfD* stems from the difference in consumer surplus, as the respective emission-free production capacity are identical. Since the consumer surplus is a convex function, the welfare difference is positive:[99]

$$E[\mathcal{W}_{\sigma_D}^{CCfD}] - E[\mathcal{W}_{\sigma_D}^{Com}] = E[\int_D^\infty Q(z)dz] - \int_{\mu_D}^\infty Q(z)dz \geq 0 \tag{B.21}$$

Hence, it holds that $E[\mathcal{W}_{\sigma_D}^{CCfD}] \geq E[\mathcal{W}_{\sigma_D}^{Com}]$.

Whether the difference in expected welfare between *Flex* and *Com* is positive or not, is ambiguous. The difference is equal to

$$\begin{aligned} E[\mathcal{W}_{\sigma_D}^{Flex}] - E[\mathcal{W}_{\sigma_D}^{Com}] = &\underbrace{E[\int_D^\infty Q(z)dz] - \int_{\mu_D}^\infty Q(z)dz}_{\geq 0} \\ &\underbrace{+(\mu_D - c_v)(\overline{\chi}^{Flex} - \overline{\chi}^{Com}) - \int_{\overline{\chi}^{Com}}^{\overline{\chi}^{Flex}} (c_i z)dz}_{\leq 0}, \end{aligned} \tag{B.22}$$

where the first part, i.e., difference in consumer surplus, is positive and the second part, i.e., the difference in abatement benefit, is negative.

B.3. Proof of Proposition 3

For the proof of Proposition 3.3.2, we derive the optimal solutions in the respective carbon pricing regimes and under the assumption of a social planner.

[99]This relation is also known, as Jensen gap stemming from Jensen's inequality.

Social optimum

In the social optimum, the social planner sets in t_3 the carbon price p after the actual level of variable costs revealed. She optimises:

$$\begin{aligned} \max_p \mathcal{W} &= \int_p^{\infty} Q(z)dz + (p-d)Q(p) \int_0^{\overline{\chi}} (d - \hat{c_v} - c_i z)dz \\ \frac{\partial \mathcal{W}}{\partial p} &= -Q(p) + Q(p) + Q'(p)(p-d) = 0 \end{aligned} \tag{B.23}$$

Given the first-order condition, the optimal solution is equal to $p^{Opt} = d$.

The investments are due before the level of variable costs reveals. Hence, the social planner must set the emission-free production capacity in the presence of risk. The social planner optimises the expected welfare with respect to the emission-free production capacity $\overline{\chi}$, as depicted in (B.12). Given the expected variable costs, the optimal solution is equal to:

$$\frac{\partial E[\mathcal{W}]}{\partial \overline{\chi}} = d - E[c_v]c_i - \overline{\chi} = 0 \longrightarrow \overline{\chi}^{Opt} = \frac{d - \mu_{C_v}}{c_i} \tag{B.24}$$

Regulatory flexibility

As under the assumption of a social planner, the regulator sets the carbon price in t_3. Again, the regulator and the social planner have the same objective function. Hence, under *Regulatory flexibility*, the regulator optimises (B.23), which yields $p^{Flex} = d$.

In t_2, the firms take their investment decision, anticipating the risk in variable costs that arises in the subsequent stage:

$$\begin{aligned} EU(\pi(\overline{\chi})) &= p^{Flex} - E[c_v] - c_i\overline{\chi} - \lambda\sigma_{C_v} = 0 \\ \longrightarrow \overline{\chi}^{Flex} &= \frac{p^{Flex} - \mu_{C_v} - \lambda\sigma_{C_v}}{c_i} = \frac{d - \mu_{C_v} - \lambda\sigma_{C_v}}{c_i} \end{aligned} \tag{B.25}$$

where the last step stems from replacing the optimal carbon price in *Flex*. The emission-free production capacity decreases with the volatility of the variable costs and the firms' risk aversion, as $\frac{\partial \overline{\chi}^{Flex}}{\partial \lambda} = -\frac{\sigma_{C_v}}{c_i}$ and $\frac{\partial \overline{\chi}^{Flex}}{\partial \sigma_{C_v}} = -\frac{\lambda}{c_i}$, which both are smaller than zero.

Commitment

When the regulator commits to a carbon price, she faces no decision in t_3. In t_2, the firms choose to invest in the emission-free technology given the announced carbon price level. In this setting, the firms still face a risk, stemming from the variable costs. The firms invest if their expected utility is greater than zero.

Hence, the marginal firm investing in the emission-free technology is characterised by:

$$\begin{aligned} EU(\pi(\overline{\chi})) &= p^{Com} - E[c_v] - c_i\overline{\chi} - \lambda\sigma_{C_v} = 0 \\ \longrightarrow \overline{\chi}^{Com} &= \frac{p^{Com} - \mu_{C_v} - \lambda\sigma_{C_v}}{c_i} \end{aligned} \tag{B.26}$$

In t_1, the regulator sets the carbon price maximising expected welfare and accounting for the reaction function of the firms to the announced price:

$$\begin{aligned} \max_p E[\mathcal{W}] &= E\left[\int_p^{\infty} Q(z)dz + (p-d)Q(p) \int_0^{\overline{\chi}} (d - \hat{c_v} - c_i z)dz\right] \\ \frac{\partial E[\mathcal{W}]}{\partial p} &= Q'(p)(p-d) + \overline{\chi}'(p)(d - \mu_{C_v} - c_i\overline{\chi}(p)) = 0 \\ \longrightarrow p - d &= \frac{\overline{\chi}'(p)}{-Q'(p)}(d - \mu_{C_v} - c_i\overline{\chi}(p)) \end{aligned} \tag{B.27}$$

Rearranging the first-order condition and substituting $\epsilon(p) = -\frac{\partial Q(p)}{\partial p}\frac{p}{Q(p)}$ yields the expression in (3.14). Additionally, we define $\eta = \frac{\overline{\chi}'(p)}{-Q'(p)}$. Substituting η in (B.27) and using $\overline{\chi}(p)^{Com}$ from (B.26), yields

$$p^{Com} = d + \frac{\eta}{1+\eta}\lambda\sigma_{C_v} \tag{B.28}$$

The resulting carbon price is greater than the environmental damage d, as η is a positive number.

CCfD

When the regulator offers a CCfD, she sets the carbon price in t_3 after the firms made their investment decision. Hence, she optimises (B.23), and the solution is identical with the one of the social planner and under *Regulatory flexibility*, i.e., $p^{CCfD} = d$.

In t_2, the firms invest in the emission-free technology accounting for the strike price of the CCfD. As in the other carbon pricing regimes, the firms face a risk in variable costs. The marginal firm investing in the emission-free technology is characterised by:

$$\begin{aligned} EU(\pi(\overline{\chi})) &= p_s - E[c_v] - c_i\overline{\chi} - \lambda\sigma_{C_v} = 0 \\ \longrightarrow \overline{\chi}^{CCfD} &= \frac{p_s - \mu_{C_v} - \lambda\sigma_{C_v}}{c_i} \end{aligned} \tag{B.29}$$

In t_1, the regulator chooses the strike price that maximises the expected social welfare:

$$\max_{p_s} E[\mathcal{W}] = E\left[\int_p^\infty Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}(p_s)} (d - c_v - c_i z)dz\right]$$
$$\frac{\partial E[\mathcal{W}]}{\partial p_s} = d - \mu_{C_v} - c_i\overline{\chi}(p_s) = 0 \tag{B.30}$$

Inserting the optimal investment level $\overline{\chi}^{CCfD}$ from (B.29), the first-order condition is equal to

$$\left(\frac{d-\mu_{C_v}}{c_i} - \frac{p_s - \mu_{C_v} - \lambda\sigma_{C_v}}{c_i}\right) = 0 \tag{B.31}$$

, which yields $p_s^{CCfD} = d + \lambda\sigma_{C_v}$. Inserting p_s^{CCfD} into (B.29) shows that the emission-free production capacity is equal to tne one under a social planner, i.e., $\overline{\chi}^{CCfD} = \frac{d-\mu_{C_v}}{c_i}$

Welfare ranking

As shown before, the carbon price and the emission-free production capacity are identical in the social optimum and in the *CCfD* regime. Thus, welfare in the *CCfD* regime and in the social optimum is identical, i.e., $E[\mathcal{W}_{\sigma_{C_v}}^{Opt}] = E[\mathcal{W}_{\sigma_{C_v}}^{CCfD}]$.

Similar to the case of damage risk in B.2, the emission-free production capacity under *Regulatory flexibility* is lower than the under the *CCfD* regime, as:

$$\overline{\chi}^{CCfD} - \overline{\chi}^{Flex} = \frac{\lambda\sigma_{C_v}}{c_i} \geq 0 \tag{B.32}$$

Expected welfare increases in the emission-free production capacity $\overline{\chi}$, as long as $\overline{\chi} \leq \overline{\chi}^{CCfD} = \frac{d-\mu_{C_v}}{c_i}$, since $\frac{\partial E[\mathcal{W}]}{\partial \overline{\chi}} = d - \mu_{C_v} - c_i\overline{\chi}$. Hence, welfare in *Flex* is lower than socially optimal, i.e., $E[\mathcal{W}_{\sigma_{C_v}}^{CCfD}] \geq E[\mathcal{W}_{\sigma_{C_v}}^{Flex}]$.

To show that offering a CCfD is welfare superior to *Commitment*, we first compare the strike price with optimal carbon price in *Com*. Inserting $\overline{\chi}^{Com}$ and rearranging (B.28), yields:

$$p^{Com} - p_s = d + \frac{\eta}{1+\eta}\lambda\sigma_{C_v} - (d + \lambda\sigma_{C_v}) = \left(\frac{\eta}{1+\eta} - 1\right)\lambda\sigma_{C_v} \tag{B.33}$$

As η is a positive number, the first expression is negative and the difference is negative. Hence, we see that the optimal carbon price under commitment p^{Com} is smaller than the strike price of the CCfD. Consequently, the emission-free production capacity in *Com* is lower than when offering a CCfD, i.e., $\overline{\chi}^{CCfD} \geq \overline{\chi}^{Com}$. Similarly, it is straightforward to show that the carbon price under the *Com* regime is higher than under the *CCfD* regime. Both variables lead to lower welfare and, hence, we show that $E[\mathcal{W}^{CCfD}_{\sigma_{C_v}}] \geq E[\mathcal{W}^{Com}_{\sigma_{C_v}}]$.

To show that in this setting, *Commitment* to a carbon price is welfare superior to *Regulatory flexibility*, we can make use of the optimality of the carbon price in *Com*. The regulator sets a price above the marginal environmental damage to incentivise additional investments. She could, however, choose not to. We show the optimality by comparing:

$$\begin{aligned}
&E[\mathcal{W}^{Com}_{\sigma_{C_v}}] \\
&\quad = E\left[\int_{p^{Com}}^{\infty} Q(z)dz + (p^{Com} - d)Q(p) + d\overline{\chi}^{Com} - \frac{c_i}{2}(\overline{\chi}^{Com})^2 - c_v\overline{\chi}^{Com})\overline{Q}\right] \\
&\quad \geq E\left[\int_{p^{Flex}}^{\infty} Q(z)dz + (p^{Flex} - d)Q(p) + d\overline{\chi}^{Com} - \frac{c_i}{2}(\overline{\chi}^{Com})^2 - c_v\overline{\chi}^{Com})\overline{Q}\right] \\
&\quad \geq E\left[\int_{p^{Flex}}^{\infty} Q(z)dz + (p^{Flex} - d)Q(p) + d\overline{\chi}^{Flex} - \frac{c_i}{2}(\overline{\chi}^{Flex})^2 - c_v\overline{\chi}^{Flex})\overline{Q}\right] \\
&\quad = E[\mathcal{W}^{Flex}_{\sigma_{C_v}}],
\end{aligned} \tag{B.34}$$

where the first inequality is given by the optimality of p^{Com} and the second by the fact that $\overline{\chi}^{Flex} \leq \overline{\chi}^{Com}$ (c.f. Chiappinelli and Neuhoff, 2020).

B.4. Proof of Proposition 4

For the proof of Proposition 4, we derive the optimal solutions in the respective carbon pricing regimes and under the assumption of a social planner.

Social optimum

In t_3, the social planner sets the carbon price p after the actual environmental damage revealed, by optimising (B.10). Hence, the optimal carbon price is equal to $p^{Opt} = \hat{d}$.

In t_2, the social planner sets the emission-free production capacity under risk such that it maximises the expected welfare. She considers the cases in which production may not be optimal, i.e., $c_v > \hat{d}$.

$$\begin{aligned}\max_{\overline{\chi}} E[\mathcal{W}] =&P\Big(\int_p^{\infty} Q(z)dz + (p-\hat{d})Q(p) + \int_0^{\overline{\chi}}(\hat{d} - c_v - c_i z)dz \mid c_v \leq p\Big)\\ &+P\Big(\int_p^{\infty} Q(z)dz + (p-\hat{d})Q(p) - \int_0^{\overline{\chi}}(c_i z)dz \mid c_v > p\Big)\\ &= \int_p^{\infty} Q(z)dz + (p-\mu_D)Q(p) - \int_0^{\overline{\chi}}(c_i z)dz + \int_{c_v}^{\infty} \overline{\chi}(z-c_v)f_D(z)dz\end{aligned} \tag{B.35}$$

, where $f_D(z)$ is the density function of the environmental damage. Given the first-order condition, the optimal solution is equal to:

$$\frac{\partial E[\mathcal{W}]}{\partial \overline{\chi}} = \int_{c_v}^{\infty}(z-c_v)f_D(z)dz - c_i\overline{\chi} = 0 \longrightarrow \overline{\chi}^{Opt} = \frac{\int_{c_v}^{\infty}(z-c_v)f_D(z)dz}{c_i} \tag{B.36}$$

Regulatory flexibility

As under the assumption of a social planner, the regulator sets the carbon price after the actual damage revealed with the same objective function. Hence, she sets $p^{Flex} = \hat{d}$.

In t_2, the firms invest in the emission-free technology if the associated expected utility is positive. They anticipate that the Pigouvian carbon tax depends on the damage level that is not yet revealed. The marginal firm investing in the emission-free technology is characterised by:

$$\begin{aligned}EU(\pi(\overline{\chi})) &= P\Big(p^{Flex} - c_v - c_i\overline{\chi} \mid c_v \leq p^{Flex}\Big) + P\Big(-c_i\overline{\chi} \mid c_v > p^{Flex}\Big)\\ &= \int_{c_v}^{\infty}(z-c_v)f_D(z)dz - c_i\overline{\chi} = 0\\ &\longrightarrow \overline{\chi}^{Flex} = \frac{\int_{c_v}^{\infty}(z-c_v)f_D(z)dz}{c_i}\end{aligned} \tag{B.37}$$

The emission-free production capacity equals the socially optimal level, as the carbon price set in t_3 equals the marginal damage ($p^{Flex} = \hat{d}$), i.e. $\overline{\chi}^{Flex} = \overline{\chi}^{Opt}$.

Commitment

In t_2, the firms make their investment decision given the announced carbon price level. In this setting, the firms know all parameters affecting their profits, such that the firms face no risk. However, the profit functions of firms depend on the carbon price level, and they have to distinguish two cases.

$$\pi(\chi) = \begin{cases} p - c_v - c_i\chi, & \text{for } c_v \leq p \\ -c_i\chi, & \text{else} \end{cases} \tag{B.38}$$

Given the indifference condition of the marginal firm investing in the emission-free technology:

$$\overline{\chi}^{Com} = \begin{cases} \frac{p^{Com} - c_v}{c_i}, & \text{for } c_v \leq p \\ 0, & \text{else} \end{cases} \tag{B.39}$$

In t_1, the regulator sets the carbon price anticipating that her choice impacts the firms' investment decision:

$$\max_p E[\mathcal{W}] = \begin{cases} \int_p^\infty Q(z)dz + (p - \mu_D)Q(p) & \\ \quad + \int_0^{\overline{\chi}(p)} \int_{-\infty}^\infty (t - c_v) f_D(t) - (c_i z) dt dz, & \text{if } c_v \leq p \\ \int_p^\infty Q(z)dz + (p - \mu_D)Q(p), & \text{else} \end{cases} \tag{B.40}$$

For the second case, is straightforward to show that the regulator sets carbon price equal to the expected damage. The solution for the first case is identical to the optimisation in (B.16). In both cases, the optimal carbon price equals the expected environmental damage and, thus,

$$p^{Com} = \begin{cases} \mu_D, & \text{if } c_v \leq p \\ \mu_D, & \text{else} \end{cases} \tag{B.41}$$

As by assumption the expected damage is higher than the variable costs, i.e., $\mu_D > c_v$, only the first case materialises. Thus, the optimal emission-free production capacity is equal to $\overline{\chi}^{Com} = \frac{\mu_D - c_v}{c_i}$.

CCfD

When the regulator offers a CCfD, she sets the carbon price in t_3 after the firms made their investment decision. Hence, she optimises (B.10), and the solution is identical with the one of the social planner and under *Regulatory flexibility*, i.e., $p^{CCfD} = \hat{d}$.

In t_2, the firms take their investment decision and account for the strike price of the CCfD. The carbon price is irrelevant to the firms. However, the firms only invest, if the strike price is above the variable costs.

$$\pi(\chi) = \begin{cases} p_s - c_v - c_i\chi, & \text{for } c_v \leq p_s \\ -c_i\chi, & \text{else} \end{cases} \longrightarrow \overline{\chi}^{CCfD} = \begin{cases} \frac{p_s - c_v}{c_i}, & \text{for } c_v \leq p_s \\ 0, & \text{else} \end{cases} \tag{B.42}$$

In t_1, the regulator chooses the strike price that maximises the expected social welfare:

$$\max_{p_s} E[\mathcal{W}] = \begin{cases} \int_p^\infty Q(z)dz + (p-\mu_D)Q(p) \int_0^{\overline{\chi}(p_s)} \int_{-\infty}^\infty (t-c_v)f_D(t) - & (c_i z)dtdz, \\ & \text{if } c_v \leq p_s \\ \int_p^\infty Q(z)dz + (p-\mu_D)Q(p), & \text{else} \end{cases} \tag{B.43}$$

For the second case, the strike price can take any realisation between zero and c_v, as firms would not invest. For the first case, the solution is identical to (B.30). Hence, the result is equal to

$$p_s = \begin{cases} \mu_D \\ 0 \leq p_s < c_v \end{cases} \tag{B.44}$$

Again, only the first case materialises, as by assumption $\mu_D > c_v$. Inserting p_s^{CCfD} into (B.42) shows that the investment level is equal to $\overline{\chi}^{CCfD} = \frac{\mu_D - c_v}{c_i}$.

Welfare ranking

As shown before, the carbon price and the emission-free production capacity are identical in the social optimum and under *Regulatory flexibility*. Thus, welfare in this carbon pricing regime is identical to the social optimum, i.e., $E[\mathcal{W}_{\sigma_D}^{Opt}] = E[\mathcal{W}_{\sigma_D}^{Flex}]$.

To compare *Flex* and *CCfD*, we evaluate the difference of expected welfare. Since $p^{Flex} = p^{CCfD}$, there is only a difference regarding welfare from production with the emission-free technology. Taking the derivatives of (3.19), we see that the expected social welfare is increasing in investments as long as $\overline{\chi} \leq \overline{\chi}^{Opt} = \overline{\chi}^{Flex}$:

$$\begin{aligned} \frac{\partial E[\mathcal{W}]}{\partial \overline{\chi}} &= \int_{c_v}^\infty (z-c_v)f_D(z)dz - c_i\overline{\chi} > 0 \;\forall\, \overline{\chi} < \frac{\int_{c_v}^\infty (z-c_v)f_D(z)dz}{c_i} \\ \frac{\partial^2 E[\mathcal{W}]}{\partial \overline{\chi}^2} &= -c_i < 0 \end{aligned} \tag{B.45}$$

As $\overline{\chi}^{CCfD} \leq \overline{\chi}^{Flex}$, we conclude that $E[\mathcal{W}_{\sigma_D}^{Flex}] \geq E[\mathcal{W}_{\sigma_D}^{CCfD}]$.

Lastly, it is straightforward to show that *Commitment* is welfare-inferior to the *CCfD* regime. As investments are identical in both regimes, the difference in welfare stems form the consumer surplus. Again, applying Jensen's inequality, it holds that

$$E[\mathcal{W}_{\sigma_D}^{CCfD}] - E[\mathcal{W}_{\sigma_D}^{Com}] = E[\int_D^\infty Q(z)dz] - \int_{\mu_D}^\infty Q(z)dz \geq 0. \tag{B.46}$$

B.5. Regulatory solutions with variable cost risk and potentially socially not optimal production

Under variable cost risk and potentially welfare-reducing production, the increase in marginal production costs might be so high that firms using the emission-free technology do not produce in t_4. As the investments in abatement are sunk, they do not impact the production decision. Overall welfare in t_4 is given by:

$$\mathcal{W} = \begin{cases} \int_p^\infty Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}}(d - \hat{c_v} - c_i z)dz, & \text{for } \hat{c_v} < d \\ \int_p^\infty Q(z)dz + (p-d)Q(p) - \int_0^{\overline{\chi}}(c_i z)dz, & \text{for } \hat{c_v} \geq d \end{cases} \tag{B.47}$$

Social optimum

In the social optimum, the social planner sets the carbon price p^{Opt} after the level of variable costs revealed. The optimisation is identical to maximising (B.10). Hence, it holds that $p^{Opt} = d$. The social planner sets the emission-free production capacity $\overline{\chi}^{Opt}$ such that it maximises expected welfare:

$$\begin{aligned} \max_{\overline{\chi}} E[\mathcal{W}] =& P\left(\int_p^\infty Q(z)dz + (p-d)Q(p) + \int_0^{\overline{\chi}}(d - c_v - c_i z)dz \mid c_v \leq d \right) \\ =& \int_p^\infty Q(z)dz + (p-d)Q(p) - \int_0^{\overline{\chi}}(c_i z)dz + P((d-c_v)\overline{\chi} \quad \mid c_v < d) \\ =& \int_p^\infty Q(z)dz + (p-d)Q(p) - \int_0^{\overline{\chi}}(c_i z)dz + \int_{-\infty}^d (d-z)\overline{\chi} f_{C_v}(z)dz \end{aligned} \tag{B.48}$$

We solve the problem using the first-order conditions:

$$\begin{aligned} \frac{\partial E[\mathcal{W}]}{\partial \overline{\chi}} =& - c_i \overline{\chi} + \int_{-\infty}^d (d-z) f_{C_v}(z)dt = 0 \\ \longrightarrow \overline{\chi}^{Opt} =& \frac{\int_{-\infty}^d (d-z) f(z)dt}{c_i} \end{aligned} \tag{B.49}$$

The integral of the distribution function represents the marginal benefit from abatement (damage minus variable costs) weighted by its probability of realisation. The integral is limited to d as beyond this point production does not occur and the marginal benefit, hence, is zero.

Regulatory flexibility

As under the assumption of a social planner, the regulator sets the carbon price after the firms made their investment. Hence, she optimises (B.23) and sets $p^{Flex} = \hat{d}$, which is the Pigouvian tax.

In t_2, the firms choose to invest if their expected utility is greater than zero, given the risk regarding its future variable costs and anticipating the Pigouvian carbon tax rational of the regulator. The marginal firm investing in the emission-free technology is characterised by:

$$\begin{aligned} EU(\pi(\overline{\chi})) &= P\Big(p^{Flex} - c_v - c_i\overline{\chi} \mid c_v \leq p^{Flex}\Big) + P\Big(- C(\overline{\chi}) \mid c_v > p^{Flex}\Big) = 0 \\ &= \int_{-\infty}^{d} (d - z) f_{C_v}(z) dz - c_i\overline{\chi} = 0 \\ &\longrightarrow \overline{\chi}^{Flex} = \frac{\int_{-\infty}^{d} (d - z) f_{C_v}(z) dz}{c_i} \end{aligned} \tag{B.50}$$

, where we inserted the optimal carbon price ($p^{Flex} = d$). As in the case of damage risk without risk aversion, *Regulatory flexibility* reaches the social optimum.

Commitment

Under *Commitment*, the firms choose to invest after the regulator has announced the carbon price. The rationale for investments is identical to the one of *Regulatory flexibility*, as no damage risk exists. Hence, the structural solution is identical with one under the flexible carbon price regime.

$$\overline{\chi}^{Com} = \frac{\int_{-\infty}^{p} (p - z) f_{C_v}(z) dz}{c_i} \tag{B.51}$$

In t_1, the regulator sets the carbon price anticipating that her choice impacts the firms' investment decision:

$$\begin{aligned} \max_p E[\mathcal{W}] &= \int_p^{\infty} Q(z) dz + (p - d) Q(p) - \int_0^{\overline{\chi}(p)} (c_i z) dz + \int_{-\infty}^{p} \overline{\chi}(d - t) f_{C_v}(t) dt \\ &\longrightarrow p^{Com} = d \end{aligned} \tag{B.52}$$

The result is identical to the one of *Regulatory flexibility* and the social planner. As the firms are not risk averse, the regulator chooses the Piguvian tax level, that they can perfectly anticipate.

CCfD

When the regulator can offer firms a CCfD in t_1, she sets the carbon price in t_3 after the actual variable costs revealed and the firms made their investment decision. The firms using the emission-free production technology produce, if their variable costs are lower than the conventional technology, i.e., if $c_v < p_s$. The solution yields the socially optimal Pigouvian tax, i.e. $p^{CCfD} = d$. In t_2, the firms invest in the emission-free technology given the announced strike price. The costs remain risky, hence the marginal firm investing in the emission-free technology is characterised by:

$$\begin{aligned} EU(\pi(\overline{\chi})) &= P\Big(p_s - c_v - c_i\overline{\chi} \mid c_v \leq p_s\Big) + P\Big(-c_i\overline{\chi} \mid c_v > p_s\Big) = 0 \\ &= \int_{p_s}^{\infty} (p_s - z) f_{C_v}(z) dz - c_i\overline{\chi} = 0 \\ \longrightarrow \overline{\chi}^{CCfD} &= \frac{\int_{-\infty}^{p_s} (p_s - t) f_{C_v}(z) dz}{c_i} \end{aligned} \tag{B.53}$$

In t_1, the regulator chooses a strike price that maximises expected welfare. She accounts for the firms' reaction to the strike price:

$$\begin{aligned} \max_{p_s} E[\mathcal{W}] &= \int_{p}^{\infty} Q(z)dz + (p-d)Q(p) - \int_{0}^{\overline{\chi}(p)} (c_i z)dz + \int_{-\infty}^{p_s} \overline{\chi}(d-t) f_{C_v}(t) \\ \longrightarrow p_s &= d \end{aligned} \tag{B.54}$$

Welfare ranking

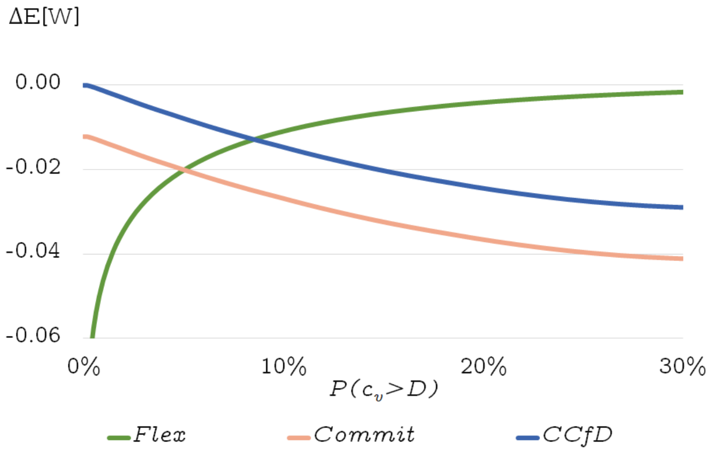
As all carbon pricing regimes result in the socially optimal carbon price and emission-free production capacity, there is no difference in welfare. The absence of risk aversion in this setting leads to equivalent welfare expectations.

B.6. Welfare difference compared to the social optimum in the presence of damage risk, and (ex post) potentially socially not optimal abatement due to an increase in σ_D

Figure B.1 shows a similar effect, when varying the probability of socially not optimal production, $P(C_v > D)$, by altering the expected value of the marginal damage, μ_D.

The welfare of *CCfD* and *Commitment* is not affected by the presence of risk aversion (compare Figure B.1 (with risk aversion) with Figure 7b (no risk aversion)). Hence, as explained in section 4.2, the shortfall in welfare increases with an increased probability of socially not optimal production. Furthermore, the effect is concave in the probability of socially not optimal emission-free production as the welfare-deferring effect is mitigated by decreasing socially optimal investments.

The *Regulatory flexibility* regime does not result in the social optimum if the firms are risk averse. However, as the socially optimal emission-free production capacity decrease, the absolute gap in welfare compared to the social optimum decreases.



$D \sim N(\mu_D \in [4.25; 5.5], \sigma_D^2 = 0.25)$, $\lambda = 1.5$, $Q(p) = 5 - 0.1p$, $c_v = 4$, $c_i = 1$.

Figure B.1.: Difference in welfare compared to social optimum due to change in $P(c_v > D)$ by altering μ_D in the presence of damage risk and potentially welfare-reducing production.

C. Supplementary Material for Chapter 4

C.1. Formal representation of the theoretical model

The cost minimisation problem can be formulated as the Lagrangian $\mathcal{L}$ with the Lagrange multipliers $\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6$:

$$\begin{aligned}
&\mathcal{L}(S, C_1, C_2, C_{12}, \mu_1, \mu_2, ..., \mu_6) = \\
&a\,(d_1 + d_2) + \frac{1}{2}\left[b\,(d_1 + S)^2 + b\,(d_2 - S)^2\right] \\
&+\ \left[m\,(C_1 + C_2) + 2\,C_{12}\right]\tau_c \\
&+\ S\,\tau_s \\
&+ \mu_1\,(d_1 + S - C_{12} - C_1) \\
&+ \mu_2\,(d_2 - S - C_{12} - C_2) \\
&+ \mu_3\,(-S) + \mu_4\,(-C_1) + \mu_5\,(-C_2) + \mu_6\,(-C_{12}))
\end{aligned}$$

The Karush-Kuhn-Tucker (KKT) conditions that need to be fulfilled are as follows:
Stationarity conditions:

$$\frac{\partial \mathcal{L}}{\partial C_1} = m\,\tau_c - \mu_1 - \mu_4 = 0 \tag{C.1}$$

$$\frac{\partial \mathcal{L}}{\partial C_2} = m\,\tau_c - \mu_2 - \mu_5 = 0 \tag{C.2}$$

$$\frac{\partial \mathcal{L}}{\partial C_{12}} = 2\,\tau_c - \mu_1 - \mu_2 - \mu_6 = 0 \tag{C.3}$$

$$\frac{\partial \mathcal{L}}{\partial S} = \tau_s + 2\,b\,(d_1 + 2\,S - d_2) + \mu_1 - \mu_2 - \mu_3 = 0 \tag{C.4}$$

Dual feasibility and complementary slackness:

$$\mu_1 \left(d_1 + S - C_{12} - C_1\right) = 0 \tag{C.5}$$
$$\mu_2 \left(d_2 - S - C_{12} - C_2\right) = 0 \tag{C.6}$$
$$\mu_3 \, S = 0 \tag{C.7}$$
$$\mu_4 \, C_1 = 0 \tag{C.8}$$
$$\mu_5 \, C_2 = 0 \tag{C.9}$$
$$\mu_6 \, C_{12} = 0 \tag{C.10}$$
$$\mu_1,\, \mu_2,\, \mu_3,\, \mu_4,\, \mu_5,\, \mu_6 \,\geq 0 \tag{C.11}$$

Primal feasibility:

$$C_{12} + C_1 \geq d_1 + S \tag{C.12}$$
$$C_{12} + C_2 \geq d_2 - S \tag{C.13}$$
$$S,\, C_1,\, C_2,\, C_{12} \,\geq 0. \tag{C.14}$$

C.2. KKT points

In order to find the optimal KKT points of the optimisation problem and identify the conditions under which they apply, we consider in this section all the realistically possible cases. Those cases correspond to the possible combinations of the Lagrange multipliers of the capacity bookings, C_1, C_2, and C_{12}. The combinations that cannot result in demand being satisfied at both time points, i.e. ($C_1 = C_2 = C_{12} = 0$), ($C_1 = C_{12} = 0$, $C_2 > 0$) and ($C_2 = C_{12} = 0$, $C_1 > 0$), are ruled out. The remaining possible cases are as follows:

1. C_1, $C_2 > 0$ and $C_{12} = 0$ (i.e. $\mu_4 = \mu_5 = 0$ and $\mu_6 \geq 0$)
2. C_1, $C_{12} > 0$ and $C_2 = 0$ (i.e. $\mu_4 = \mu_6 = 0$ and $\mu_5 \geq 0$)
3. C_1, C_2, $C_{12} > 0$ (i.e. $\mu_4 \,, \mu_5 \,, \mu_6 = 0$)
4. $C_{12} > 0$ and C_1, $C_2 = 0$ (i.e. $\mu_6 = 0$ and μ_4, $\mu_5 \geq 0$)
5. C_2, $C_{12} > 0$ and $C_1 = 0$ (i.e. $\mu_5 = \mu_6 = 0$ and $\mu_4 \geq 0$)

In addition to the main cases listed above, all four sub-cases arising from supply constraints (C.5) and (C.6) and their respective Lagrange multipliers μ_1 and μ_2 are considered. For clarity, the storage constraint (C.7) and its respective Lagrange multiplier, μ_3, if applicable, are considered within the four sub-cases.

1. **Case: $\boldsymbol{C_1}$, $\boldsymbol{C_2 > 0}$ and $\boldsymbol{C_{12} = 0}$**
   This case corresponds to $\mu_4 = \mu_5 = 0$ and $\mu_6 \geq 0$. In order to obtain the conditions under which this case becomes valid, we need to go through the associated sub-cases.

a) <u>Supply constraints are binding in t_1 and t_2 (i.e. $\mu_1 \geq 0$, $\mu_2 \geq 0$):</u> From Equations C.1 and C.2 $\mu_1 = \mu_2 = m\,\tau_c$ is obtained. Substituting these into Equation C.3 yields:

$$\mu_6 = 2\,\tau_c\,(1 - m)$$

Since $\mu_6 \geq 0$, the condition for the validity of this case is $m \leq 1$. We now consider two sub-cases where storage S is equal to zero or non-zero, i.e. $\mu_3 \geq 0$ or $\mu_3 = 0$, respectively.

i. <u>$S = 0$:</u> From Equation C.4 with $\mu_1 = \mu_2 = m\,\tau_c$ and $S = 0$, we obtain:

$$\mu_3 = \tau_s + 2\,b\,(d_1 - d_2)$$

Since $\mu_3 \geq 0$, the condition for the storage tariff becomes $\tau_s \geq 2\,b\,(d_2 - d_1)$.

From Equations C.5 and C.6 the optimal values for the capacity bookings are obtained:

$$C_1 = d_1$$
$$C_2 = d_2$$

ii. <u>$S > 0$:</u> From Equation C.4 with $\mu_1 = \mu_2 = m\,\tau_c$ and $\mu_3 = 0$, we obtain:

$$S = \frac{d_2 - d_1}{2} - \frac{\tau_s}{4\,b}$$

Since $S > 0$, the condition for the storage tariff becomes $\tau_s < 2\,b\,(d_2 - d_1)$. From Equations C.5 and C.6 the optimal values for the capacity bookings are obtained:

$$C_1 = \frac{d_1 + d_2}{2} - \frac{\tau_s}{4\,b}$$
$$C_2 = \frac{d_1 + d_2}{2} + \frac{\tau_s}{4\,b}$$

The results indicate that when $m \leq 1$ only ST capacity products (C_1 and C_2) are booked and LT product (C_{12}) is not booked. If the storage tariff is sufficiently low ($\tau_s < 2\,b\,(d_2 - d_1)$), then the traders utilise storages by booking and transporting more than the required demand in t_1 period ($C_1 > d_1$) and less than the demand in t_2 period ($C_2 < d_2$). However, if the storage tariff is sufficiently high ($\tau_s \geq 2\,b\,(d_2 - d_1)$), then the traders do not use storages and book in both periods the respective demand ($C_1 = d_1$, $C_2 = d_2$).

b) Supply constraint is binding in t_1 but not in t_2 (i.e. $\mu_1 \geq 0$, $\mu_2 = 0$):
Substituting $\mu_2 = 0$ into Equation C.2 with $\mu_5 = 0$ yields $m\,\tau_c = 0$. Since by definition $m > 0$ and $\tau_c > 0$, this is not a valid case.

c) Supply constraint is binding in t_2 but not in t_1 (i.e. $\mu_1 = 0$, $\mu_2 \geq 0$):
Substituting $\mu_1 = 0$ into Equation C.1 with $\mu_4 = 0$ yields $m\,\tau_c = 0$. Since by definition $m > 0$ and $\tau_c > 0$, this is not a valid case.

d) Supply constraints are neither binding in t_1 nor in t_2 (i.e. $\mu_1 = 0$, $\mu_2 = 0$):
Substituting $\mu_1 = 0$ into Equation C.1 with $\mu_4 = 0$ yields $m\,\tau_c = 0$. Since by definition $m > 0$ and $\tau_c > 0$, this is not a valid case.

2. **Case: C_1, $C_{12} > 0$ and $C_2 = 0$**
This case corresponds to $\mu_4 = \mu_6 = 0$ and $\mu_5 \geq 0$. This case is possible for $m > 1$ only if $d_1 > d_2$. However, since by definition $d_2 > d_1$, this is not a valid case.

3. **Case: C_1, C_2, $C_{12} > 0$**

This case corresponds to $\mu_4 = \mu_5 = \mu_6 = 0$. In order to obtain the conditions under which this case becomes valid, we need to go through the associated sub-cases.

a) Supply constraints are binding in t_1 and t_2 (i.e. $\mu_1 \geq 0$, $\mu_2 \geq 0$):
From Equations C.1 and C.2 $\mu_1 = \mu_2 = m\,\tau_c$ is obtained. Substituting these into Equation C.3 yields:

$$m = 1$$

We now consider two sub-cases where storage S is equal to zero or non-zero, i.e. $\mu_3 \geq 0$ or $\mu_3 = 0$.

i. $S = 0$: From Equation C.4 with $\mu_1 = \mu_2 = m\,\tau_c$ and $S = 0$, we obtain:

$$\mu_3 = \tau_s + 2\,b\,(d_1 - d_2)$$

Since $\mu_3 \geq 0$, the condition for the storage tariff becomes $\tau_s \geq 2\,b\,(d_2 - d_1)$.

By rearranging the condition for τ_s to obtain $d_2 - d_1 \leq \frac{\tau_s}{2\,b}$ and plugging into Equation C.5 subtracted from Equation C.6, we obtain:

$$C_2 - C_1 \leq \frac{\tau_s}{2\,b}$$

We do not obtain unique results for C_1, C_2, and C_{12}. Instead, all combinations of positive C_1, C_2, and C_{12} that fulfil the condition above in addition to the constraints stated in Equations C.12) and C.13 are KKT points and hence optimal solutions.

ii. $S > 0$: From Equation C.4 with $\mu_1 = \mu_2 = m\,\tau_c$ and $\mu_3 = 0$, we obtain:

$$S = \frac{d_2 - d_1}{2} - \frac{\tau_s}{4\,b}$$

Since $S > 0$, the condition for the storage tariff becomes $\tau_s < 2\,b\,(d_2 - d_1)$. By rearranging the condition for τ_s to obtain $d_2 - d_1 > \frac{\tau_s}{2\,b}$ and plugging into Equation C.5 subtracted from Equation C.6, we obtain:

$$C_2 - C_1 > \frac{\tau_s}{2\,b}$$

Again, we do not obtain unique results for C_1, C_2, and C_{12}. All combinations of positive C_1, C_2, and C_{12} that fulfil the condition above in addition to the constraints stated in Equations C.12 and C.13 are KKT points and hence optimal solutions.

b) Supply constraint is binding in t_1 but not in t_2 (i.e. $\mu_1 \geq 0$, $\mu_2 = 0$): Substituting $\mu_2 = 0$ into Equation C.2 with $\mu_5 = 0$ yields $m\,\tau_c = 0$. Since by definition $m > 0$ and $\tau_c > 0$, this is not a valid case.

c) Supply constraint is binding in t_2 but not in t_1 (i.e. $\mu_1 = 0$, $\mu_2 \geq 0$): Substituting $\mu_1 = 0$ into Equation C.1 with $\mu_4 = 0$ yields $m\,\tau_c = 0$. Since by definition $m > 0$ and $\tau_c > 0$, this is not a valid case.

d) Supply constraints are neither binding in t_1 nor in t_2 (i.e. $\mu_1 = 0$, $\mu_2 = 0$): Substituting $\mu_1 = 0$ into Equation C.1 with $\mu_4 = 0$ yields $m\,\tau_c = 0$. Since by definition $m > 0$ and $\tau_c > 0$, this is not a valid case.

4. **Case: $C_1 = C_2 = 0$ and $C_{12} > 0$**
This case corresponds to μ_4, $\mu_5 \geq 0$ and $\mu_6 = 0$. In order to obtain the conditions under which this case becomes valid, we need to go through the associated sub-cases.

a) Supply constraints are binding in t_1 and t_2 (i.e. $\mu_1 \geq 0$, $\mu_2 \geq 0$): From Equations $C.5$ and $C.6$ it follows that $S = \frac{d_2 - d_1}{2}$ and the corresponding Lagrange multiplier $\mu_3 = 0$. The value for the long-term capacity booking is also obtained as $C_{12} = \frac{d_2 + d_1}{2}$. Stationarity conditions then take the form:

$$\begin{aligned} m\,\tau_c - \mu_1 - \mu_4 &= 0 \\ m\,\tau_c - \mu_2 - \mu_5 &= 0 \\ 2\,\tau_c - \mu_1 - \mu_2 &= 0 \\ \tau_s + \mu_1 - \mu_2 &= 0 \end{aligned}$$

Solving the system of equations above yields the following results:

$$\begin{aligned} \mu_1 &= \tau_c - \frac{\tau_s}{2} \\ \mu_2 &= \tau_c + \frac{\tau_s}{2} \\ \mu_4 &= \tau_c\,(m-1) + \frac{\tau_s}{2} \\ \mu_5 &= \tau_c\,(m-1) - \frac{\tau_s}{2} \end{aligned}$$

From the condition $\mu_1,\, \mu_2,\, \mu_4,\, \mu_5 \,\geq\, 0$ it follows:

$$\begin{aligned} 2\,\tau_c &\geq \tau_s \\ m &\geq 1 + \frac{\tau_s}{2\,\tau_c} \end{aligned}$$

To fulfil both equations simultaneously, $m \leq 2$ is required. This implies when multiplier is sufficiently high, but still below 2, and the storage tariff is sufficiently low, then the transported volumes align and only long-term capacity is booked.

b) Supply constraint is binding in t_1 but not in t_2 (i.e. $\mu_1 \geq 0$, $\mu_2 = 0$):

In this case, the stationary conditions reduce to:

$$\begin{aligned} (m-2)\,\tau_c &= \mu_4 \\ m\,\tau_c &= \mu_5 \\ 2\,\tau_c &= \mu_1 \\ \tau_s + 2\,b\,(d_1 + 2S - d_2) + 2\,\tau_c &= \mu_3 \end{aligned}$$

So $m \geq 2$ since $\mu_4 \geq 0$.
In addition we get from Equations C.5 and C.9 that:

$$\begin{aligned} C_{12} &= d_1 + S \\ S &= \frac{d_1 - d_2}{2} - \frac{1}{2b}\,\tau_c - \frac{1}{4b}\,\tau_s \end{aligned}$$

Substituting C_{12} into Equation C.13, we obtain:

$$S \geq \frac{d_2 - d_1}{2}$$

Substituting the previously obtained storage value into the inequality above yields:

$$0 \geq 2\tau_c + \tau_s.$$

This is not possible since τ_s, $\tau_c > 0$. Hence, this case is not valid.

c) Supply constraint is binding in t_2 but not in t_1 (i.e. $\mu_1 = 0$, $\mu_2 \geq 0$): In this case, the stationary conditions reduce to:

$$\begin{aligned} m\,\tau_c &= \mu_4 \\ 2\,\tau_c &= \mu_2 \\ (m-2)\,\tau_c &= \mu_5 \\ \tau_s + 2\,b\,(d_1 + 2S - d_2) - 2\,\tau_c &= \mu_3 \end{aligned}$$

The case is valid for $m \geq 2$ since $\mu_5 \geq 0$.

We now consider two sub-cases where storage S is equal to zero or non-zero, i.e. $\mu_3 \geq 0$ or $\mu_3 = 0$:

i. $S = 0$: From Equations C.12 and C.13, and the assumption $d_2 > d_1$ we derive:

$$C_{12} = d_2$$

To ensure $\mu_3 \geq 0$ the following condition needs to hold:

$$\tau_s \geq 2\tau_c + 2b\,(d_2 - d_1)$$

It can be seen that in this case a portion of C_{12} equal to $d_2 - d_1$ is not utilised i.e. wasted in t_1.

ii. $S > 0$: In this case $\mu_3 = 0$.

Plugging the given information into Equations C.4 and C.6 allows to solve for S and C_{12}:

$$\begin{aligned} S &= \frac{d_2 - d_1}{2} + \frac{1}{2b}\tau_c - \frac{1}{4b}\,\tau_s \\ C_{12} &= \frac{d_2 + d_1}{2} - \frac{1}{2b}\tau_c + \frac{1}{4b}\,\tau_s \end{aligned}$$

To ensure $S > 0$ and that the supply constraint as shown in Equation C.12 is satisfied, τ_s has to lie in the range between:

$$2\tau_c \leq \tau_s < 2\tau_c + 2b\,(d_2 - d_1)$$

If $\tau_s > 2\tau_c$, a portion of C_{12} equal to $\frac{\tau_s}{2\,b} - \frac{\tau_c}{b}$ is wasted in t_1. For $\tau_s = 2\tau_c$, this term becomes zero and thus transmissions in t_1 and t_2 align and no capacity booking is wasted.

The results indicate that under the condition $m \geq 2$ only LT capacity is booked for both periods. When the storage tariff is sufficiently high ($\tau_s > 2\,\tau_c$), storage utilisation is not sufficient to align transports in t_1 and t_2 such that some LT capacity is wasted in t_1. For $\tau_s = 2\,\tau_c$, transported volumes in both periods align, such that no LT capacity is wasted.

d) Supply constraints are neither binding in t_1 nor in t_2 (i.e. $\mu_1 = 0$, $\mu_2 = 0$):
In this case the stationary conditions reduce to:

$$\begin{aligned} m\,\tau_c &= \mu_4 \\ m\,\tau_c &= \mu_5 \\ 2\,\tau_c &= 0 \\ \tau_s + 2\,b\,(d_1 + 2\,S - d_2) &= \mu_3 \end{aligned}$$

This is not a valid case since it yields $\tau_c = 0$, where by definition $\tau_c > 0$.

5. **Case: $C_1 = 0$ and C_2, $C_{12} > 0$**

This case corresponds to $\mu_4 \geq 0$ and $\mu_5 = \mu_6 = 0$. In order to obtain the conditions under which this case becomes valid, we need to go through the associated sub-cases.

a) Supply constraints are binding in t_1 and t_2 (i.e. $\mu_1 \geq 0$, $\mu_2 \geq 0$):
From Equations $C.5$ and $C.6$ it follows:

$$S = \frac{d_2 - d_1}{2} - \frac{C_2}{2} \qquad \text{(C.15)}$$

Since $\mu_5 = \mu_6 = 0$, from Equations C.1, C.2 and C.3 we obtain:

$$\begin{aligned} \mu_1 &= \tau_c\,(2 - m) \\ \mu_2 &= m\,\tau_c \\ \mu_4 &= 2\,\tau_c\,(m - 1) \end{aligned}$$

From the condition that μ_1, μ_2, $\mu_4 \geq 0$ it follows that:

$$1 \leq m \leq 2$$

Substituting the previously obtained μ_1 and μ_2 into Equation C.4 yields the following:

$$\tau_s + 2\,b\,(d_1 + 2\,S - d_2) + 2\,\tau_c\,(1 - m) - \mu_3 = 0 \tag{C.16}$$

We now consider two sub-cases where storage, S, is equal to zero or non-zero, i.e. $\mu_3 \geq 0$ or $\mu_3 = 0$:

i. <u>$S = 0$:</u> In this case $\mu_3 \geq 0$. Setting Equation C.15 to zero, we obtain:

$$C_2 = d_2 - d_1$$
$$C_{12} = d_1$$

Similarly, substituting $S = 0$ in Equation C.16 yields:

$$\mu_3 = \tau_s + 2\,b\,(d_1 - d_2) + 2\,\tau_c\,(1 - m)$$

Since $\mu_3 \geq 0$, the condition for this case becomes:

$$\tau_s \geq 2\,b\,(d_2 - d_1) + 2\,\tau_c\,(m - 1)$$

which can be rewritten as:

$$m \leq 1 + \frac{\tau_s}{2\,\tau_c} - \frac{b}{\tau_c}\,(d_2 - d_1) \tag{C.17}$$

The implication of this finding is that given the multiplier m and model parameters, when the storage tariff τ_s is sufficiently large no gas will be stored in the storage. Similarly, given the parameters, when the multiplier m is less than or equal to the right-hand side of the condition presented in Equation C.17 no gas will be stored in the storage.

ii. <u>$S > 0$:</u> In this case $\mu_3 = 0$. From Equation C.16 the optimal storage value then becomes:

$$S = \frac{d_2 - d_1}{2} - \frac{\tau_s}{4b} + \frac{\tau_c(m-1)}{2b} \tag{C.18}$$

From Equations C.5 and C.6, we similarly obtain the optimal values for the capacities:

$$C_2 = \frac{\tau_s}{2b} - \frac{\tau_c(m-1)}{b} \tag{C.19}$$

$$C_{12} = \frac{d_2 + d_1}{2} - \frac{\tau_s}{4b} + \frac{\tau_c(m-1)}{2b} \tag{C.20}$$

Taking into account that S, C_{12}, $C_2 > 0$, the conditions for the validity of the case are obtained as follows:

$$\tau_s < 2\,b\,(d_2 - d_1) + 2\,\tau_c\,(m-1)$$

which can be rewritten as:

$$m > 1 + \frac{\tau_s}{2\,\tau_c} - \frac{b}{\tau_c}\,(d_2 - d_1) \tag{C.21}$$

and:

$$m < 1 + \frac{\tau_s}{2\,\tau_c} \tag{C.22}$$

The results indicate that while the conditions stated in Equations C.21 and C.22 are valid, i.e.

$$1 + \frac{\tau_s}{2\,\tau_c} - \frac{b}{\tau_c}\,(d_2 - d_1) < \; m \; < \; 1 + \frac{\tau_s}{2\,\tau_c} \tag{C.23}$$

long-term capacity C_{12} is booked for both periods, short-term capacity C_2 is booked for t_2, no short-term capacity C_1 is booked for t_1, and the storages are utilised.

b) Supply constraint is binding in t_1 but not in t_2 (i.e. $\mu_1 \geq 0$, $\mu_2 = 0$):

Substituting $\mu_5 = \mu_2 = 0$ into Equation C.2 yields:

$$m\,\tau_c = 0$$

Since both m and τ_c are by definition non-zero, this case is not valid.

c) Supply constraint is binding in t_2 but not in t_1 (i.e. $\mu_1 = 0$, $\mu_2 \geq 0$):

Considering that $\mu_5 = \mu_6 = \mu_1 = 0$, we obtain from Equations C.1, C.2 and C.3:

$$2\,\tau_c = m\,\tau_c$$
$$m = 2$$

We now consider two sub-cases where storage S is equal to zero or non-zero, i.e. $\mu_3 \geq 0$ or $\mu_3 = 0$, respectively.

i. $S = 0$: In this case $\mu_3 \geq 0$. From Equation C.4 we obtain:

$$\mu_3 = \tau_s + 2\,b\,(d_1 - d_2) - m\,\tau_c$$

Since $\mu_3 \geq 0$, the condition for this case becomes:

$$\tau_s \geq 2\,b\,(d_2 - d_1) + m\,\tau_c$$

The conditions from the supply constraints are as follows:

$$C_{12} \geq d_1$$
$$C_{12} = d_2 - C_2$$

It can be seen that there exists no unique solution for C_2, and C_{12}. All combinations of positive C_2 and C_{12} that fulfil the conditions above are KKT points and hence optimal solutions.

ii. $S > 0$: In this case $\mu_3 = 0$. From Equation C.4 we obtain:

$$S = \frac{d_2 - d_1}{2} + \frac{m\tau_c}{4\,b} - \frac{\tau_s}{4\,b}$$

Since $S > 0$, the condition for this case becomes:

$$\tau_s < 2\,b\,(d_2 - d_1) + m\,\tau_c$$

The conditions from the supply constraints are as follows:

$$C_{12} \geq d_1 + S$$
$$C_{12} = d_2 - C_2 - S$$

Again, there exists no unique solution for C_2, and C_{12}. All combinations of positive C_2 and C_{12} that fulfil the conditions above are KKT points and hence optimal solutions.

d) Supply constraints are neither binding in t_1 nor in t_2 (i.e. $\mu_1 = 0$, $\mu_2 = 0$): Again, substituting $\mu_5 = \mu_2 = 0$ in Equation C.2 yields:

$$m\,\tau_c = 0$$

Similarly, substituting $\mu_1 = \mu_2 = \mu_6 = 0$ in Equation C.3 yields:

$$2\,\tau_c = 0$$

Since both m and τ_c are by definition non-zero, this case is not valid.

C.3. Prices in region A

Deriving prices in region A is less straightforward, since for the sake of simplicity no demand in region A is integrated. To derive the prices in region A one can add a fictional demand d_{A1} and d_{A2} to the procurement cost equation and differentiate it by d_{A1} and d_{A2}. Alternatively, one can subtract the Lagrange multipliers μ_1

and μ_2 from the prices in region B, since the Lagrange multipliers represent the marginal costs for transporting gas from region A to B.

$$P_{A1} = P_{B1} - \mu_1 = \begin{cases} a + b\, d_1 & \text{for } m < \underline{m} \\ a + b\left(\dfrac{d_1 + d_2}{2}\right) + \tau_c\,(m-1) - \dfrac{\tau_s}{2} & \text{for } \underline{m} < m < \overline{m} \\ a + b\left(\dfrac{d_1 + d_2}{2}\right) & \text{for } \overline{m} < m \end{cases}$$

$$P_{A2} = P_{B2} - \mu_2 = \begin{cases} a + b\, d_2 & \text{for } m < \underline{m} \\ a + b\left(\dfrac{d_1 + d_2}{2}\right) - \tau_c\,(m-1) + \dfrac{\tau_s}{2} & \text{for } \underline{m} < m < \overline{m} \\ a + b\left(\dfrac{d_1 + d_2}{2}\right) & \text{for } \overline{m} < m \end{cases} \tag{C.24}$$

The functions describing the consumer prices in region A are plotted in Figure C.1. Although individual consumer prices are influenced by m for $m < \overline{m}$, unweighted average prices remain constant.

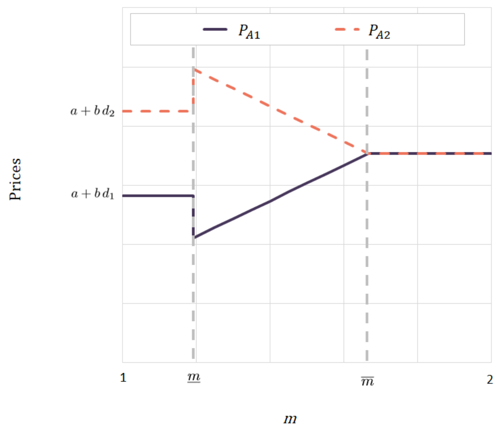


Figure C.1.: Development of prices in region A at time periods t_1 and t_2 with respect to the multiplier

The average price in region A is equal to the gas procurement prices which arise when overall demand is split evenly among periods:

$$\frac{P_{A2} + P_{A1}}{2} = a + b\,(\frac{d_2 + d_1}{2})$$

As can be seen in Figure C.1, when $m \leq \underline{m}$, the prices in region A are independent of the multiplier due to storages not being used and prices solely reflecting the costs for gas production. In the domain $\underline{m} < m < \overline{m}$, as storages start being utilised and the marginal costs of storage utilisation is included in the prices, an offset in prices (decrease in t_1, increase in t_2) occurs. With increasing m, prices in region A start converging as production volumes increasingly align. With $m \geq \overline{m}$, production volumes fully converge and the same prices in both periods are observed in region A.

C.4. Surpluses and deadweight loss when no feasible $\underline{m}$ and $\overline{m}$ exist

Depending on the tariff structures (i.e. the proportion of τ_s and τ_c), $\underline{m}$ and $\overline{m}$ may not exist in the feasible multiplier range of $1 \leq m \leq 2$. In such a case, the previously identified domains $m < \underline{m}$ and $m > \overline{m}$ do not exist. Hence, Proposition 4.3.4 holds throughout the feasible multiplier range (i.e. $1 \leq m \leq 2$) and storages are utilised as well as ST and LT capacities are booked for all such multipliers.

The surpluses of the agents in the model and the deadweight loss are plotted in Figure C.2.[100]

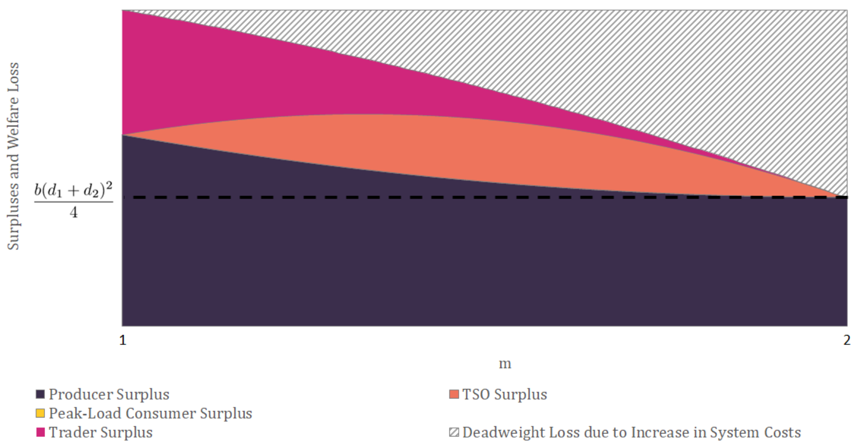


Figure C.2.: Surpluses and deadweight loss when no feasible $\underline{m}$ and $\overline{m}$ exist

[100]The parameters assumed for the figure are as follows: $d_1 = 11$, $d_2 = 30$, $\tau_c = 2.9$, $\tau_s = 5.7$, $a = 4$, $b = 0.15$.

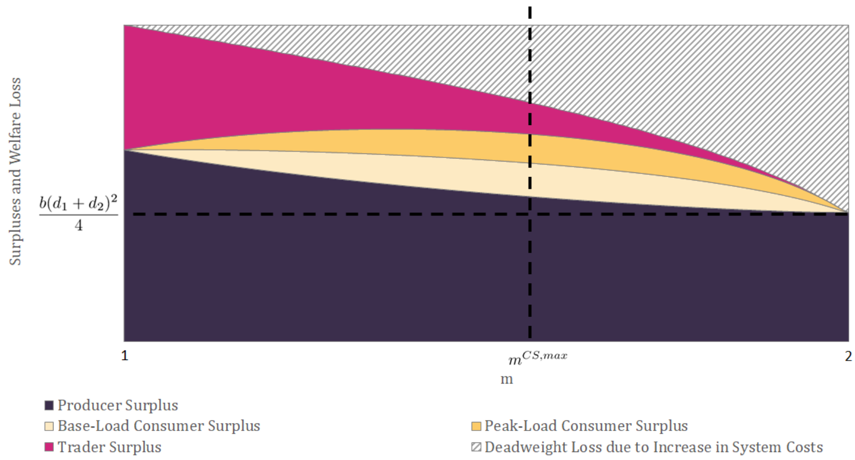


Figure C.3.: Surpluses and deadweight loss when no feasible $\underline{m}$ and $\overline{m}$ exist in the case where τ_c is adjusted

For the case when transmission tariffs (τ_c) are adjusted such that the TSO does not earn a surplus, the surpluses of the agents in the model and the deadweight loss are plotted in Figure C.3. The multiplier level that maximises the total consumer surplus is equal to $m^{CS,max} = 1 + \frac{\tau_s}{\tau_c^{adj}}$.

D. Supplementary Material for Chapter 5

D.1. Theoretical analysis

Lemma D.1.1. *With T being the total number of time periods, it is optimal to solely book long-term capacity covering all periods if the duration of the short-term capacities products multiplied by the respective multiplier exceed T.*

Proof. The cost for a short-term (ST) capacity product is equal to $t_p m_p \tau_c$, with t_p being the duration of the capacity product p, m_p being the multiplier of the respective capacity product and τ_c being the tariff for the long-term (LT) capacity. For LT capacity that covers all the periods, no multiplier is applied and the cost is equal to $T\tau_c$. It is clear that if $t_p m_p > T$ the cost for the ST capacity product becomes higher than the cost of LT capacity. In this situation, it is always optimal to book only LT capacity. This concludes the proof. ∎

In the paper at hand we assess the effects of multipliers in a setting with twelve periods, in which each period represents one month. A yearly (LT) capacity covering all the twelve periods, a quarterly capacity covering three periods and a monthly capacity covering one period are offered.

The cost of one unit of quarterly capacity, covering three periods, is equal to $3m_q \tau_c$, with m_q being the quarterly multiplier. For LT capacity, covering all the twelve periods, no multiplier is applied, so the cost is equal to $12\tau_c$. If $m_q > 4$, the cost of the quarterly capacity becomes higher than the LT capacity. The cost of one unit of monthly capacity, covering one period, is equal to $m_m \tau_c$, with m_m being the multiplier for monthly capacity. If $m_m > 12$, the cost of the monthly capacity becomes higher than the LT capacity. Therefore, in the setting with twelve periods, and multipliers of $m_q > 4$ and $m_m > 12$, it is always optimal for a cost-minimising trader to book only LT capacity.

Lemma D.1.2. *If demand for transmission capacity is fully inelastic where it equals to $X - e$ in t_p periods and X in the remaining consecutive $T - t_p$ periods, under the condition $m_p > \frac{T}{T-t_p}$, only LT capacity is booked in the optimal solution and some capacity rights remain unused.*

Proof. A trader can either book a combination of LT and ST capacity or choose to book LT capacity only. In case it is decided to mix both types of capacities, the trader procures $X - e$ units of LT capacity, valid in all T periods, and buys additionally e units of ST capacity for the remaining consecutive $T - t_d$ periods with higher demand. t_p represents the duration of the ST capacity product p. Other combinations would result in higher costs. If it is decided to book only LT capacity instead, the trader books X units of LT capacity for the whole period. It would be optimal to book only LT capacity if the associated costs were lower, i.e. if the inequality below would hold:

$$\tau_c[(X - e)\,T + e\,(T - t_p)\,m_p] > \tau_c X\,T$$

which then simplifies to:

$$m_p > \frac{T}{(T - t_p)}$$

■

The situation of fully inelastic demand as assumed in the Lemma would occur if storages are exhausted. Applying the Lemma to a setting with twelve periods where each period represents one month—and a yearly capacity (LT) covers all the twelve periods, a quarterly capacity covers three periods and a monthly capacity covers one period—results in the following thresholds for multipliers:

In case demand equals X in eleven months and is lower in the remaining one month, solely LT capacity is booked if the monthly multiplier exceeds $m_p > \frac{12}{(12-1)} = 1.\overline{09}$. In case demand equals X in nine months and is lower in the remaining consecutive three months, solely LT capacity is booked if the monthly multiplier exceeds $m_p > \frac{12}{(12-3)} = 1.\overline{33}$. The multiplier threshold in this case is higher, as a larger share of LT capacity is wasted. The two examples show that, even in the presence of moderate multipliers, it can be optimal for traders to let some capacity remain unused.

If demand is not fully elastic, but transports are not fully aligned even in the presence of multipliers that induce a capacity pricing regime (see Lemma D.1.1), then multipliers causing only LT capacity to be booked would lie between the thresholds resulting from Lemma D.1.1 and Lemma D.1.2. This would be the case if flexibility is available but the marginal cost curve for flexibility is steep.

D.2. Reference case and model validation

We validate our model against historical results for the 2018 gas year covering the period 01. October 2017–30. September 2018. For this purpose, we consider the reference case where every region has the default EU average multiplier ($m4$) levels. The simulated storage levels, imports from Russia and the price levels are then compared with the historical levels.

In Figure D.1 the simulated monthly storage levels in the EU are plotted against the historical levels.[101] Note that LNG storages are not included. It can be seen that the simulated storage levels during the winter period lie slightly below the historical levels. Nevertheless, the storage levels then follow the historical levels very closely in the summer period.

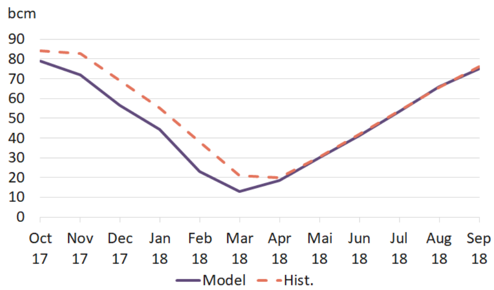


Figure D.1.: Simulated and the historical monthly storage levels in the EU

In Figure D.2 the simulated monthly imported gas volumes from Russia are plotted against the historical volumes.[102] The simulated import volumes lie slightly above the historical volumes in the winter period, while they lie slightly below the historical volumes in the summer period. The difference between the simulated and the historical results in the total yearly imported volumes is less than 1%.

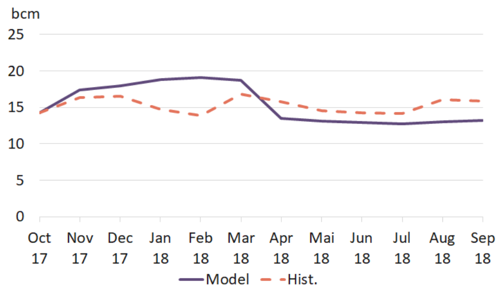


Figure D.2.: Simulated and the historical monthly import volumes from Russia into the EU

In Figure D.3 the average prices in the considered regions for the gas year 2018 and the historical TTF price during this period are plotted. It can be seen that the average price in the Central region is very close to the average TTF price. The price levels in the other regions are higher than the price level in the Central and lie in realistic ranges. Note that the prices for the Baltic and the South East

[101] Historical storage levels for European countries are obtained from the AGSI+ platform (`https://agsi.gie.eu/`).

[102] Historical imports are derived from the IEA Gas Trade Flows (GTF) service (`https://www.iea.org/reports/gas-trade-flows`).

regions include on top of the simulated prices markups of 3 EUR/MWh and 1.5 EUR/MWh, respectively. This is done in order to represent the realistic price levels observed in these regions due to having less competitive market structures.

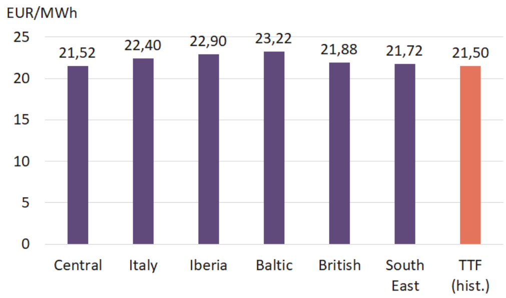
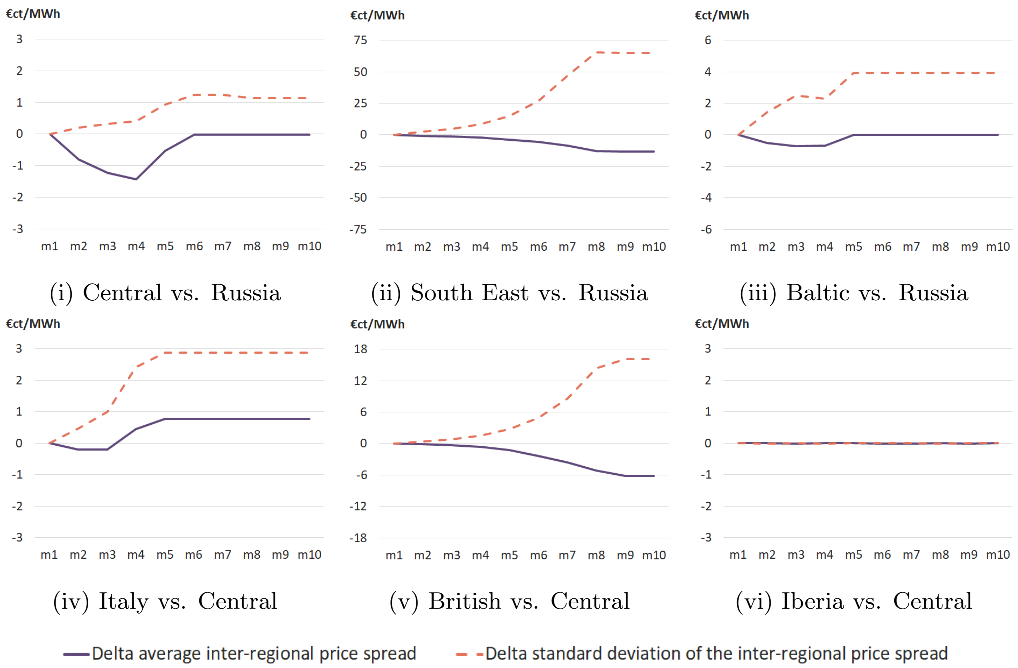


Figure D.3.: Simulated regional price levels for the gas year 2018 and the historical TTF price in the corresponding period

D.3. Overview of regional price spreads



(i) Central vs. Russia (ii) South East vs. Russia (iii) Baltic vs. Russia

(iv) Italy vs. Central (v) British vs. Central (vi) Iberia vs. Central

Figure D.4.: Change in the average inter-regional price spread and its standard deviation with respect to import region when each region adjusts their multipliers individually

Figure D.4 plots the average inter-regional price spread as well as its standard deviation with respect to multipliers when regions adjust their multipliers individually in the default case. It can be seen that the change in the average regional price spreads directly follow the change in average prices due to tariff adjustments (see Figure 5.5). The standard deviation of the regional price spreads, which can also be referred to as the volatility of the regional price spreads, is shown to be increasing with multipliers in all regions except Iberia.

D.4. Overview of external effects on consumer surplus

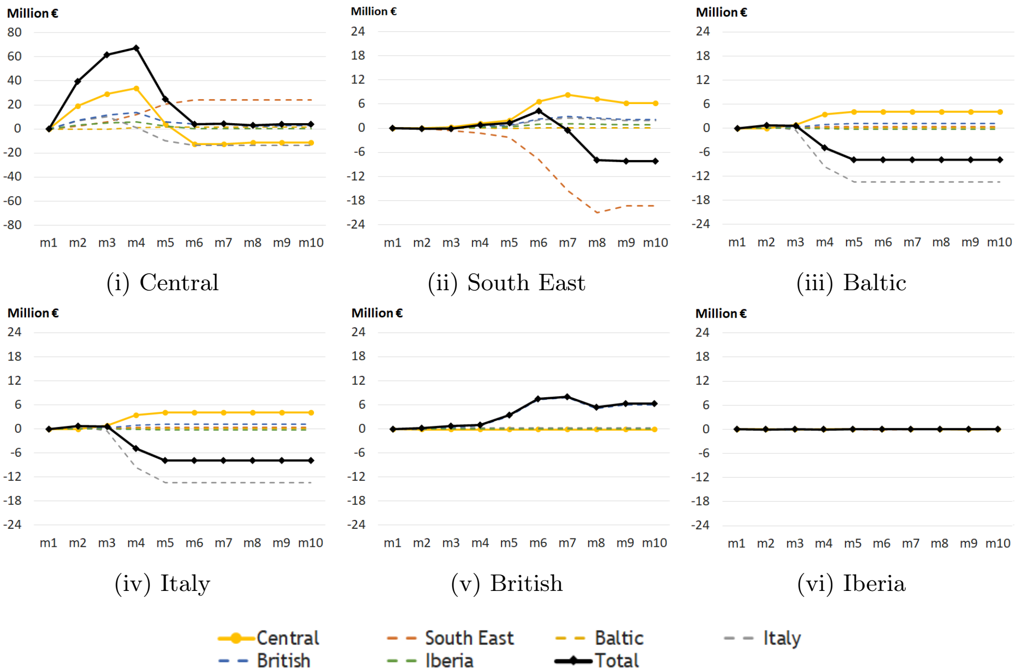


Figure D.5.: The changes in consumer surplus in the regions and the total impact in the EU when multipliers are adjusted individually in the regions: (a) Central, (b) South East, (c) Baltic, (d) Italy, (e) British, (f) Iberia.

Bibliography

ACER (2019a). ACER Market Monitoring Report 2018 - Gas Wholesale Market Volume. Technical report, ACER, Ljubljana.

ACER (2019b). Annual Report on Contractual Congestion at Interconnection Points, Period Covered: 2018. Technical report, ACER, Ljubljana.

ACER (2019c). Annual Report on Contractual Congestion at Interconnection Points, Period Covered: 2018. Technical report, ACER, Ljubljana.

ACER (2020a). 7th ACER Report on congestion in the EU gas markets and how it is managed. Period covered: 2019. Technical report, ACER, Ljubljana.

ACER (2020b). ACER Market Monitoring Report 2019 - Gas Wholesale Market Volume. Technical report, ACER, Ljubljana.

ACER (2022). Decision No 11/2022 of the European Union Agency for the Cooperation of Energy Regulators of 8 August 2022 on the alternative bidding zone configurations to be considered in the bidding zone review process.

ACER and CEER (2019). The Bridge Beyond 2025 - Conclusions Paper. Technical report, ACER, Ljubljana and CEER, Brussels.

Agency, I. E. (2019). The future of hydrogen. `https://www.iea.org/reports/the-future-of-hydrogen`, accessed: 20.08.2021.

Alexander, P. and Moran, D. (2013). Impact of perennial energy crops income variability on the crop selection of risk averse farmers. *Energy Policy*, 52:587–596. Special Section: Transition Pathways to a Low Carbon Economy.

António Guterres (2022). Cop27 climate summit.

Arrow, K. and Lind, R. C. (1970). Uncertainty and the evaluation of public investment decisions. *American Economic Review*, 60(3):364–78.

Arrow, K. J. and Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, pages 265–290.

Baldursson, F. M. and Von der Fehr, N.-H. M. (2004). Price volatility and risk exposure: on market-based environmental policy instruments. *Journal of Environmental Economics and Management*, 48(1):682–704.

Ballard, C. L. and Fullerton, D. (1992). Distortionary taxes and the provision of public goods. *Journal of Economic Perspectives*, 6(3):117–131.

Banal-Estañol, A. and Ottaviani, M. (2006). Mergers with product market risk. *Journal of Economics & Management Strategy*, 15(3):577–608.

Bermúdez, A., González-Díaz, J., González-Diéguez, F. J., and González-Rueda, Ã. M. (2016). Gas transmission networks in Europe: Connections between different entry-exit tariff methodologies. *Applied Energy*, 177:839–851.

Bird, L., Lew, D., Milligan, M., Carlini, E. M., Estanqueiro, A., Flynn, D., Gomez-Lazaro, E., Holttinen, H., Menemenlis, N., Orths, A., Eriksen, P. B., Smith, J. C., Soder, L., Sorensen, P., Altiparmakis, A., Yasuda, Y., and Miller, J. (2016). Wind and solar energy curtailment: A review of international experience. *Renewable and Sustainable Energy Reviews*, 65:577–586.

Blundell, W., Gowrisankaran, G., and Langer, A. (2020). Escalation of scrutiny: The gains from dynamic enforcement of environmental regulations. *American Economic Review*, 110(8):2558–85.

BMU (2021). Eckpunkte für eine förderrichtlinie klimaschutzverträge zur umsetzung des pilotprogramms čarbon contracts for differencë. `https://www.bmu.de/fileadmin/Daten_BMU/Download_PDF/Klimaschutz/eckpunktepapier_klimaschutzvertraege_ccfd_bf.pdf`, accessed: 20.08.2021.

BNetzA (2019). Veröffentlichung von Stellungnahmen gemäß Art.26 Abs.3 S.1 der Verordnung (EU) Nr.2017/460. Technical report, BNetzA, Bonn.

BNetzA (2022). SMARD - Electricity market data for Germany.

Boiteux, M. (1949). La tarification des demandes en point: application de la theorie de la vente au cout marginal. *Revue General de l'Electricite*, 58:321–340. (translated as Peak Load Pricing. Journal of Business. 33, 157-179).

Borch, K. (1962). Equilibrium in a reinsurance market. *Econometrica*, 30(3):424–444.

Borenstein, S. (2016). The economics of fixed cost recovery by utilities. *The Electricity Journal*, 29:5–12.

BP (2019). Statistical Review of World Energy. Technical report, BP, London.

Brändle, G., Schönfisch, M., and Schulte, S. (2021). Estimating long-term global supply costs for low-carbon hydrogen. *Applied Energy*, 302:117481.

Bundesverfassungsgericht (2021). Constitutional complaints against the federal climate change act partially successful. `https://www.bundesverfassungsgericht.de/SharedDocs/Pressemitteilungen/EN/2021/bvg21-031.html;jsessionid=C62AC122F5A4EAFDC942FEB0AB0F9B96.1_cid377`, accessed: 20.08.2021.

Burke, P. J. and Yang, H. (2016). The price and income elasticities of natural gas demand: International evidence. *Energy Economics*, 59:466–474.

Çam, E. and Lencz, D. (2021a). Internal and external effects of pricing short-term gas transmission capacity via multipliers. EWI Working Papers 2021-4, Energiewirtschaftliches Institut an der Universitaet zu Koeln (EWI).

Çam, E. and Lencz, D. (2021b). Pricing short-term gas transmission capacity: A theoretical approach to understand the diverse effects of the multiplier system. *Energy Economics*, 95:105095.

Çam, E. and Lencz, D. (2021). Pricing short-term gas transmission capacity: A theoretical approach to understand the diverse effects of the multiplier system. *Energy Economics*, 95:105095.

Cervigni, G., Conti, I., Glachant, J.-M., Tesio, E., and Volpato, F. (2019). Towards an Efficient and Sustainable Tariff Methodology for the European Gas Transmission Network. Technical report, Florence School of Regulation.

Chao, H.-P. and Wilson, R. (1993). Option value of emission allowances. *Journal of Regulatory Economics*, 5(3):233–249.

Chiappinelli, O., Gerres, T., Neuhoff, K., Lettow, F., de Coninck, H., Felsmann, B., Joltreau, E., Khandekar, G., Linares, P., Richstein, J., et al. (2021). A green covid-19 recovery of the eu basic materials sector: identifying potentials, barriers and policy solutions. *Climate Policy*, pages 1–19.

Chiappinelli, O. and Neuhoff, K. (2020). Time-consistent carbon pricing: The role of carbon contracts for differences. *DIW Berlin Discussion Paper*, 1859:1–35.

Cholteeva, Y. (2020). Constraint payments: rewarding wind farms for switching off.

Christiansen, V. and Smith, S. (2015). Emissions taxes and abatement regulation under uncertainty. *Environmental and Resource Economics*, 60(1):17–35.

Cialani, C. and Mortazavi, R. (2018). Household and industrial electricity demand in Europe. *Energy Policy*, 122:592–600.

Cramton, P., MacKay, D. J., Ockenfels, A., and Stoft, S. (2017). *Global carbon pricing: the path to climate cooperation*. The MIT Press.

Czock, B., Sitzmann, A., and Zinke, J. (2022). The place beyond the lines - Efficient storage allocation in a spatially unbalanced power system with a high share of renewables. Bonn. display.

Datta, A. and Somanathan, E. (2016). Climate policy and innovation in the absence of commitment. *Journal of the Association of Environmental and Resource Economists*, 3(4):917–955.

Diamond, P. A. (1978). The role of a stock market in a general equilibrium model with technological uncertainty. In *Uncertainty in Economics*, pages 209–229. Elsevier.

Dieckhöner, C. (2012). Simulating security of supply effects of the nabucco and south stream projects for the european natural gas market. *The Energy Journal*, 33(3):153–181.

Dieckhöner, C., Lochner, S., and Lindenberger, D. (2013). European natural gas infrastructure: The impact of market developments on gas flows and physical market integration. *Applied Energy*, 102:994–1003.

Dixit, A. K., Dixit, R. K., and Pindyck, R. S. (1994). *Investment under uncertainty*. Princeton university press.

DNV-GL (2018). Study on the Future Role of Gas from a Regulatory Perspective. Technical report, CEER, Brussels.

Dorsey, J. (2019). Waiting for the courts: Effects of policy uncertainty on pollution and investment. *Environmental and Resource Economics*, 74(4):1453–1496.

Dueñas, P., Leung, T., Gil, M., and Reneses, J. (2015). Gas-Electricity Coordination in Competitive Markets Under Renewable Energy Uncertainty. *IEEE Transactions on Power Systems*, 30(1):123–131.

Economist (2021). A court ruling triggers a big change in Germanyś climate policy. *The Economist*. https://www.economist.com/europe/2021/05/08/a-court-ruling-triggers-a-big-change-in-germanys-climate-policy, accessed: 20.08.2021.

Elberg, C. and Hagspiel, S. (2015). Spatial dependencies of wind power and interrelations with spot price dynamics. *European Journal of Operational Research*, 241(1):260–272.

Enervis (2012). Systemwert von Gasspeichern - Intelligenz statt Stahl. Technical report, INES.

ENTSOG (2017). Implementation Document for NC TAR - 2016. Technical Report September, ENTSOG, Brussels.

ENTSOG (2018). Implementation Monitoring of NC - 2017. Technical Report July, ENTSOG, Brussels.

ENTSOG (2019). ENTSOG Transmission Capacity Map 2019. Available at: https://www.entsog.eu/sites/default/files/2020-01/ENTSOG_CAP_2019_A0_1189x841_FULL_401.pdf(Accessed: 29 April 2021).

ENTSOG (2019). Implementation Monitoring of NC - 2017 (2$^{\text{nd}}$ Edition revised). Technical report, ENSTOG, Brussels.

Eser, P., Chokani, N., and Abhari, R. (2019). Impact of Nord Stream 2 and LNG on gas trade and security of supply in the European gas network of 2030. *Applied Energy*, 238:816–830.

European Commission (2010). Commission Staff Working Paper - Interpretative Note on Directive 2009/73/EC Concerning Common Rules for the Internal Market in Natural Gas - Third-Party Access to Storage Facilities. Technical report, European Commission, Brussels.

European Commission (2017). Commission regulation 2017/460. `https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0460&from=EN%29`, accessed: 09.04.2023.

European Commission (2020). Balticconnector gas pipeline up and running since 1 January 2020. *European Commission News*, 8 January 2020. Available at: https://ec.europa.eu/info/news/balticconnector-gas-pipeline-ready-use-1-january-2020-2020-jan-08_en (Accessed: 5 April 2020).

European Commission (2021a). Proposal for a Directive of the European Parliament and of the Council - Amending Directive 2003/87/EC establishing a system for greenhouse gas emission allowance trading within the Union, Decision (EU) 2015/1814 concerning the establishment and operation of a market stability reserve for the Union greenhouse gas emission trading scheme and Regulation (EU) 2015/757. *Official Journal of the European Union.*

European Commission (2021b). Towards competitive and clean european steel. , Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions.

EY and REKK (2018). Quo Vadis EU gas market regulatory framework? - Study on Gas Market Design in Europe. Technical report, European Commission, Brussels.

Fernandez, V. (2018). Price and income elasticity of demand for mineral commodities. *Resources Policy*, 59:160–183. Sustainable management and exploitation of extractive waste: towards a more efficient resource preservation and waste recycling.

Fisher, A. C. (1973). Environmental externalities and the arrow-lind public investment theorem. *The American Economic Review*, 63(4):722–725.

Fürsch, M., Hagspiel, S., JÃ€gemann, C., Nagl, S., Lindenberger, D., and Tröster, E. (2013). The role of grid extensions in a cost-efficient transformation of the European electricity system until 2050. *Applied Energy*, 104:642–652.

German Bundestag (2000). Gesetz fÃŒr den vorrang erneuerbarer energien (erneuerbare-energien-gesetz - eeg) sowie zur änderung des energiewirtschaftsgesetzes und des mineralÃ¶lsteuergesetzes.

http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl100s0305.pdf, accessed: 09.04.2023.

GIE (2018). GIE Storage Map 2018. Available at: https://www.gie.eu/download/maps/2018/GIE_STOR_2018_A0_1189x841_FULL_FINAL.pdf(Accessed: 29 April 2021).

GIE (2019). GIE LNG Map 2019. Available at: https://www.gie.eu/download/maps/2019/GIE_LNG_2019_A0_1189x841_FULL_Final3.pdf(Accessed: 29 April 2021).

Gravelle, H. (1976). The peak load problem with feasible storage. *The Economic Journal*, 86(342):256–277.

Green, R. (2007). Nodal pricing of electricity: how much does it cost to get it wrong? *Journal of Regulatory Economics*, 31(2):125–149.

Grimm, V., Schewe, L., Schmidt, M., and Zöttl, G. (2019). A multilevel model of the European entry-exit gas market. *Mathematical Methods of Operations Research*, 89(2):223–255.

Habermacher, F. and Lehmann, P. (2020). Commitment versus discretion in climate and energy policy. *Environmental and Resource Economics*, 76(1):39–67.

Hallack, M. and Vazquez, M. (2013). European Union regulation of gas transmission services: Challenges in the allocation of network resources through entry/exit schemes. *Utilities Policy*, 25:23–32.

Harstad, B. (2012). Climate contracts: A game of emissions, investments, negotiations, and renegotiations. *The Review of Economic Studies*, 79(4):1527–1557.

Hauser, P., Heidari, S., Weber, C., and Möst, D. (2019). Does Increasing Natural Gas Demand in the Power Sector Pose a Threat of Congestion to the German Gas Grid? A Model-Coupling Approach. *Energies*, 12:1–22.

Hecking, H. (2015). Rethinking Entry-Exit: Two new tariff models to foster competition and security of supply in the EU gas market. *EWI Policy Brief*.

Helm, D., Hepburn, C., and Mash, R. (2003). Credible carbon policy. *Oxford Review of Economic Policy*, 19(3):438–450.

Hepburn, C. (2006). Regulation by prices, quantities, or both: a review of instrument choice. *Oxford review of economic policy*, 22(2):226–247.

Heutel, G. (2019). Prospect theory and energy efficiency. *Journal of Environmental Economics and Management*, 96:236–254.

Höffler, F. (2006). Monopoly prices versus ramsey-boiteux prices: Are they š̈imilar,̈ and: Does it matter? *Journal of Industry, Competition and Trade*, 6(1):27–43.

Höffler, F. (2014). *Ökonomische Analyse des Energieumweltrechts*, volume 2.

IEA (2021). *World Energy Outlook 2021*. IEA. https://www.oecd-ilibrary.org/content/publication/14fcb638-en.

IEA (2022a). Evolution of solar pv module cost by data source, 1970-2020. https://www.iea.org/data-and-statistics/charts/evolution-of-solar-pv-module-cost-by-data-source-1970-2020, accessed: 09.04.2023.

IEA (2022b). VRE share in annual electricity generation in selected countries, 2016-2022.

IEA (2022c). *World Energy Outlook 2022*. Paris.

IPCC (2021). Summary for policymakers. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. YelekÃ§i, R. Yu and B. Zhou (eds.)]. Cambridge University Press. In Press.

Jääskeläinen, J. J., Höysniemi, S., Syri, S., and Tynkkynen, V. (2018). Finland's Dependence on Russian Energy - Mutually Beneficial Trade Relations or an Energy Security Threat? *Sustainability*, 10:1–25.

Jakob, M. and Brunner, S. (2014). Optimal commitment under uncertainty: adjustment rules for climate policy. *Strategic Behavior and the Environment*, 4(3):291–310.

Jeddi, S., Lencz, D., and Wildgrube, T. (2021). Complementing carbon prices with Carbon Contracts for Difference in the presence of risk - When is it beneficial and when not? EWI Working Papers 2021-9, Energiewirtschaftliches Institut an der Universitaet zu Koeln (EWI).

Kanbur, R., Pirttilä, J., and Tuomala, M. (2006). Non-welfarist optimal taxation and behavioural public economics. *Journal of Economic Surveys*, 20(5):849–868.

Kaufman, N. (2014). Why is risk aversion unaccounted for in environmental policy evaluations? *Climatic change*, 125(2):127–135.

Keller, J. T., Kuper, G. H., and Mulder, M. (2019). Mergers of Germany's natural gas market areas: Is transmission capacity booked efficiently? *Utilities Policy*, 56:104–119.

Kies, A., Schyska, B., and von Bremen, L. (2016). Curtailment in a Highly Renewable Power System and Its Effect on Capacity Factors. *Energies*, 9(7):510.

Kristiansen, T. (2007). Cross-border transmission capacity allocation mechanisms in South East Europe. *Energy Policy*, 35:4611–4622.

Laffont, J.-J. and Tirole, J. (1996). Pollution permits and environmental innovation. *Journal of Public Economics*, 62(1-2):127–140.

Le Fevre, C. (2013). Gas storage in Great Britain. *Oxford Institute for Energy Studies, OIES Paper: NG 72*.

Lencz, D. (2023). How curtailment affects the spatial allocation of variable renewable electricity - What are the drivers and welfare effects? EWI Working Papers 2023-2, Energiewirtschaftliches Institut an der Universitaet zu Koeln (EWI).

Lochner, S. (2011a). Identification of congestion and valuation of transport infrastructures in the european natural gas market. *Energy*, 36(5):2483–2492.

Lochner, S. (2011b). Modeling the european natural gas market during the 2009 russian-ukrainian gas conflict: Ex-post simulation and analysis. *Journal of Natural Gas Science and Engineering*, 3(1):341–348.

Lochner, S. (2012). *The Economics of Natural Gas Infrastructure Investments - Theory and Model-Based Analysis for Europe*. PhD thesis, University of Cologne.

López Rodríguez, J. M., Sakhel, A., and Busch, T. (2017). Corporate investments and environmental regulation: The role of regulatory uncertainty, regulation-induced uncertainty, and investment history. *European Management Journal*, 35(1):91–101.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Meunier, G. (2013). Risk aversion and technology mix in an electricity market. *Energy Economics*, 40:866–874.

Mosácula, C., Chaves-Ávila, J. P., and Reneses, J. (2019). Reviewing the design of natural gas network charges considering regulatory principles as guiding criteria in the context of the increasing interrelation of energy carriers. *Energy Policy*, 126:545–557.

Müller, T. (2017). The Role of Demand Side Management for the System Integration of Renewable Energies. In *IEEE*.

Neumann, A. and Zachmann, G. (2009). Expected vs. observed storage usage: Limits to intertemporal arbitrage. In *The Economics of Natural Gas Storage*, pages 13–29. Springer.

Newbery, D. M., Reiner, D. M., and Ritz, R. A. (2019). The political economy of a carbon price floor for power generation. *The Energy Journal*, 40(1).

Nguyen, D. T. (1976). The problems of peak loads and inventories. *The Quarterly Journal of Economics*, 7:242–248.

Nordhaus, W. (2015). Climate Clubs: Overcoming Free-riding in International Climate Policy. *American Economic Review*, 105(4):1339–1370.

Norgaard, R. and Killeen, T. (1980). Expected utility and the truncated normal distribution. *Management Science*, 26(9):901–909.

Obermüller, F. (2017). Build wind capacities at windy locations? Assessment of system optimal wind locations. page 31.

OECD (2002). Glossary of statistical terms - homogenous products.

OECD (2021). *Managing Climate Risks, Facing up to Losses and Damages.*

OEIS (2021). The role of hydrogen in the energy transition. *Oxford Energy Forum*. https://www.oxfordenergy.org/wpcms/wp-content/uploads/2021/05/OEF-127.pdf, accessed: 20.08.2021.

Oren, S., Smith, S., and Wilson, R. (1985). Capacity Pricing. *Econometrica*, 53:545–566.

Pechan, A. (2017). Where do all the windmills go? Influence of the institutional setting on the spatial distribution of renewable energy installation. *Energy Economics*, 65:75–86.

Peŕez-Arriaga, I. J. and Olmos, L. (2005). A plausible congestion management scheme for the internal electricity market of the European Union. *Utilities Policy*, 13:117–134.

Petrov, K., Anton, D., Nußberger, M., and Gonzales, J.-J. D. (2019). Expert Opinion on the Economic Suitability of the REGENT Regulations and the Possibility of Cost-Orientated Reference Price Methods. Technical report, DNV-GL.

Pigou, A. C. (1920). The economics of welfare.

Quiggin, J. C., Karagiannis, G., and Stanton, J. (1993). Crop insurance and crop production: an empirical study of moral hazard and adverse selection. *Australian Journal of Agricultural Economics*, 37(429-2016-29192):95–113.

Redpoint (2012). Gas security of supply report: Further measures modelling. Technical report, Ofgem.

Requate, T. and Unold, W. (2003). Environmental policy incentives to adopt advanced abatement technology: Will the true ranking please stand up? *European Economic Review*, 47(1):125–146.

Richstein, J. C. (2017). Project-based carbon contracts: A way to finance innovative low-carbon investments. *DIW Berlin Discussion Paper.*

Richstein, J. C., Kröger, M., Neuhoff, K., Chiappinelli, O., and Lettow, F. (2021). Carbon Contracts for Difference. *DIW Berlin.*

Rüster, S., von Hirschhausen, C., Marcantonini, C., He, X., Egerer, J., and Glachant, J.-M. (2012). EU Involvement in Electricity and Natural Gas Transmission Grid Tarification. Technical report, European University Institute - THINK Report.

Scharf, H., Arnold, F., and Lencz, D. (2021). Future natural gas consumption in the context of decarbonization - a meta-analysis of scenarios modeling the german energy system. *Energy Strategy Reviews*, 33:100591.

Schlund, D., Schulte, S., Sprenger, T., et al. (2021). The whoś who of a hydrogen market ramp-up: A stakeholder analysis for germany. Technical report, Energiewirtschaftliches Institut an der Universitaet zu Koeln (EWI).

Schmidt, L. and Zinke, J. (2020). One price fits all? Wind power expansion under uniform and nodal pricing in Germany. 20(20):44.

Schulte, S. and Weiser, F. (2019). Natural Gas Transits and Market Power: The Case of Turkey. *The Energy Journal*, 40(2):77–100.

Schweppe, F. C., Caramanis, M. C., Tabors, R. D., and Bohn, R. E. (1988). *Spot Pricing of Electricity*. Springer US, Boston, MA.

Sinden, G. (2007). Characteristics of the UK wind resource: Long-term patterns and relationship to electricity demand. *Energy Policy*, 35(1):112–127.

Sinn, H.-W. (2017). Buffering volatility: A study on the limits of Germany's energy revolution. *European Economic Review*, 99:130–150.

Staffell, I. and Pfenninger, S. (2016). Using bias-corrected reanalysis to simulate current and future wind power output. *Energy*, 114:1224–1239.

Steiner, P. O. (1957). Peak Loads and Efficient Pricing. *The Bell Journal of Economics*, 71:585–610.

Strategy& and PwC (2015). Study supporting the Impact Assessment concerning Rules on Harmonised Transmission Tariff Structures for Gas and Allocation of New Gas Transmission Capacity. Technical report, European Commission, Brussels.

Tietjen, O., Lessmann, K., and Pahle, M. (2020). Hedging and temporal permit issuances in cap-and-trade programs: the market stability reserve under risk aversion. *Resource and Energy Economics*, page 101214.

Unold, W. and Requate, T. (2001). Pollution control by options trading. *Economics Letters*, 73(3):353–358.

Vogl, V., Åhman, M., and Nilsson, L. J. (2018). Assessment of hydrogen direct reduction for fossil-free steelmaking. *Journal of Cleaner Production*, 203:736–745.

Willems, B. and Morbee, J. (2010). Market completeness: How options affect hedging and investments in the electricity sector. *Energy Economics*, 32(4):786–795.

Yasuda, Y., Bird, L., Carlini, E. M., Eriksen, P. B., Estanqueiro, A., Flynn, D., Fraile, D., Gomez Lazaro, E., Martín-Martínez, S., Hayashi, D., Holttinen, H., Lew, D., McCam, J., Menemenlis, N., Miranda, R., Orths, A., Smith, J. C., Taibi, E., and Vrana, T. K. (2022). C-E (curtailment - Energy share) map: An objective and quantitative measure to evaluate wind and solar curtailment. *Renewable and Sustainable Energy Reviews*, 160:112212.

Zerrahn, A., Schill, W.-P., and Kemfert, C. (2018). On the economics of electrical storage for variable renewable energy sources. *European Economic Review*, 108:259–279.