UNIVERSITY OF COPENHAGEN
FACULTY OF SOCIAL SCIENCES

PhD Dissertation

Rasmus Kehlet Berg

Technology and the Environment in General Equilibrium

Supervisor: Peter Birch Sørensen
Submitted: December 31, 2021

Technology and the Environment in General Equilibrium

PhD Dissertation

Rasmus Kehlet Berg

Department of Economics,
University of Copenhagen

Submitted: December 31, 2021

Supervisor: Peter Birch Sørensen

This dissertation is part of the Green REFORM research project financed by Carlsbergfondet and the KR foundation.

Contents

Disclaimer:

I enrolled in the so-called '4+4' PhD programme, which means that I wrote my Master's thesis by the end of my second year in the PhD programme. The research presented in the current dissertation thus, naturally, builds on research ideas from the thesis. The thesis covered a preliminary version of the model for the electricity and district heat system that I use throughout parts 1 and 2 in the dissertation. While the model itself and both parts 1 and 2 have changed significantly since, they may still repeat text from the Master's thesis. The thesis is cited as Berg and Eskildsen (2019) and can be downloaded from the CURIS database.

Foreword

At the tender age of 18, after combing through reports on environmental sustainability, I decided that economists needed a good talking to; it seemed to me that everyone except economists cared or understood that economic growth might be environmentally unsustainable. While the following decade has not been especially comforting for my general concerns over the state of the environment, it has restored a faith in my fellow environmental economists. I am particularly grateful to my supervisor, Peter Birch Sørensen, for creating a platform for environmental economic research at the department of economics by initiating large research projects as the development of the Green REFORM model. I am deeply grateful for his guidance and, beyond that, for growing the team of environmental economics researchers that I have thoroughly enjoyed being a part of.

I am also indebted to Peter Stephensen, the Research Director at the Danish Research Institute for Economic Analysis and Modeling (DREAM). Beyond committing resources and time early on to developing the Green REFORM model, his insights into both economic modeling and energy economics were essential for the research outputs in this dissertation. I also want to thank my collaborators Janek Eskildsen, Jonathan Leisner, and August T. Nielsen, and the many colleagues from DREAM including my brother, Asbjørn K. Berg, and the programming wizard, Martin K. Bonde.

Finally, a huge shout-out goes to my family and friends. To my parents for showing me the beauty in mathematics from a young age with 'equation of the day', and reminding me that math is most beautiful when it can tell you something about society. Ultimately, though, the loudest roar goes out to my girlfriend, Boline K. Skjødt: Somehow, you managed to make the last few years, with 90 hour work weeks and a pandemic, some of the best yet.

English Summary

The PhD dissertation consists of three self-contained parts that deal with representing technology in environmental economic models. All three parts offer ways for improving economic models by leveraging knowledge on technologies collected by other academic disciplines. The first two parts explore new ways for economic models to capture the potential for renewable energy technologies. One of the main challenges of reaching net-zero greenhouse gas emissions is that most cheap, renewable energy generation is intermittent; the availability of wind or solar power depends on natural conditions, such as wind speed and cloud cover. We develop a technology-rich model for electricity and district heat systems and show that it captures the key effects of intermittency. We show that the model can be solved efficiently alongside conventional economic models using smoothing techniques. This integrated framework can be used to evaluate policies e.g. aimed at lowering emissions through electrification of polluting economic activity, such as transportation, using intermittent renewable energy.

The problem of intermittency can be solved using a combination of four technologies: Back-up capacity of more flexible dispatchable electricity producing plants, expansion of the transmission grid to enable more trade in electricity, flexible demand technologies, and electricity storage. The second part of the dissertation uses the technology-rich model from part one to investigate the value and potential of electricity storage. We show that in the Danish context, investing in electricity storage reduces greenhouse gas emissions and lowers average prices significantly. However, the investment costs of lithium-ion batteries prohibit them from being cost-effective until after 2030. We show that this is largely a result of neighboring countries' strategy for investments in storage technologies and intermittent renewables, thus emphasizing the importance of a technology-rich model that includes details on the transmission grid.

Finally, while parts one and two focus on the scope for emission reductions through renewable energy, part three focuses on modeling reductions through abatement technologies. We provide a novel framework that allows us to use much more detailed data on abatement technologies than state of the art environmental economic models. We show under what conditions current best practices over- and underestimates abatement costs when technology data is not appropriately represented in the economic model.

Below I go into greater detail for each part.

1: Electricity and District Heat Systems in a Computable General Equilibrium Framework: A New Approach to Integrated Bottom-up and Top-down Modeling.

Part one presents a new method for integrating bottom-up models of the electricity and district heat systems in economic models. The bottom-up model derives market supply of electricity and district heat from aggregation of heterogeneous energy producing plants, taking technological constraints in the energy system into account. Compared to previous models of electricity and district heat systems that generally rely on linear optimization problems, we use smoothing techniques to

specify the model as a square nonlinear system of equations. Our smoothing approach allows us to derive closed-form solutions for supply and demand decisions, thus reducing the numerical problem to identifying equilibrium prices in electricity and heat markets. Furthermore, to capture the issues related to intermittency in energy production, the model solves for intra-yearly equilibria taking into account hourly variation in generation capacity of intermittent plants, demand for electricity and heat, and capacity of transmission lines.

Finally, we apply the model to the case of Denmark. We calibrate the model to the base year 2019 and show that the model reproduces key characteristics related to intermittency, such as relatively low average prices of wind energy. We simulate the effects of increasing the domestic CO_2-e tax and compare the results from a computable general equilibrium model with and without integrating the bottom-up model. While the two models predict similar aggregate reductions in emissions, they provide different predictions for where these reductions occur. This has important implications e.g. for the estimation of tax revenue generated by the carbon tax.

2: The Value and Potential of Storage

This part deals with one of the main instruments that are expected to mitigate the economic consequences of intermittency and increase the share of renewables. In this part, we develop a convex optimization representation of electricity storage technologies that is particularly easy to solve numerically and nests a traditional linear optimization framework. In a structural estimation using data on Norwegian hydroelectric power plants, we show that our framework is able to replicate observed patters in power generation. Importantly, the estimated parameters suggest that our framework is a significant improvement on the traditional linear optimization approach.

To asses the value and potential of electricity storage, we include our representation of electricity storage in a detailed model for Northwestern European electricity and district heat systems. The model accounts for existing generation technologies, access to trade in electricity, and potential of demand flexibility technologies. We focus on Denmark, which has a high share of capacity based on intermittent energy sources. Compared to previous estimates, we find that electricity storage lowers both average electricity prices and greenhouse gas emissions significantly. This happens even though the considered storage technologies are net consumers of energy. Furthermore, although electricity storage increases the market value of technologies based on intermittent energy sources, electricity storage is, in itself, unlikely to become profitable before 2030 in a Danish context. This is because other strategic substitutes for mitigating intermittency - such as trade in electricity and dispatchable back-up capacity - reduce the scope to profit from large price variations. We emphasize that our conclusions likely depend on the valuation of price stability and emission reductions as well as ambitions for electrification of economic activity.

3: Solving the Big MAC Challenge

The focus of this part is to provide a method for integrating technology data on emissions abatement in economic models. We propose a methodology that provides better estimates of the marginal costs of abating - a key variable for designing effective environmental and climate policy - as it

addresses four issues with existing methods: First, it distinguishes between end-of-pipe and input-displacing technologies and shows that these are substitutes; adopting end-of-pipe technologies reduces emissions related to inputs, thus making dirty technologies cheaper relative to competing clean input-displacing ones. Second, it ties the use of technologies to the relevant use of resources, thus making the costs of technologies a function of general equilibrium prices. For instance, if a large portion of abatement technologies are related to electrification, potential feedback effects from electricity prices are taken into account. Third, it features sluggish technology adoption both across firms and over time. Fourth, we allow technologies to be overlapping, such that adoption of one technology reduces the potential of another. We illustrate the core features through simulations in a toy model.

Finally, we develop general calibration methods that allow us to fit the model to input-output data as well as technology data characteristics, such as the degree of adoption of technologies in a baseline year.

Resumé

PhD afhandlingen består af tre separate dele, der beskæftiger sig med hvordan teknologisk data kan anvendes i miljøøkonomiske modeller. Alle tre dele præsenterer metoder, der kan forbedre økonomiske modeller, ved at udnytte viden om teknologiske løsninger indsamlet i andre akademiske discipliner. De første to dele udforsker nye metoder til at indfange potentialet for vedvarende energiteknologier. En af de primære udfordringer forbundet med at opnå klimaneutralitet er, at billige kilder til vedvarende energi mestendels er af fluktuerende karakter; adgangen til vind- og solenergi afhænger af udsving i vejrforhold som vindhastighed og skydække. Vi udvikler en teknologisk detaljeret model for el- og fjernvarmesystemer og viser, at den fanger de vigtigste karakteristika ved et system med fluktuerende energi. Vi viser, at modellen kan løses sammen med konventionelle økonomiske modeller ved brug af udglatningsteknikker. Denne integrerede model kan bruges til at evaluere effekterne af politiske tiltag, for eksempel rettet mod emissionsreducering gennem elektrificering af forurenende økonomiske aktiviteter, såsom transport, ved brug af fluktuerende vedvarende energi.

Problemet ved at have en høj grad af fluktuerende energi i forsyningssystemet, kan løses med en kombination af fire teknologier: Back-up kapacitet af mere fleksible *dispatchable* energiproducerende værker, udvidelse af transmissionsnettet, der tillader mere handel med el, fleksible efterspørgselsteknologier og lagerteknologier. Afhandlingens andel del anvender modellen fra første del til at undersøge værdien af- og potentialet for ellagringsteknologier. Vi viser - i en dansk kontekst - at investeringer i ellagringsteknologier kan sænke udledningen af drivhusgasser og gennemsnitlige elpriser væsentligt. Dog betyder investeringsomkostningerne for lithium-ion batterier, at de først bliver omkostningseffektive efter 2030. Vi viser, at dette resultat i høj grad er drevet af nabolandenes strategi for investeringer i lagringsteknologier og fluktuerende, vedvarende energi, hvilket understreger vigtigheden af en detaljeret model, der indeholder detaljer om transmissionsnettet.

I forhold til de to første dele, der fokuserer på muligheden for at reducere emissioner gennem vedvarende energi, fokuserer tredje del mere direkte på modellering af emissionsreducerende teknologier. Vi præsenterer en ny modelramme, der kan udnytte langt mere detaljerede data på emissionsreducerende teknologier end *state of the art* miljøøkonomiske modeller. Vi viser under hvilke forhold økonomiske modeller, der ikke udnytter data på emissionsreducerende teknologier på passnde vis, vil over- og undervurdere reduktionsomkostningerne.

Det følgende giver en grundigere gennemgang af afhandlingens tre dele.

1: Electricity and District Heat Systems in a Computable General Equilibrium Framework: A New Approach to Integrated Bottom-up and Top-down Modeling.

Afhandlingens første del præsenterer en ny metode til at integrere *bottom-up* modeller for elektricitets- og fjernvarmesystemer i økonomiske modeller. *Bottom-up* modellen udleder markedsudbud af elektricitet og fjernvarme ved aggregering af heterogene energiproducerende anlæg, under hensyntagen til teknologiske begrænsninger i energisystemet. I modsætning til tidligere modeller for el- og fjernvarmesystemer, der generelt anvender lineære optimeringsmetoder, benytter vi udglatningsteknikker til at specificere modellen som et *square* ikke-lineært ligningssystem. Vores udglatningstilgang tillader os at udlede analytiske løsninger for udbuds- og efterspørgselsbeslutninger, hvilket reducerer det numeriske problem til at identificere ligevægtspriserne i elektricitets- og varmemarkederne. For at tage højde for udfordringerne ved at have en høj grad af fluktuerende energi i forsyningssystemet, løser modellen desuden for intra-årlige ligevægte, under hensyntagen til timevariationen i fluktuerende anlægs' produktionskapacitet, efterspørgsel efter elektricitet og varme, samt transmissionsledningers kapacitet.

Endeligt anvender vi modellen på den danske case. Vi kalibrerer modellen med 2019 som basisår og viser, at modellen indfanger de primære egenskaber ved et forsyningssystem med høj grad af fluktuerende energi, såsom relativt lave gennemsnitspriser på vindenergi. Vi simulerer effekterne af at øge den indenlandske CO_2 afgift og sammenligner resultaterne fra en anvendt generel ligevægtsmodel med og uden integrering af *bottom-up* modellen. De to modeller forudsiger relativt ens reduktioner af aggregerede emissioner, men afviger med hensyn til hvor disse reduktioner finder sted. Dette har vigtige implikationer fx for estimeringen af skatteprovenuet fra CO_2 afgiften.

2: The Value and Potential of Storage

Afhandlingens anden del omhandler en af de primære instrumenter, der forventes at kunne reducere de økonomiske konsekvenser af at have en høj grad af fluktuerende energi i forsyningssystemet. I denne del opstiller vi et konvekst optimeringsproblem for ellagringsteknologier, der er særligt let at løse numerisk, og som nester en mere traditionel lineær optimeringstilgang. I en strukturel estimering, der anvender data på norske vandkraftværker, viser vi, at vores model er i stand til at replikere de observerede mønstre i elproduktionen. De estimerede parametre antyder, at vores metode er en væsentlig forbedring i forhold til den traditionelle lineære optimeringsmetode.

For at vurdere værdien og potentialet af ellagring, inkluderer vi vores repræsentation af ellagring i *bottom-up*, der præsenteres i afhandlingens første del. Modellen tager højde for eksisterende

produktionsteknologier, adgang til handel med elektricitet og anvendelsen af fleksible efterspørgselsteknologier. Vi fokuserer på Danmark, hvor en stor del af elforsyningen er baseret på fluktuerende energikilder. Vi finder, at ellagring både sænker gennemsnitlige elpriser og udledninger af drivhusgasser betydeligt mere end tidligere undersøgelser har fundet. Dette er på trods af, at ellagringsteknologier er nettoforbrugere af energi. Vi finder desuden, at ellagring øger markedsværdien af fluktuerende strøm, men sandsynligvis ikke er privatøkonomisk profitabelt før 2030 i en dansk kontekst. Det skyldes at andre strategiske substitutter for afbødning af effekterne ved fluktuerende energi – såsom handel med elektricitet og *dispatchable* back-up kapacitet – reducerer muligheden for at profitere på store prisudsving. Vi understreger, at vores konklusioner formentlig afhænger af hvordan prisstabilitet og emissionsreduktioner samfundsøkonomisk værdisættes, såvel som af hvor høje ambitioner, der sættes for elektrificering af forurenende økonomiske aktiviteter såsom transport.

3: Solving the Big MAC Challenge

Afhandlingens tredje del fokuserer på udvikle en metode til at integrere teknologiske data for emissionsreducerende tiltag i økonomiske modeller. Vi præsenterer en metode, der giver bedre estimater af de marginale reduktionsomkostninger for emissioner, idet den afhjælper fire problemer ved eksisterende metoder: 1) Vi skelner mellem *end-of-pipe* og *input-displacing* teknologier og viser, at disse er substitutter; anvendelse af *end-of-pipe* teknologier reducerer inputrelaterede emissioner og gør dermed forholdsvist beskidte teknologier billigere i forhold til konkurrerende rene *input-displacing* teknologier. 2) Vi kobler brugen af emissionsreducerende teknologier til den relevante brug af ressourcer og gør dermed teknologiomkostningerne til en funktion af priser i den generelle ligevægt. Hvis reduktionsteknologierne for eksempel er baseret på strategien om elektrificering, tages der højde for potentielle feedback-effekter fra elpriserne. 3) Vores metode giver naturligt anledning til træg indtrænging af nye teknologier, både på tværs af virksomheder og over tid. 4) Vi tillader, at teknoliger har overlappende reduktionspotentiale, således at investering i én teknologi kan reducere reduktionspotentialet for en anden. Vi illustrerer disse kerneegenskaber gennem simuleringer i en legetøjsmodel.

Endeligt udvikler vi en generel kalibreringsmetode, der tillader os at tilpasse modellen til både input-output data og teknologiske data, såsom graden af indtrængning for en teknologi i et givet basisår.

Part 1

Electricity and District Heat Systems in a Computable General Equilibrium Framework: A New Approach to Integrated Bottom-up and Top-down Modeling.

Disclaimer: *The current part extends the work from my Master's thesis. Some text may be repeated from Berg and Eskildsen (2019).*

Electricity and District Heat systems in a Computable General Equilibrium Framework*

A new approach to integrated bottom-up and top-down modelling

Rasmus K. Berg† Janek B. Eskildsen‡

December 29, 2021

Abstract

Expansion of intermittent, renewable energy is a key strategy for decarbonizing economic activity. To estimate the economic consequences of higher shares of intermittent energy supply, we develop a dynamic bottom-up model for electricity and district heat systems that can be directly integrated with large-scale general equilibrium models. The bottom-up model derives market supply of electricity and district heat from aggregation of heterogeneous energy producing plants, taking technological constraints in the energy system into account. Compared to previous models of electricity and district heat systems that generally rely on linear optimization problems, we use smoothing techniques to specify the model as a square nonlinear system of equations. Our smoothing approach allows us to derive closed-form solutions for supply and demand decisions, thus reducing the numerical problem to identifying equilibrium prices in electricity and heat markets. Furthermore, to capture the issues related to intermittency in energy production, the model solves for intra-yearly equilibria taking into account hourly variation in generation capacity of intermittent plants, demand for electricity and heat, and capacity of transmission lines.

Finally, we apply the model to the case of Denmark. We calibrate the model to the base year 2019 and show that the model reproduces key characteristics related to intermittency, such as relatively low average prices of wind energy. We simulate the effects of increasing the domestic CO_2-e tax and compare the results from a computable general equilibrium model with and without integrating the bottom-up model. While the two models predict similar aggregate reductions in emissions, they provide different predictions for where these reductions occur. This has important implications e.g. for the estimation of tax revenue generated by the carbon tax.

*For useful comments and discussions, we thank Peter Stephensen, Peter Birch Sørensen, Kristoffer Steen Andersen, and seminar participants at EAERE-ETH Winter School, University of Copenhagen, and University of Oslo. We are grateful to the modelling group DREAM for developing and giving us access to the CGE model GreenREFORM.

†University of Copenhagen. Oster Farimagsgade 5, 1353 Copenhagen K, Denmark. E-mail: rasmus.kehlet.berg@econ.ku.dk.

‡University of Copenhagen. Oster Farimagsgade 5, 1353 Copenhagen K, Denmark. E-mail: jbe@econ.ku.dk

Keywords: Computable general equilibrium; Bottom-up; Top-down; Intermittency; Environment and trade; Demand flexibility.

JEL classification codes: C2; C6; D5; F1; Q4.

1 Introduction

In addressing the problem of anthropogenic climate change, economies are facing large structural changes in the coming decades as society transitions to a future of net-zero greenhouse gas emissions. A key component of the green transition is the provision of renewable electricity and heat. Electricity and heat generation is the largest driver of global CO_2 emissions, accounting for roughly 44% of CO_2 emissions in 2019 (IEA, 2021a). The importance of renewable electricity provision is compounded by the fact that electrification is essential for decarbonizing other key drivers of CO_2 emissions such as the transportation sector (see e.g. Zhang and Fujimori, 2020). However, the process of electrifying economic activity requires a massive expansion of capacity based on intermittent renewable energy sources such as wind, water, and sunlight (Jacobson et al., 2017; IEA, 2021b).

From a partial equilibrium perspective, a large scale penetration of intermittent renewables raises fundamental concerns for the energy system: Given the technological constraints in the energy system, how can society safeguard energy supply if most of it has to come from intermittent sources, and what are the costs and potential of various technologies aimed at mitigating the issue of intermittency? Answering such questions requires a model that accounts for the problem of intermittency, but also captures key characteristics of potential solutions, such as storage technologies, expansion of transmission networks, and flexible demand technologies. Considering the need for electrifying other sectors, however, we also require a model that accounts for general equilibrium effects and that can estimate the economy-wide effects of integrating a much higher share of intermittent renewable energy in the energy system.

To provide answers to the above mentioned concerns, we develop a so-called 'bottom-up' model for the electricity and district heat system that can be directly embodied in large-scale general equilibrium models. At its core, the bottom-up model derives market supply for electricity and district heat at the hourly level by aggregating dispatch decisions from a discrete set of heterogeneous energy producing plants. Our framework can be used to address a wide range of climate, energy, and economic policies within a unified conceptual framework based on the concept of a decentralized equilibrium with profit maximizing agents.

The primary contribution of this paper is to advance the state of the art for modeling and solving integrated general equilibrium models with technologically detailed bottom-up models for electricity and district heat systems. Generally, electricity and district heat systems include a wide array of technological constraints represented by weak inequalities. These constraints primarily come in the form of capacity constraints. For instance, electricity and district heat generation is

bounded by a maximum generation capacity, storage technologies are subject to storage capacity constraints, and trade in electricity with neighboring countries is constrained by capacity limits for transmission lines. The presence of these occasionally binding constraints implies that deriving aggregate market supply in bottom-up models involves searching through potential corner solutions where capacity constraints are binding. We propose to replace traditional assumptions of linear costs with strictly increasing and convex marginal costs that guarantee an interior solution and - in all instances except for storage technologies - provide closed-form solutions. Doing so implies that the only numerical challenge consists of identifying equilibrium prices where aggregate market supply equals demand. We show that solving the model boils down to finding the solution to a square system of nonlinear equations.

Specifically, our method relies on recasting the profit maximization problem of plants as a smooth and convex optimization problem. We do this for a number of different technologies including simple dispatchable generation plants (e.g. condensation plants) and combined heat and power plants. In each of these cases, there is a closed-form solution to the generation decision at the plant level in the form of a smooth and increasing function of equilibrium prices. In a companion paper (Berg and Eskildsen, 2021), we also show how to recast the dynamic profit maximization problem of plants with capacity for energy storage such as hydro power plants with reservoirs, batteries, and heat storage. While there is no analytical solution to this problem, we show that the optimality conditions can be topologically ordered, which makes it straightforward to solve numerically. Furthermore, we model trade in electricity as a homogeneous good where trade is constrained by the physical limits of transmission lines. In a similar exercise as with individual plants, we show that assuming a specific set of convex marginal costs of trade, the total net exports in electricity between any two neighboring areas can be expressed in closed-form as a smoothly decreasing function of the price difference between the areas.

Our main theoretical result can be summarized as follows: The aggregate market excess demand function is a smooth function of equilibrium prices and the short run equilibria of the bottom-up model is the solution to a square system of equations based on smooth differentiable functions; thus, equilibria are in the interior and can be computed using gradient-based methods. Furthermore, we show that the short run equilibrium is linearly homogeneous in the sense that increasing (i) all generation capacities, (ii) yearly levels of electricity and heat demand, and (iii) trade capacities by $\lambda\% > 0$ in all states of the model, markets, and geographical areas leaves equilibrium prices unchanged, the share of electricity demand covered by net exports is unchanged, and all plants increase production by $\lambda\%$. Finally, when smoothing parameters converge to zero, the solution to our bottom-up model collapses to the solution of a traditional bottom-up linear optimization framework.[1] We also show under which conditions the smoothing methodology is not appropriate. In our framework, it can be difficult for gradient-based methods to identify equilibrium prices when three conditions are met simultaneously: (i) Demand is perfectly inelastic, (ii) marginal costs of generation and transmission are approximately constant, and (iii) there are large intervals in the

[1] For practical applications, however, gradient based solvers are only applicable for sufficiently positive smoothing parameters.

distribution of average generation costs that have zero density.

Our model for the electricity sector and the district heat system draws on existing bottom-up models such as Ramses (Danish Energy Agency, 2018), Balmorel (EA Energy Analyses, 2018), and TIMES (Loulou et al., 2016). We also rely on the extensive body of data used to inform these models. To fully account for intermittency, the bottom-up model includes hourly variation in generation capacity of intermittent plants, demand for electricity and heat, and capacity of transmission lines. We take into account the most important instruments aimed at mitigating the issue of intermittency: Trade in electricity, energy storage, flexible demand technologies, and dispatchable back-up capacity. For the first three instruments, we provide novel formulations and features. First, we estimate the hourly availability of transmission lines using a double truncated Tobit type I model with two latent variables. The estimation allows us to simulate patterns in availability of transmission lines that feature realistic correlations with peaks in demand and available intermittent wind power. Second, we introduce an intuitive and simple way of modeling flexible short run demand for electricity and heat. Our formulation encompasses both habits as well as partial price-responsiveness. While flexibility in short run energy demand has been studied theoretically (see e.g. Ambec and Crampes, 2017), our approach allows us to match the current pattern of short run demand, while enabling the researcher to simulate effects of increased demand flexibility (e.g. by the use of smart meters and load switching devices). Third, storage plants optimize discounted profits under convex and significantly nonlinear costs based on smoothing parameters that are structurally estimated. While the model can be solved with the full 8760 hourly states for each year, we mainly use an aggregated version with 25 representative states. For details on storage modeling and clustering of intra-yearly states, we refer to the companion paper Berg and Eskildsen (2021).

As an illustrative example, we apply the bottom-up model to the case of the Denmark. The techno-economic parameters of the model are based on the national inventory of domestic power and heat producing plants (known as 'Energiproducenttællingen') and aggregated data for neighboring countries collected by the Danish Energy Agency. We calibrate the model to annual moments in the Danish energy statistics for 2019 and show that the model is able to replicate other relevant characteristics of the energy system, such as the average prices received by intermittent technologies. We also document the nonlinearity of the aggregate supply curve, which is a direct implication of the heterogeneity in generation capacity and marginal costs. The nonlinearity is crucial because the marginal effects generated from such a supply function are significantly different from what one would see in conventional CGE models. The bottom-up model is then integrated with a large-scale, fully dynamic, multi-sector CGE model of the Danish Economy called GreenREFORM. We simulate the effects of gradually increasing the domestic CO_2-e tax on greenhouse gas emissions not covered by the EU Emissions Trading System (ETS). To illustrate the importance of the bottom-up model, we compare the results from simulations with the integrated model and a 'stand-alone' version of GreenREFORM without the bottom-up model. While the results are

similar in the aggregate, we document non-trivial differences in where reductions in greenhouse gas emissions occur. These differences take place in the electricity and district heat sectors represented by the bottom-up model and have important implications e.g. for prices, consumption, net exports in the case of electricity, and tax revenue. Importantly, the 'stand-alone CGE' model underestimates the effect of additional tax revenue by roughly 1.2 billion DKK in 2030. Finally, the integrated model allows us to investigate how the change in Danish CO_2-e tax affects allowance prices in the EU ETS. We show that the Danish CO_2-e tax leads to roughly 5 Mt accumulated CO_2-e reductions by 2040 in domestic emissions from activity covered by the EU ETS. With lower demand for emission allowances the related price decreases, which in turn leads to increased foreign emissions of 4.8 MtCO_2-e that primarily comes from increased consumption of coal by foreign electricity generating plants.

The rest of the paper is organized as follows: Section 1.1 outlines existing literature and how it relates to this paper. Section 2 provides an overview of the electricity and district heat sector in Denmark as well as the plant level bottom-up data used by the model. Section 3 outlines the fundamentals of the CGE model GreenREFORM. Section 4 describes the bottom up model for the Danish electricity and district heat systems: 4.1 presents the concept of smoothing, 4.2 details the optimal dispatch of various types of plants, 4.3 outlines how we model trade in electricity including the estimation of availability in transmission lines, 4.4 elaborates on short run demand for electricity and district heat, and finally 4.5 summarizes the equilibrium concept applied in the bottom-up model and its main features. Section 5 calibrates the bottom-up model to replicate aggregate data in a baseline year, and provides some evidence for how the model behaves in a baseline scenario. Section 6 shows the effects of increasing the tax on CO_2 emissions in the CGE model with and without the bottom-up model. Finally, section 7 concludes.

1.1 Related literature

The paper relates to three strands of existing literature: (i) Bottom-up models of the energy system, (ii) integration of bottom-up models with general equilibrium models, and (iii) intermittency in energy production. We review the different strands of literature below.

Energy and power system models: The scope of the bottom-up model

A large number of bottom-up models are dedicated to describing the energy system, including the TIMES-MARKAL model family (Loulou et al., 2016), EnergyPLAN (Lund et al., 2021), and PRIMES (Siskos et al., 2018) to name a few. All of these energy system models cover the entire energy system from the extraction of primary energy to the final consumption of energy services. Compared to such large-scale energy system models, the bottom-up model presented in this paper only deals with electricity and district heat systems. In the full integrated model, however, the CGE model, GreenREFORM, still encompasses the entire Danish energy system: Sectors in the CGE model demand energy-services, which are translated into a demand for secondary energy

goods as electricity and district heat through the application of end-use technologies.[2]

In its description of the electricity system, we simplify the bottom-up model in a number of ways for computational purposes. While we include trade in electricity across price regions using a transmission network, we do not include a representation of the far more high-dimensional distribution network within price regions. Furthermore, we do not model a balancing market for electricity, nor do we generally consider the need for ancillary services. Also, we disregard the issue of market power in the electricity markets and assume simple competitive markets. In all of these respects, our simplifying assumptions are in line with those adopted by other bottom-up models that are linked to general equilibrium models (see e.g. Balyk et al., 2019; Krook-Riekkola et al., 2017).

Integrating bottom-up and general equilibrium models

Integrating bottom-up and general equilibrium models has been a long-standing issue in energy economics. The immediate challenge is that bottom-up and general equilibrium models require fundamentally different solution methods. Bottom-up models generally minimize the total system costs of satisfying a given energy demand, where supply derives from heterogeneous plants with capacity constraints and linear costs of generation. Therefore, nearly all bottom-up models of the energy system are formulated as (mixed integer) linear programming optimization problems (Bloess et al., 2018). They are most efficiently solved using gradient-free methods, e.g. the dual simplex method. In contrast, CGE models are formulated as a square system of smooth nonlinear equations and solved using gradient-based methods (see Horridge et al., 2013, for a review).

Böhringer and Rutherford (2009) classify previous efforts to couple detailed bottom-up models for the energy system with general equilibrium models into three categories: (i) Having a reduced form representation of one model in the other, (ii) soft-linking existing models and iterating between the two models until convergence, and finally (iii) direct coupling by formulating the economic model as a mixed complementarity problem (MCP). In this paper we add a fourth category to this classification: Direct coupling by specifying the bottom-up model as a square nonlinear system of equations using smoothing techniques.

Fully integrated models can solve the bottom-up and general equilibrium models simultaneously and were pioneered by Rutherford (1995), Böhringer (1998), and Böhringer and Rutherford (2008). Mathiesen (1985) showed that a competitive equilibrium model can formulated as a MCP in which profits and excess supply are non-negative and associated with non-negative prices and output, respectively. Böhringer and Rutherford (2009) used this result to show that a standard bottom-up model for the electricity sector formulated as a MCP is a special case of the competitive equilibrium model also formulated as a MCP. The MCP format therefore allows for a direct integration of the two models. The generality of this approach, however, is also the reason why it suffers from the curse of dimensionality. For each inequality constraint, the MCP defines a non-negative complementary variable to represent the shadow cost of the constraint, often doubling the size of the model (Böhringer and Rutherford, 2008). For this reason, the MCP has limited

[2]For more on the GreenREFORM model, see https://dreamgroup.dk/greenreform/.

practical appeal, in particular if the size of the bottom-up model is large and thus requires a large number of inequality constraints.

Large-scale applications have therefore focused on soft-linking models where the two models are solved separately. Most applications use a simple successive approximation approach where a set of connection points are initially identified: A connection point is a variable that is exogenous in one model and endogenous in the other (Krook-Riekkola et al., 2017). The models are then solved iteratively where connection points are updated for each iteration until convergence in the connection points is attained. Böhringer and Rutherford (2009) improve on the simple successive approximation approach by including a linear approximation of the endogenous connection point in the CGE model directly in the bottom-up model.[3] Thus, instead of only exchanging information on the level of relevant connection points in each iteration, they utilize information on the first-order derivatives as well. Compared to this, our integration approach using smoothing techniques allow us to use additional information such as second-order derivatives when linking. This is done automatically when solving the models simultaneously using off-the-shelf solution software such as CONOPT (Drud, 2019).

The main advantage of the approach presented in this paper, and the overall purpose, is the full information exchange between the bottom-up and general equilibrium model when computing equilibria in the electricity and district heat markets. All the information is implicitly exchanged in the formulation of the integrated model's Jacobian and Hessian. However, we emphasize that solving the bottom-up model as a system of nonlinear equations is not always more appropriate. The only numerical challenge in the smooth bottom-up model is to identify equilibrium prices such that excess demand is zero. However, if the excess demand function is flat around large price intervals, a gradient-based method can have problems with achieving convergence. In relation to this, the convergence properties, and more general the solution time, for gradient-based methods for solving nonlinear systems depends on initial values of the system. We suggest that, if necessary, smoothing parameters can be used as a globalization method; the model can easily be solved for large smoothing values and then gradually decreased towards a preferred level.[4]

The effect of intermittent renewables in energy production

The costs and complications of intermittent electricity generation can be described in both technical and economic terms. In technical terms, increasing the share of intermittent electricity generation makes it difficult to safeguard electricity supply. Power transmission grids operate at a certain frequency (typically 50 Hertz) and for the grid to be stable, supply and demand has to be balanced at all times. With an increasing share of fluctuating supply, however, it becomes more difficult to balance load and supply (Delarue and Morris, 2015). In economic terms, intermittency creates a more rigid energy supply giving rise to price variations due to fluctuating natural conditions and technological constraints. Ideally, we would like to move energy from states where energy is valued

[3]The endogenous connection is the Marshallian energy demand around the equilibrium price.

[4]In computational economics a globalization procedure is a strategy for reaching convergence when the initial starting point is not in the domain of attraction ('not near' the solution).

at a low price to states where it is valued at a high price (Ambec and Crampes, 2017). Our focus is on the economic aspect of intermittency.

To mitigate intermittency, there are generally four instruments that can be applied: First, investment in dispatchable back-up capacity that can be used when intermittent productivity is low (Hittinger and Lueken, 2015; Delarue and Morris, 2015). Second, investment in energy storage implies that supply can be moved to hours where it is needed the most (Bistline et al., 2020; Zerrahn et al., 2018). Third, trade in electricity allows countries to access a more diverse portfolio of energy producing technologies if intermittent productivity is low (Brown et al., 2018; Bahar and Sauvage, 2013). Fourth and finally, similar to making supply more flexible, demand can be made flexible by means of smart meters or load switching devices (Ambec and Crampes, 2017).[5] The present model includes all these potential channels for mitigating intermittency.

Hirth et al. (2015) argue that the cost of intermittency is summarized by the downlift defined as the price difference between electricity provided by intermittent technologies and the average yearly price. Because intermittent technologies have near zero marginal cost electricity prices are driven down when intermittent productivity is high. This implies that the average price intermittent technologies receive is lower than the average price that dispatchable plants receive. In absence of subsidies or mandated feed-in tariffs there is less incentive to undertake further investments in intermittent renewable-based technologies. In practice, the downlift for intermittent energy supply is caused by a combination of technological constraints and imperfect foresight of fluctuating weather conditions. We show that our model is able to replicate the observed downlift even though we assume that plant managers have perfect foresight, as in most large scale models of the energy system. We leave the case of uncertainty for future research.

2 Institutional setting and data

In the current section, we provide an overview of the data used in the bottom-up model, and an overview of the market structure of the Danish electricity and district heat systems including some descriptive evidence on the current state of the systems.

2.1 Data

The primary source is the techno-economic data collected and maintained by the Danish Energy Agency that is used to inform their model for the electricity and heat systems, Ramses. This data contains geographical, technical, and economic information on all electricity and district heat producing plants coupled to the public grid in Denmark. This plant-specific data is based on the national inventory of domestic energy producing plants (known as *Energiproducenttællingen*). Beyond this, the database contains more aggregated data on electricity producing plants in relevant

[5] Another instrument is demand curtailment/load shedding. However, it is more of a sign of intermittency than an instrument. It is the process of curtailing demand if demand is higher than supply for any given price. This instrument is typically a last resort because it involves consumers having a zero consumption. The option for curtailing demand is also represented in the model.

neighboring countries. Plants are described by a finite, discrete, seven-dimensional state space. This information includes which electricity market the plant competes on (g_E), district heat market (g_H), type of fuel mix (f_F), production technology (τ), type of intra-yearly variation in generation capacity (ν), subsidy type (s), and the bid type (b). Furthermore, plants are described by numerous continuous techno-economic variables such as fuel efficiency, generation capacity, variable operating and maintenance costs, non-fuel related emission intensities etc.. Section 4 shows how we utilize this data.

Beyond covering the current state of electricity and district heat producing plants, data covers the future composition of plants as projected by the Danish Energy Agency under a frozen policy assumption. We note that the projection of foreign electricity producing plants are based on the scenario "National Trends" from the Ten-Year Network Development Plan 2020, which is based on National Energy and Climate plans. This means that the baseline scenario for Danish plants does not live up to the political target of 70% reductions in greenhouse gas emissions by 2030, whereas foreign countries live up to their respective national targets. As investments in new generation capacity is highly influenced by policy, we make the simplifying assumption that the entire path of capacity is fixed; we compute the value of investments in additional generation capacity for various types of technologies, but leave the challenge of endogenizing investments under significant political constraints for future research.

We use three additional data sources: First, we use the yearly energy statistics (Danish Energy Agency, 2021) to calibrate the bottom-up model to a baseline year 2019. This includes information on production levels of electricity and district heat, yearly aggregate consumption of primary energy, net exports of electricity, levels of consumption etc.. Second, we use hourly data on consumption and prices of electricity from the Danish Transmission System Operator (TSO) Energinet (Energy Data Service, 2018; Energy Data Service, 2021a). This is applied to estimate habits in consumption and to compute weighted average consumer prices on electricity. Third, we use hourly data on net transfer capacities from ENTSO-E to estimate intra-yearly variation in transmission capacities; section 4.3 elaborates on this.

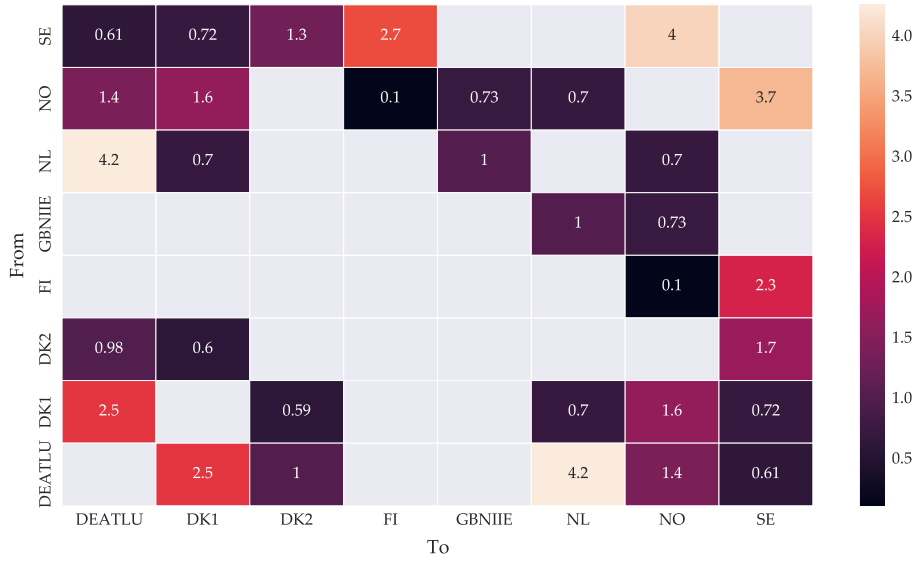
2.2 The electricity system

Electricity markets are generally divided into a wholesale market, a distribution market, and a retail market for end-consumers. The primary wholesale market in Denmark is the Nordic electricity exchange Nord Pool where most Northern European countries operate. In this market, trade in electricity with neighboring partners is determined and as of 2016 more than 90% of the Danish electricity consumption is traded on Nord Pool (Energy Commission, 2016). As a decentralized market with many producers the wholesale market exhibits (near) perfect competition. The equilibrium concept in our framework resembles that of the wholesale market.

The Danish electricity system is comprised of two interconnected areas, Western Denmark (DK1) and Eastern Denmark (DK2). The two areas are separated by the Great Belt. As of 2021 DK1 is further connected with Germany, the Netherlands, Norway, and Sweden while DK2

is connected with Germany and Sweden. Furthermore, significant expansions of the European transmission grid are currently underway, including a connection between DK1 and the United Kingdom (ENTSO-E, 2021). Between 2019 and 2025, the total access to imports and exports for DK1 more than doubles, while for DK2 it increases by roughly 50%. Figure 2.1 provides an overview of the capacities in existing transmission lines in 2021 for the electricity areas included in our model. Appendix A provides supplementary descriptive evidence.

Figure 2.1: Hourly Transmission Capacities in GW, 2021

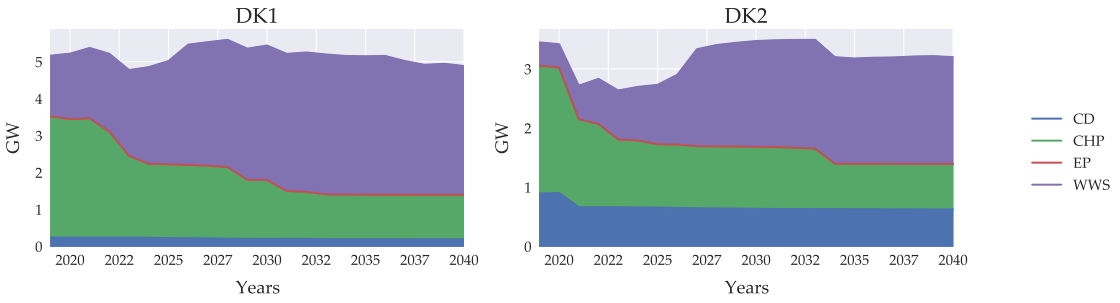


Note: The abbreviations denote the geographic areas Western Denmark (DK1), Eastern Denmark (DK2), Germany, Austria, and Luxembourg (DEATLU), Finland (FI), Great Britain, Northern Ireland, and Ireland (GBNIIE), Netherlands (NL), Norway (NO), and Sweden (SE).

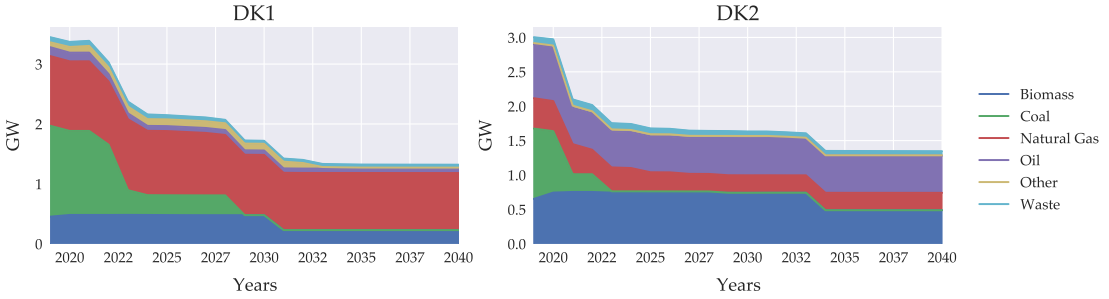
In 2021, the electricity generation capacity of Danish plants comes mostly from wind-turbines as well as combined heat and power plants. As figure 2.2 illustrates, however, this mix of plant types is expected to skew much more heavily towards wind and photovoltaics towards 2040. As more wind and photovoltaics enter the system, a large block of the dispatchable generation capacity is retired: As figure 2.2 shows, dispatchable plants that rely on coal are almost entirely phased out, while a share of biomass plants are also retired throughout the 2030s.

Even though the share of production that comes from intermittent renewables, and in particular wind, has increased during the last decades, the price volatility has for the most part been kept down. This is primarily due to the access to imports from Norway and Sweden that have large hydro power plants with reservoirs. Figure 2.3 shows that prices between 2011-2020 generally fluctuated between -500 and 2000 DKK/MWh with almost all of the distribution being centered around prices of 0-500 DKK/MWh. As we will revisit later in section 5, it is important for the bottom-up model to replicate this price variation to ensure that the model accurately can assess the value and potential of intermittent renewables.

Figure 2.2: Hourly Average Generation Capacity, Electricity



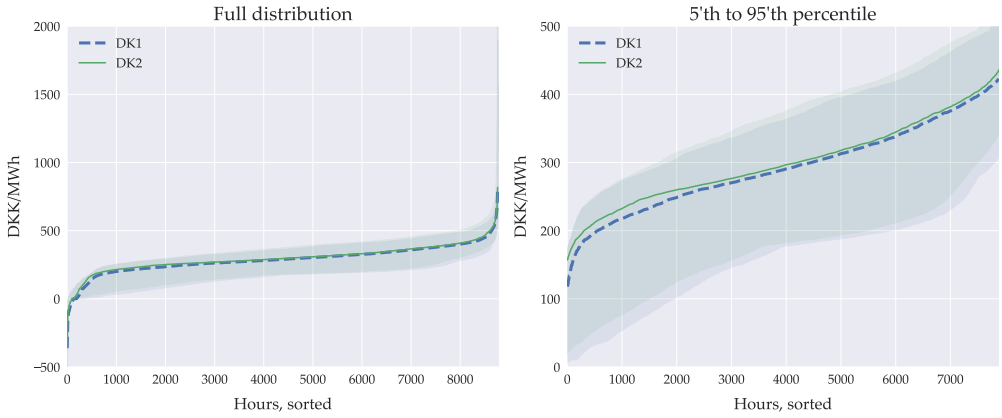
(a) Plant types



(b) Fuel types

[1] By hourly average generation capacity, we mean that capacities are scaled by their yearly capacity factors. I.e. a wind-turbine with installed capacity of 1GW has an average generation capacity equal to its capacity factor.
[2] The abbreviations for plant types denote condensation plants (CD), Combined Heat and Power plants (CHP), Exogenous Production (EP), and Water, Wind, and Sunlight (WWS) technologies.
[3] The fuel type "natural gas" includes bio methane/synthetic natural gas.. The category "Other" includes bio gas, bio oil, hydrogen, and uranium.

Figure 2.3: Hourly variation in prices on day ahead market, 2011-2020



Note: The graphs shows the sorted day ahead prices for DK1 and DK2 for 2019. The shaded areas indicate the minimum/maximum interval of prices between 2011-2020.
Source: Energy Data Service (2021a).

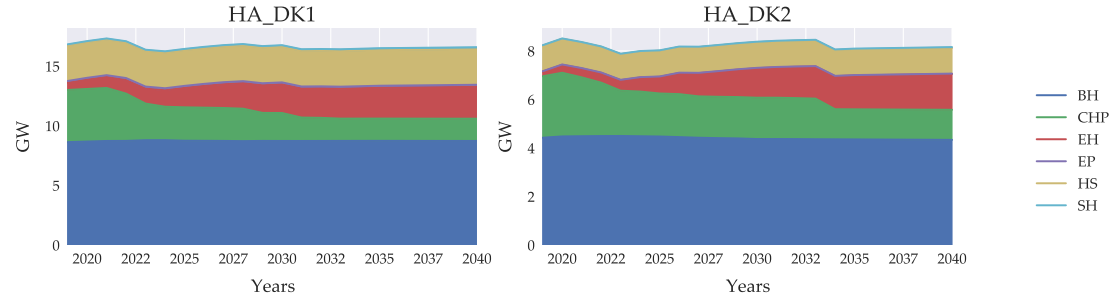
2.3 The district heat system

The Danish heat system differs from the electricity market in three major ways. Firstly, the Danish heat system is comprised of two supply systems: District- and individual heating. The district heat

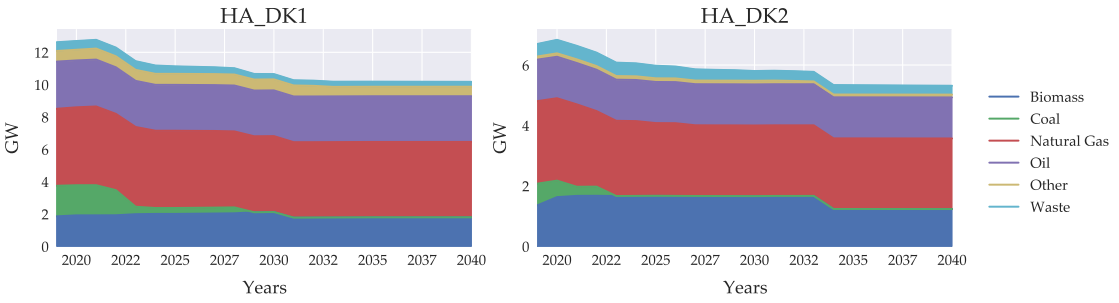
system currently covers around 2/3 of Danish households' consumption and half of all consumption of heating services (Grøn Energi & Ea Energianalyse, 2016). In the current version of the model individual heat supply is not included due to data limitations. Secondly, the market for district heat is not characterized by perfect competition. Thus there is no spot price clearing the market at each hour. Instead, prices are determined by cost-plus pricing, i.e. on the basis of the costs of the plants. Third and finally, district heat is produced by local plants distributed within district heat areas and is (generally) not traded across domestic districts or countries. Consequently, a wholesale market does not exist for heating services.

Figure 2.4 shows that, compared to the generation capacity for electricity, we expect less significant changes to the district heat supply: The overall generation capacity stays more or less constant. As we saw in figure 2.2 as well, part of the capacity from combined heat and power plants is retired. In the district heat system, this is expected to be replaced by electrical heaters and heat pumps. However, there is not the same expected increase of intermittent renewables as in the electricity system.

Figure 2.4: Hourly Average Generation Capacity, District Heat



(a) Plant types



(b) Fuel types

[1] By hourly average generation capacity, we mean that capacities are scaled by their yearly capacity factors. I.e. a solar heating plant with installed capacity of 1MW has an average generation capacity equal to its capacity factor.
[2] The abbreviations for plant types denote Boiler Heaters (BH), Combined Heat and Power plants (CHP), Exogenous Production (EP), Electrical Heaters, including heat pumps (EH), and Water, Wind, and Sunlight (WWS) technologies. For fuel types, the category "other" covers bio gas, bio oil, and hydrogen.

3 The computable general equilibrium model

The conventional computable general equilibrium model typically consists of a number of agents that either maximize profits or utility, or simply behave in accordance with some rules of thumb. When aggregated across all agents, the demand and supply for all goods in the economy are balanced in a way that in some baseline years corresponds to the data from national accounts.[6]

In the current paper we link our bottom up model of the electricity and district heat system to the GreenREFORM model, a dynamic, multisector model of the Danish economy. The specifics of the CGE model is not essential for the methodological contribution of the paper. Thus, for our purposes, it suffices to briefly outline a very general model.

The economy consists of N_G goods that are supplied and/or demanded by N_A aggregate agents. Let $(\boldsymbol{x}_t^d, \boldsymbol{x}_t^s)$ denote vectors of goods demanded and supplied in year t, and let $(x_{a,t}^d, x_{a,t}^s)$ denote the elements specific to agent a. Similarly, $p_{x,t}^a$ denotes the price agent a pays for good x in year t, $\boldsymbol{p}_t^a$ denotes the vector of of prices on the N_G goods, and $\boldsymbol{P}_t$ denotes the matrix of all prices. Importantly, note that from the perspective of integrating a detailed model of the electricity and district heat systems, it is essential that the CGE model has such a counterpart, as well as a representation of foreign suppliers/consumers of electricity. Furthermore, it is essential that the CGE model operates with commodities such as electricity and heat, various fuel inputs, and capital goods.

For each year t the economy is in general equilibrium as long as:[7]

$$\forall x \in \boldsymbol{x}: \qquad \sum_a x_{a,t}^d\left(\boldsymbol{P}_t\right) = \sum_a x_{a,t}^s\left(\boldsymbol{P}_t\right).$$

When a CGE model is formulated as a square system of nonlinear equations, the implication is that all demand and supply components $(\boldsymbol{x}_t^d, \boldsymbol{x}_t^s)$ are continuously, differentiable functions. Most often, the demand functions $(x_{a,t}^d)$ will be decreasing in the respective price $(p_{x,t}^a)$, and supply functions $(x_{a,t}^s)$ will be increasing in all relevant prices $(p_{x,t}^a)$.

The role of the bottom-up model is to formulate demand and supply functions for the electricity and district heat sector. The models are roughly linked as follows: The CGE model takes yearly electricity and district heat prices as given as well as the electricity and district heat sector's demand for inputs, such as fuels, capital, labor etc..[8] These are the CGE connection points. Given this, the CGE model essentially solves for all other equilibrium prices, levels of production, and income streams. For the bottom-up model, the relevant connection points include the level of demand for electricity and district heat and equilibrium prices on fuels, labor, capital etc.. Given the bottom-up connection points, the bottom-up model solves for the CGE connection points. The

[6] We note that balancing demand and supply is in general less restrictive than it first sounds; For instance, modelling inventory investments allow for temporary disequilibria, and including matching frictions on the labor market allows for involuntary unemployment.

[7] Naturally, as the model is dynamic, some of the decision variables will inherently also depend on future prices. We ignore this here for notational simplicity.

[8] In principle, there might be more than one sector that produces electricity and district heat. In the case of the GreenREFORM, the sector is split into a waste and a non-waste sector. For more specific details on the linking between the CGE model and the bottom up model see Berg, Kirk, et al. (2020).

integrated model is solved when the outlined connection points have converged. As discussed in the introduction, this convergence can be achieved by iterating between the models or simply by solving the two simultaneously.

4 The short run equilibrium for electricity and district heat

In the following, we distinguish between types of outputs, individual plants, year, hour, technology type, and more. For notational convenience, many of these sub- and superscripts are left out when it is not essential.

4.1 Smoothing methodology

In the bottom-up literature, the energy sector is represented by a discrete set of energy producing plants. While the optimization problem of each plant depends crucially on e.g. the technology that is employed, a somewhat general representation is to define the optimal dispatch problem of a plant i as a decision rule or policy function $E^*_{i,h}$:

$$E^*_{i,h}\left(\boldsymbol{p}^E;\mathcal{S}_i\right) \equiv \underset{E}{\text{argmax}}\left\{\nu_{i,h}\left(E,\boldsymbol{p}^E;\mathcal{S}_i\right)\right\}, \tag{1}$$

where $\nu_{i,h}$ is the plant's value function in hour h, E is the choice variable (e.g. how much electricity to dispatch in hour h), $\boldsymbol{p}^E$ is a vector of hourly prices on E, and $\mathcal{S}_i$ is the relevant state space that characterizes the specific plant. Identifying the policy function $E^*_{i,h}$ typically involves solving a non-differentiable problem. The most frequent non-differentiable scenario in techno-economic models of the energy system arises due to capacity constraints resulting in 0/1 discrete events. As the following sections will explain more thoroughly, these types of non-differentiabilities occur (1) in the supply from so-called standard electricity or heat generating plants, (2) in the optimal output split of joint production technologies, and (3) in the way trade in electricity depends on relative prices between neighboring areas. In all of these scenarios, the bottom-up literature models the outcome, e.g electricity generation ($E^*_{i,h}$), as a jump from o to q_i, when the price exceeds a certain threshold value (c_i):

$$E^*_{i,h} = \begin{cases} \text{o}, & p^E_h < c_i \\ q_i, & p^E_h \geq c_i \end{cases} \tag{2}$$

To identify aggregate outcomes the bottom-up approach sums over the corresponding variables on plant level. For instance, yearly supply is given by $E_{agg} = \sum_{i,h} E^*_{i,h}$. Bottom-up models then proceed to solve for equilibrium prices as a linear programming problem. To formulate the model as a square system of nonlinear, smooth equations, we recast the problem of the individual plants as a differentiable one that nests the original formulation.[9] To be precise, in cases where the policy

[9] Another approach can be found in the literature on heterogeneous agents models (see e.g. Algan et al., 2014). Here the original non-differentiable problem in (1) is unchanged, and the aggregation is approximated by e.g. low- or high-order polynomials, depending on the degree of non-linearity of E_{agg} in $\boldsymbol{p}^E$. Newton's method can then be

function in (2) contains a jump, we smooth the jump by applying a *flexible sigmoid function*:

Definition 1 (Flexible Sigmoid Function)**.**
Define a flexible sigmoid function as a mapping $F : \mathbb{R} \to [0, 1]$ *with characteristic parameters* $\gamma \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$, *such that:*

i. F *is monotonically strictly increasing and continuously differentiable for* $\sigma > 0$.

ii. F *approaches the discrete-choice function in the limit*

$$\lim_{\sigma \to 0} F(x;\ \gamma, \sigma) = \begin{cases} 0, & x < \gamma \\ 1, & x \geq \gamma \end{cases}.$$

iii. F*'s supremum (infimum) is* 1 (0)*:*

$$\lim_{x \to \infty} F(x;\ \gamma, \sigma) = 1$$
$$\lim_{x \to -\infty} F(x;\ \gamma, \sigma) = 0.$$

We note that in the simple case (2), the optimal smooth choice is defined in closed-form:[10]

$$E^*_{i,h} = q_i F\left(p^E_h;\ c_i, \sigma\right).$$

We note that this formulation can be interpreted as a plant having increasing and convex generation costs centered around the level c_i. In this interpretation, the sigmoid function F captures the cumulative distribution of marginal costs with mean c_i and standard deviation σ. With this interpretation of F, we can further define the maximized profits (or value function) of the plant as

$$\Pi^*_{i,h} = q_i F(p_h; c_i, \sigma)\left(p_h - \mathbb{E}_F\left[x | X < p_h\right]\right),$$

where $\mathbb{E}_F$ is the conditional expectation of x when it is distributed according to F (see appendix C.1 for details). Thus, when F is defined as e.g. the cumulative normal distribution, the value function has a closed-form solution as well.

4.2 Bottom-Up Derived Short-Run Energy Supply

Individual power and district heat producing plants differ along a number of dimensions, including fuel efficiency and non-fuel related variable costs, to name a few (appendix B includes a more comprehensive description). For our purposes, the plants' technology type is of particular importance, as it determines the type of optimization problem faced by plants in the short run. Table 1

used to solve for $\boldsymbol{p}^E$. For every guess p_0, (1) is solved and E_{agg} is re-approximated. This iterative process continues until convergence.

[10] Two functions that can be written straightforwardly on this form is the standard normal cumulative density function and the logistic function. Interestingly though, the solution to more complicated scenarios, as the solution to a minimization of trade costs (cf. section 4.3 and appendix D) or the profit maximization of CHP plants (cf. appendix C.2) naturally lead to flexible *sigmoid* functions of this type.

provides the technology types and organizes them into groups of plant types:

Table 1: Technology types ($\tau \in \mathcal{T}$)

Technology type:	**Output**	**Flexibility of supply**	**Plant type**
Condensation plant	E	Dispatchable	Standard
Extraction plant	E, H	Dispatchable	CHP
Backpressure plant w. bypass	E, H	Dispatchable	CHP
Backpressure plant	E, H	Dispatchable	CHP
Electric heater	H	Dispatchable	Standard
Wind, solar PV, ROR	E	Intermittent	Intermittent
Boiler heating plant	H	Dispatchable	Standard
Solar heating plant	H	Intermittent	Intermittent
Surplus production	H	Exogenous	Standard
Heat storage	H	Dispatchable	Storage
Electricity storage	E	Dispatchable	Storage
Hydro with storage	E	Dispatchable	Storage

Note: ROR is run-off-river hydro power plants, and PV is short for photovoltaics.

In the following, each of the four plant types are described in turn. A more comprehensive review of how the technical data is used e.g. to compute marginal costs for various plant types is given in Berg and Eskildsen (2019).

Supply from standard plants

There are currently three *standard* technologies: Condensation plants, boiler heaters, and electrical heaters. The condensation plant produces electricity (E) whereas the boiler heater produces heat (H). Electrical heaters are essentially modelled as boiler heaters with the difference that electrical heaters use electricity to power the production whereas boiler heaters depend on fuels. For all three types, the standard bottom-up model approach is to assume a roughly constant fuel-efficiency (Danish Energy Agency, 2018). Thus, the energy input-to-output ratio measured in GJ or MWh is constant resulting in constant marginal costs of production.

Formally, the relevant information on a standard plant i can be summed up in the triplet $[c_i,\ q_i,\ T_i(\mathbf{t})]$, where c_i denotes the marginal cost (without taxes), q_i the capacity constraint, and $T_i(\mathbf{t})$ is a tax function that may depend on plant i's characteristics.[11] The notation $\mathbf{t}$ is shorthand for all regulation that affects marginal costs. The marginal cost component c_i generally depends on technical parameters identified by the inventory data, as well as input prices e.g. on different fuels. In this way, changes in general equilibrium prices translate into changes in the cost of producing electricity and district heat in the bottom up model. The constant marginal costs lead to the following jump function for optimal production:

$$Y_i = \begin{cases} q_i, & p_i^Y \geq c_i + T_i(\mathbf{t}) \\ 0, & p_i^Y < c_i + T_i(\mathbf{t}) \end{cases} \tag{3}$$

[11] The taxation of a plant may depend on geographic area (as rates differ across countries in the model), or individual characteristics as the fuel-efficiency, or whether the plant has abatement equipment installed to limit the emissions of SO_2.

where $Y_i \in \{E_i, H_i\}$ depending on the type of plant. Instead of assuming that marginal costs are exactly constant for $Y_i \in [0, q_i]$, we adopt the general assumption:

Assumption 1 (Nonlinear unit costs)**.**
A standard plant i is characterized by the marginal cost function $c'(Y_i)$

$$c'(Y_i) = g\left(\frac{Y_i}{q_i}; \gamma_i, \sigma\right), \qquad \gamma_i \equiv c_i + T_i(\mathbf{t}), \tag{4}$$

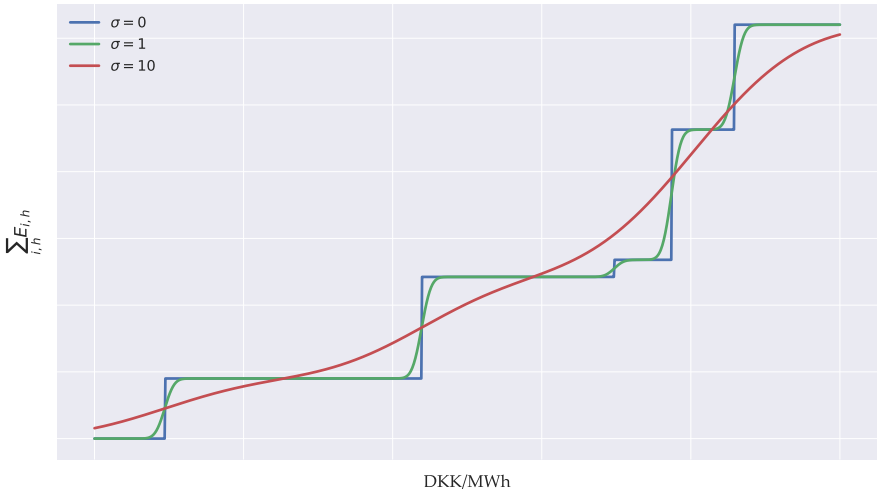
where $g : [0, 1] \to \mathbb{R}$ is the inverse of a flexible sigmoid function as defined in definition 1.

The cost function (4) nests constant unit cost assumption in (3) when $\sigma \to 0$. For $\sigma > 0$ assumption 1 implies the smooth version on the form

$$Y_i = q_i F\left(p_i^Y;\ \gamma_i, \sigma\right), \qquad \gamma_i \equiv c_i + T_i(\mathbf{t}) \tag{5}$$

where F is a flexible sigmoid function as defined in definition 1. To illustrate the effect of σ, figure 4.1 shows the simulated aggregate supply from a number of standard plants with varying marginal costs (c_i); When σ is relatively large, the supply function smoothly increases with the price whereas a small σ allows us to get closer to the step-wise bottom-up function plotted for $\sigma = 0$.

Figure 4.1: Smoothing out supply from standard plants

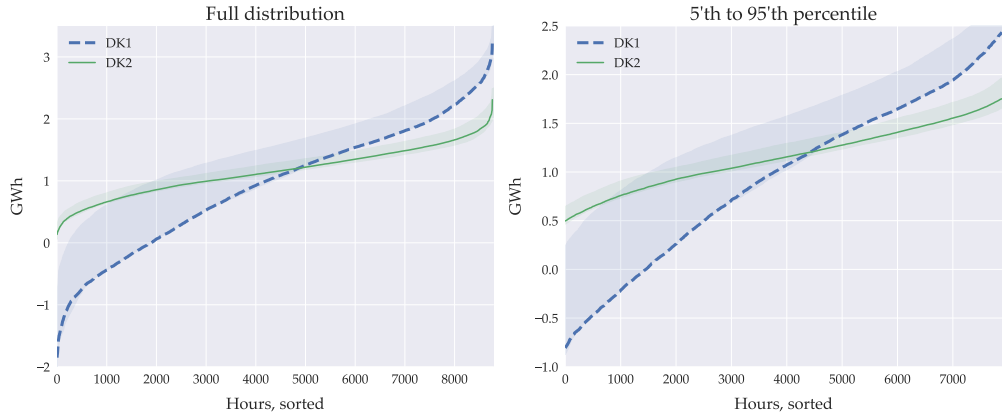


Intermittent vs. dispatchable plants

We refer to technologies as intermittent (as opposed to dispatchable), when the generation capacity varies exogenously throughout the year. In our framework this corresponds to the plant capacity (q_i) varying exogenously across the h hourly states. A second characteristic (albeit not a defining one) of intermittent technologies is that the marginal cost of production is close to zero. Thus, the short run equilibrium is found by balancing the residual demand after intermittent supply, with the supply from dispatchable plants.

Figure 4.2 illustrates the hourly variation in the difference in consumption minus supply from intermittent WWS technologies from 2011-2020. We note that the variation in residual demand is

Figure 4.2: Hourly variation in consumption less intermittent supply, 2011-2020



Note: The graphs show the gross consumption of electricity minus supply from wind, water, and sunlight technologies in 2019. The shaded areas indicate the minimum/maximum interval observed between 2011-2020.
Source: Energy Data Service (2021b).

quite large, in particular in DK1, given that the yearly average consumption per hour is around 2.3 and 1.5GWh for DK1 and DK2 respectively. These features of intermittent/dispatchable technologies are essential to the economic value of different types of plants (Joskow, 2011): Ignoring trade and subsidies, the residual load curve for Denmark in figure 4.2 shows that there are hours where wind production is much larger than the total demand which inevitably drives down equilibrium prices. Dispatchable plants that work as back-up capacity for intermittent technologies can, on the other hand, charge a considerable higher price in peak load hours.

Combined Heat and Power Plants

The model includes three different types of CHP plants: Extraction, back-pressure and back-pressure with bypass. CHP plants are modelled as a joint-production process with the optimization split into two steps: First, the plant produces a composite good $\mathcal{Q}$ at a constant marginal cost $c_i + T_i(\mathbf{t})$. Secondly, this composite good is transformed to heat and electricity. A standard bottom-up assumption is that there is a linear electricity-to-heat transformation rate (C_i^v). Depending on the type of CHP plant, there can additionally be minimum co-production constraints on either electricity (C_i^E) or heat (C_i^H). The implied production possibility set is illustrated in figure 4.3.

The relevant information for plant i can be summed up by the tuple:

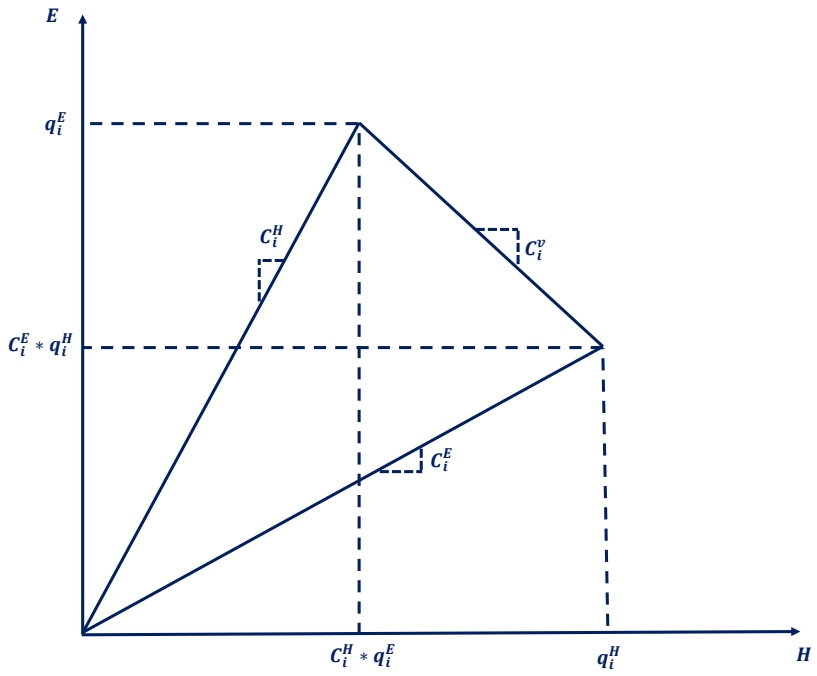
$$\Big(c_i,\ T_i(\mathbf{t}),\ T_i^E(\mathbf{t}),\ T_i^H(\mathbf{t}),\ q_i^E,\ q_i^H,\ C_i^v,\ C_i^E,\ C_i^H\Big).$$

Here $T_i^E(\mathbf{t}), T_i^H(\mathbf{t})$ are tax functions specifying how the plant is taxed depending on whether it produces electricity or heat, and q_i^E, q_i^H denotes capacities for electricity and heat production respectively.[12]

With constant unit costs of producing the composite good ($\mathcal{Q}$) as well as a linear transformation

[12] In Denmark CHP fuel-inputs are taxed differently depending on whether or not the fuel is used for electricity or heat production, see for instance Danish Council on Climate Change (2018b).

Figure 4.3: Production possibility set for CHP plants



rate, it is straightforward to show that the solution is a combination of two corner solutions:

$$E_i = \mathbb{I}_{\mathcal{Q}} \left[q_i^E \mathbb{I}_{\tilde{p}^E > \tilde{p}^H} + \left(1 - \mathbb{I}_{\tilde{p}^E > \tilde{p}^H}\right) C_i^E q_i^H \right], \qquad \mathbb{I}_{\mathcal{Q}} = \begin{cases} 1, & \tilde{p} \geq c_i + T_i(\boldsymbol{t}) \\ 0, & \tilde{p} < c_i + T_i(\boldsymbol{t}) \end{cases}, \tag{6}$$

$$H_i = \mathbb{I}_{\mathcal{Q}} \left[q_i^H \left(1 - \mathbb{I}_{\tilde{p}^E > \tilde{p}^H}\right) + \mathbb{I}_{\tilde{p}^E > \tilde{p}^H} C_i^H q_i^E \right], \qquad \mathbb{I}_{\tilde{p}^E > \tilde{p}^H} = \begin{cases} 1, & \tilde{p}^E > \tilde{p}^H \\ 0, & \tilde{p}^E \leq \tilde{p}^H \end{cases} \tag{7}$$

where the two indicator functions indicate whether or not it is profitable to turn on the plant ($\mathbb{I}_{\mathcal{Q}}$) and whether it is most profitable to produce electricity ($\mathbb{I}_{\tilde{p}^E > \tilde{p}^H}$) or heat.[13] In place of the traditional bottom-up assumptions outlined above, we propose the following:

Assumption 2 (CHP with normalized constant elasticity of transformation (NCET) assumptions)**.**

For all types of CHP plants:

i. Plants produce a composite good $\mathcal{Q}$ with marginal costs of the type

$$c'(\mathcal{Q}) = g\left(\mathcal{Q};\ \gamma_i, \sigma\right), \qquad \gamma_i \equiv c_i + T_i(\mathbf{t}), \tag{8}$$

where $g : [0,1] \to \mathbb{R}$ is the inverse of a flexible sigmoid function as defined in definition 1. The scale of production of the composite good $\mathcal{Q}$ is bounded on $[0,1]$, representing the share of capacity (q_i) that is utilized.

[13] Appendix C.2 elaborates on the derivation of (6)-(7).

ii. The composite good ($\mathcal{Q}$) is transformed into electricity and heat according to

$$q_i\mathcal{Q} = \left[\left(\theta_E^i\right)^{\frac{\epsilon-1}{\epsilon}}\left(E_i - \underline{E}\right)^{\frac{\epsilon-1}{\epsilon}} + \theta_H^{\frac{\epsilon-1}{\epsilon}}\left(H_i - \underline{H}\right)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}}, \qquad \epsilon < -1, \tag{9}$$

with coefficients

$$\theta_E = \frac{q_i^E}{q_i^E - q_i^H C_i^E}, \qquad \theta_H = \frac{q_i^E}{q_i^H - q_i^E C_i^H}, \qquad q_i = q_i^E$$
$$\underline{E} = q_i^H \mathcal{Q} C_i^E, \qquad \underline{H} = q_i^E \mathcal{Q} C_i^H.$$

Compared to the bottom-up assumptions this differs in two important ways: Firstly, part *i.* assumes that the marginal costs are not exactly constant, but take the same form as assumed for standard plants. We note that for $\sigma \to 0$ this nests the bottom-up assumption of constant marginal costs. Secondly, part *ii.* replaces the bottom-up assumption of linear transformation with one of constant elasticity of transformation. We note that for $\epsilon \to -\infty$ this nests the bottom-up assumption of a linear transformation rate. Appendix C.2 shows that under assumption 2 the non-differentiable indicator functions in (6)-(7) are replaced by flexible sigmoid functions that can approximate the bottom-up solution arbitrarily well. Furthermore, the solution in terms of optimal generation of electricity and heat can be written in closed-form as smooth functions of the relevant prices $\tilde{p}^E, \tilde{p}^H$.

Supply from plants with storage capacity

The model includes four different types of plants with storage capacity: Heat storage plants, hydro power plants with reservoirs and water inflow, hydro power plants with pumped storage, and electricity storage plants. Compared to other plants in the model, the short run supply decision for these plants is inherently dynamic and depends crucially on the variation in hourly prices.

While the Danish electricity system does not feature any significant storage capacity, imports from areas where storage is a major factor, e.g. Norway and Sweden, are crucial for formation of equilibrium prices. The companion paper Berg and Eskildsen (2021) investigates the value and potential of storage in the current model.

4.3 Trade in electricity

In large-scale CGE models the traditional approach to modeling trade is the Armington approach (Armington, 1969) that assumes domestic and foreign goods are imperfect substitutes and that the substitution pattern is given by a constant elasticity of substitution. This assumption, however, appears invalid when it comes to trade in electricity. Electricity is a near-homogenous good with capacity constraints on how much can be traded in each point in time.[14] In place of the Armington assumption, we develop a trade mechanism that emulates the one observed on Nord Pool's day-ahead market. This essentially boils down to adopting three assumptions: (i) Electricity is

[14] One might still attempt to approximate an elasticity of substitution based on system operations rather than preferences, e.g. Reguant (2019).

homogeneous, (ii) the market is competitive, and (iii) trade is constrained by the transfer capacity of transmission lines. Letting $T_{i,j}$ denote the transmission line capacity from country i to j and abstracting from trade costs, this implies a trade function of the form:

$$NX_{i,j}\begin{cases} = T_{i,j}, & p_i^E < p_j^E \\ \in [-T_{j,i}, T_{i,j}], & p_i^E = p_j^E \\ = -Tj, i, & p_i^E > p_j^E \end{cases}, \qquad \forall (i,j) \in (\mathcal{G}_E \times \mathcal{G}_E), \tag{10}$$

Note that (10) is defined for all geographical pairs (i, j). Whether or not prices are equalized between i and j depends not only on the pair of transmission cables $(T_{i,j}, T_{j,i})$; If a third area, k, has free trade capacity vis á vis i and j, electricity can flow between i and j via k. If all transmission capacities are non-binding the law of one price holds for all electricity regions.

When all agents have free access to transmission lines, a competitive, decentralized equilibrium will result in the net export function in (10).[15] As our model relies on complete and competitive markets with complete information, the fundamental welfare theorem implies that an equivalent interpretation is one of an electricity exchange that seeks to maximize the value of traded electricity subject to transmission line constraints:

$$V_{i,j} = \max_{NX_{i,j}} \left\{ \left(p_j^E - p_i^E \right) NX_{i,j} \right\} \qquad \text{st. } -T_{j,i} \leq NX_{i,j} \leq T_{i,j} \tag{11}$$

When $p_j^E > p_i^E$ it is optimal to set $NX_{i,j} = T_{i,j}$. Conversely, when $p_j^E < p_i^E$ the exchange sets $NX_{i,j} = -T_{j,i}$. When price differences are zero, the exchange is indifferent and $NX_{i,j} \in [-T_{j,i}, T_{i,j}]$. The kinked net export function in (10), however, relies on the assumption of no trade costs. Consider instead the following assumption:

Assumption 3 (Marginal trade cost)**.**
Let $C_{i,j}(NX_{i,j})$ denote the trade cost of net exports $NX_{i,j} \in [-T_{j,i}, T_{i,j}]$ between areas $i, j \in \mathcal{G}_E$, where $T_{i,j}, T_{j,i} > 0$. Let the marginal trade costs follow the form

$$C'_{i,j}\left(NX_{i,j}\right) = g\left(\frac{NX_{i,j} + T_{j,i}}{T_{i,j} + T_{j,i}}; 0, \sigma\right) - \kappa_{i,j}, \qquad \kappa_{i,j} \equiv g\left(\frac{T_{j,i}}{T_{j,i} + T_{i,j}}; 0, \sigma\right) \tag{12}$$

where $g : [0, 1] \to \mathbb{R}$ is the inverse of a flexible sigmoid function as defined in definition 1.

Figure 4.4 illustrates the shape of the marginal trade cost function under assumption 3. The constant term $\kappa_{i,j}$ is a normalizing term, which ensures that marginal costs are zero when there is no trade. Thus, trade only occurs when prices differ across electricity areas. Furthermore, note that if area i imports (exports) electricity from (to) j marginal trade costs are negative (positive). When F and its inverse F^{-1} are strictly increasing and bijective functions, the marginal trade

[15] This is the standard economic result that with free and costless trade, prices will equalize. With a physical constraint on how much electricity can be traded, however, this standard result is supplemented by the two potential corner solutions.

costs are furthermore increasing in the amount of traded electricity:[16]

$$\frac{\partial^2 C_{i,j}}{\partial NX_{i,j}^2} = \frac{1}{T_{j,i}+T_{i,j}} \frac{1}{(F)'\left[g\left(\frac{NX_{i,j}+T_{j,i}}{T_{i,j}+T_{j,i}}\right)\right]} > 0.$$

Consequently, $C_{i,j}(NX_{i,j})$ is a strictly convex function with a global minimum in $NX_{i,j} = 0$. Under assumption 3 the decentralized, competitive equilibrium still coincides with the one of a cost minimizing electricity exchange that now solves:

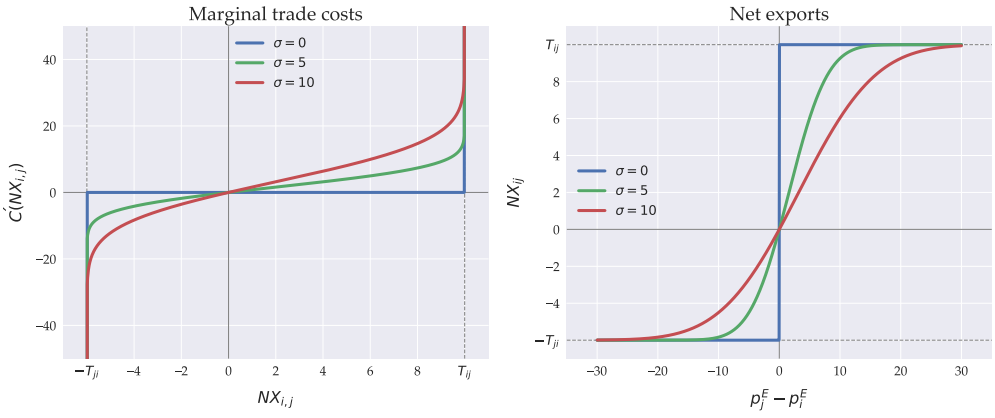
$$V_{i,j} = \max_{NX_{i,j} \in (-T_{j,i}, T_{i,j})} \left\{ \left(p_j^E - p_i^E\right) NX_{i,j} - C(NX_{i,j}) \right\} \qquad \text{s.t. assumption 3.} \tag{13}$$

In appendix D we show that the trade cost minimizing equilibrium indeed implies a net export function on the form:

$$NX_{i,j} = -T_{j,i} + \left(T_{i,j} + T_{j,i}\right) F\left(p_j^E - p_i^E + \kappa_{i,j}; 0, \sigma\right). \tag{14}$$

Figure 4.4 illustrates the shape of the net export function. Importantly, net exports are zero around $p_j^E = p_i^E$. When $p_j^E < p_i^E$ country i imports electricity and exports electricity when $p_j^E > p_i^E$. Furthermore, net exports naturally satisfy the physical bounds imposed by transmission line capacities. Finally, as σ approaches zero the net export function converges to the kinked case in (10). Thus, when the smoothing parameter is sufficiently low, the trade costs are approximately zero whenever trade is within the transmission line capacities, but increases sufficiently fast around the bounds to ensure that capacities are never exceeded. This is illustrated in figure 4.4.

Figure 4.4: Trade in electricity with smooth marginal costs



Note: The flexible sigmoid F is given by the normal cumulative density function. Transmission capacities are $T_{j,i} = 6$ and $T_{ij} = 10$.

Trade with variation in transfer capacities

As discussed in section 2.1, investments in a larger and more integrated European transmission

[16] This result follows from the inverse function theorem.

grid are underway as a part of a strategy to incorporate larger shares of intermittent renewables in the electricity system (ENTSO-E, 2021). However, just as the generation capacity of intermittent plants fluctuates across the year, so does the availability of transmission lines between regions. The difference between the technical transfer capacity of transmission lines (Total Transfer Capacity, TTC) and the actual available capacity (Net Transfer Capacity, NTC) can occur for multiple reasons. The main reasons include variability of intermittent renewables, power plant outages, outages in the transmission lines themselves, and variation in load patterns.[17]

To account for the variation in NTC over the course of a year, we augment the marginal trade cost function in assumption 3. Specifically, we add two hourly and geographic specific shocks to the transmission line capacity:

$$T_{i,j,h} = T_{i,j}\eta_{i,j,h}\upsilon_{i,j,h}, \tag{15}$$

where $\eta_{i,j,h} \in (0,1)$ and $\upsilon_{i,j,h} \in \{0,1\}$ represent a continuous and a discrete shock, respectively. Whereas the continuous shock $\eta_{i,j,h}$ is straightforward to accommodate, the discrete shock may result in $T_{i,j,h} = 0$, which is not in the supported domain for the net export function in (14). Thus, this augmentation requires adjustments to the assumptions on marginal costs of trade:

Assumption 4 (Marginal trade costs with zero events.)**.**
Let $C_{i,j,h}(NX_{i,j,h})$ denote the trade costs of net exports of electricity $NX_{i,j,h}$ between areas $i, j \in \mathcal{G}_E$ in hour h where $T_{i,j,h}$ and/or $T_{j,i,h} > 0$. The marginal trade costs follows:

$$\frac{\partial C_{i,j,h}}{\partial NX_{i,j,h}} = \sigma\left[g\left(\frac{NX_{i,j,h} + T_{j,i,h}}{T_{i,j,h} + T_{j,i,h}};\ 0, 1\right) - \kappa_{i,j,h}\right], \tag{16}$$

where $g : [0,1] \to \mathbb{R}$ is the inverse of a flexible sigmoid function as defined in definition 1 and $\kappa_{i,j,h}$ is a normalizing term:

$$\kappa_{i,j,h} = \begin{cases} 0, & \text{if } T_{i,j,h}T_{j,i,h} = 0 \\ g\left(\dfrac{T_{j,i,h}}{T_{j,i,h} + T_{i,j,h}};\ 0, 1\right), & \text{else.} \end{cases}$$

In instances where transmission lines are available in both directions ($\upsilon_{i,j,h} = \upsilon_{j,i,h} = 1$) marginal trade costs are identical to those under assumption 3, and the net export function is identical with equation (14). In cases where only one of the transmission lines are open, however, the resulting net export function is now a smoothed jump function around $p^E_{i,h} = p^E_{j,h}$. For instance, in the case where imports from area j to i are not available, the net exports function is defined as:

$$NX_{i,j,h} = T_{i,j,h}F\left(p^E_{j,h} - p^E_{i,h}; 0, \sigma\right), \qquad \text{when } T_{j,i,h} = 0 \text{ and } T_{i,j,h} > 0.$$

One caveat when using this simple augmentation of the net export function is that the net exports can be positive when foreign prices are lower than domestic ones. Appendix D.2 provides a simple

[17] See ENTSO-E's user information on Net Transfer Capacities (NTC) here: https://eepublicdownloads.entsoe.eu/clean-documents/pre2015/ntc/entsoe_NTCusersInformation.pdf.

method that alleviates this issue.

Estimation of hourly variation in NTCs

To provide a realistic description of trade in electricity, we estimate the variation in NTC using an augmented double truncated Tobit model. The current section provides an overview of the method and results. We refer to appendix H for more details.

We collect data on NTCs ($T_{i,j,h,t}$) and the yearly TTC ($T_{i,j,t}$) for geographical combinations (i,j), hours (h), and years (t). The variable of interest is the percentage reduction of NTC compared to the TTC, i.e. $\Delta T_{i,j,h,t} \equiv T_{i,j,h,t}/T_{i,j,t} - 1$. The domain of this variable is naturally confined to $\Delta T_{i,j,h,t} \in [-1, 0]$, which can be explained by a double truncated Tobit model. To explain the observed variation in NTC we include two potential sources of variation. First, we assume that transmission line availability can exist in three states (Z) that are independent to variables in our model: (i) A shock to the system can lead to an outright outage $(Z = -1)$, (ii) the system can be affected by a less severe shock that lowers transmission capacity $(Z = \xi)$, or (iii) the system can be unaffected by shocks $(Z = 0)$. Second, if there is no forced outage, there might still be disruptions to the NTC that may be correlated e.g. with variation in load patterns. To capture this, we define a latent variable $\Delta T^*_{i,j,h,t}$ defined by a linear model that depends on hourly load and wind generation. This results in the augmented Tobit model of the form:

$$\Delta T_{i,j,h,t} = \begin{cases} 0, & \text{if } Z_{i,j,h,t} = 0, \quad \text{or} \quad Z_{i,j,h,t} = \xi \text{ and } \Delta T^*_{i,j,h,t} \geq 0 \\ \Delta T^*_{i,j,h,t} & \text{if} \qquad\qquad\quad Z_{i,j,h,t} = \xi \text{ and } \Delta T^*_{i,j,h,t} \in (-1, 0) \\ -1, & \text{if } Z_{i,j,h,t} = -1, \quad \text{or} \quad Z_{i,j,h,t} = \xi \text{ and } \Delta T^*_{i,j,h,t} \leq -1 \end{cases} \tag{17}$$

Assuming the variables $Z, \Delta T^*$ are orthogonal, we estimate this using maximum likelihood. The full estimation procedure and results are reported in appendix H but as an illustrative example, we consider the case of export capacities from DK2 to Sweden for the years 2015-2020. This is illustrated in figure 4.5. Panel (a) shows the sorted hourly reduction in export capacities for each year. In all years there is a significant reduction in at least 2000 hours. While there is also signification variation in transmission capacities across years, we estimate the hourly variations for a *normal* year using observations for all years. [18] Panel (b) shows the predicted reduction in transmission capacities compared with data.[19]

4.4 Short Run Energy Demand

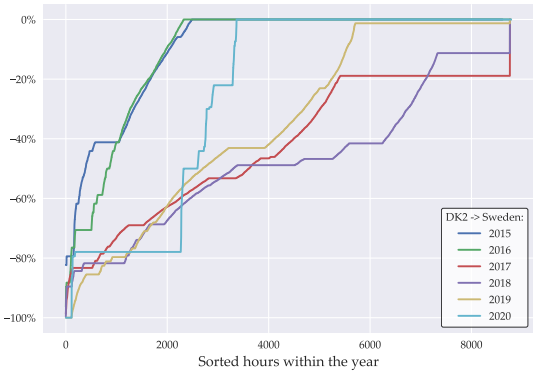
The short run demand for electricity and district heat is defined as a mix of perfectly inelastic and habit-driven demand, and a price-responsive part. Consider the demand for commodity $x^d \in \{E^d, H^d\}$, and let $\boldsymbol{p}^x$ denote a vector of all hourly prices of good x in a given year. The hourly

[18] The fundamental assumption is that there is no yearly trend in the linear model for $\Delta T^*_{i,j,h,t}$ but year-specific shocks are permissible.

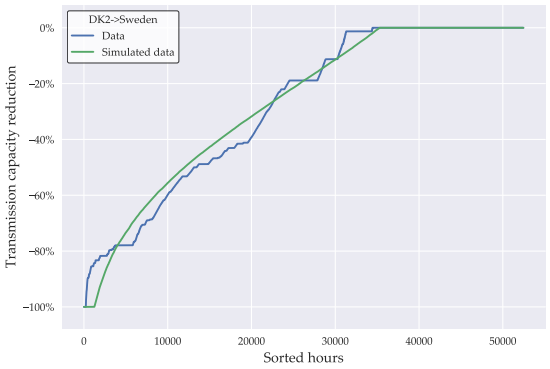
[19] Appendix H explains how we use the estimated parameters to simulate hourly transmission capacities that are consistent with the bottom-up model.

Figure 4.5: Hourly variation in export transmission capacities from DK2 to Sweden

(a) Observed hourly variation per year



(b) Estimated hourly variation for a normal year



demand x_h^d is defined as:

$$x_h^d\left(\boldsymbol{p}^x\right) = \phi\overline{x}_h + \left(1-\phi\right)\overline{x}_h L_h\left(\boldsymbol{p}^x\right), \qquad L_h\left(\boldsymbol{p}^E\right) = 1 \text{ if } \boldsymbol{p}^E \text{ is constant},$$

where ϕ is the share of short-run demand that is inelastic, $\overline{x}_h$ is the hourly level of purely habit-driven demand, and L_h is a price function that is decreasing in the price in the corresponding hour p_h^x. Rewriting the habit-driven level of demand as shares of the yearly level, $\overline{x}_h = z_h x_t^d$, we can rewrite the hourly demand function as

$$x_h^d\left(\boldsymbol{p}^x\right) = x_t^d \frac{z_h\left[\phi + (1-\phi)L_h\left(\boldsymbol{p}^x\right)\right]}{\sum_j z_j\left[\phi + (1-\phi)L_h\left(\boldsymbol{p}^x\right)\right]}, \tag{18}$$

where z_h measures habits in hourly demand in shares of yearly demand. In this way, we think of short run demand as the allocation of yearly demand (x_t) into hourly components depending on exogenous habits (z_h), how large a share of demand that cannot be re-allocated (ϕ), and how re-allocation depends on prices $(L_h(\boldsymbol{p}^x))$. The demand specification in (18) is convenient for a number of reasons: First, the simple inelastic short run demand assumption is nested by letting $\phi = 1$, while for $\phi < 1$ we can simulate the effects of varying the degree of flexibility of short run demand. Second, in this formulation it is possible to identify the exogenous habits (z_h) from hourly consumption data (appendix E.2 provides a framework for doing this). Third and finally, the formulation guarantees consistency between the hourly and yearly time-increments in the model, while allowing the yearly level of demand (x_t^d) to be determined by the CGE model.

4.5 The short run equilibrium

With demand, supply, and trade in electricity in place, we can now characterize the short run equilibrium. Let $(\boldsymbol{p}_{g_E}^E, \mathbf{p}_{g_H}^H)$ denote vectors of all hourly prices in year t. Furthermore, let $(\mathcal{I}_{g_E}, \mathcal{I}_{g_H})$ denote the set of plants in the relevant area (g_E, g_H). The short run equilibrium is given by definition 2.

Definition 2 (The Short Run Equilibrium)**.**
The short run equilibrium in hour h and year t, on electricity markets $g_E \in \mathcal{G}_E$ and heat markets $g_H \in \mathcal{G}_H$, consists of a set of electricity prices and consumption levels $\left\{p^E_{g_E}, E_{g_E}\right\}_{g_E \in \mathcal{G}_E}$, regional heat prices and consumption levels $\left\{p^H_{g_H}, H_{g_H}\right\}_{g_H \in \mathcal{G}_H}$, and net export flows $\left\{NX_{i,j}\right\}_{(i,j) \in \mathcal{G}^2_E}$ where:

i. Heat demand and supply are equalized. Both coincide with equilibrium consumption. For all $g_H \in \mathcal{G}_H$:

$$H^d\left(\mathbf{p}^H_{g_H}\right) = \sum_{i \in \mathcal{I}_{g_H}} H_i\left(\mathbf{p}^E_{g_E}, \mathbf{p}^E_{g_H}\right). \tag{19}$$

ii. Electricity demand, and sum of net exports equals domestic supply. Equilibrium consumption in area g_E coincides with demand. For all $g_E \in \mathcal{G}_E$:

$$E^d\left(\mathbf{p}^E_{g_E}\right) + \sum_{g_j \in (\mathcal{G}_E \setminus g_E)} NX_{g_E, g_j} = \sum_{i \in \mathcal{I}_{g_E}} E_i\left(\mathbf{p}^E_{g_E}, \mathbf{p}^H_{g_H}\right) \tag{20}$$

iii. Net export of electricity is determined by relative prices. For all $(g_E, g_j) \in \mathcal{G}^2_E$:

$$NX_{g_E, g_j} = -T_{g_j, g_E} + \left(T_{g_E, g_j} + T_{g_j, g_E}\right) F\left(p^E_{g_j} - p^E_{g_E};\ \mathrm{o}, \sigma\right), \tag{21}$$

where F belongs to the class of flexible sigmoid functions cf. Definition 1.

Note that due to CHP plants, electrical heaters, plants with capacity for energy storage, flexible demand components, and trade in electricity, all short run equilibria within a year in the markets for heat and electricity are determined simultaneously.

The main result characterizing this short run equilibrium in the bottom up model is summed up in the following:

Proposition 1 (A Smooth, Linearly Homogeneous Short Run Equilibrium)**.**
Consider the bottom-up model of the short run equilibrium on electricity and heat markets that consists of (i) standard plants, (ii) CHP plants, (iii) intermittent plants, (iv) storage plants, (v) trade in electricity, and (vi) short run energy demand. Assume that short run supply and trade in electricity is characterized by Assumptions 1-3. For the trade-cost minimizing short run equilibrium in definition 2 we have the following:

i. The short run equilibrium can be formulated as a differentiable, nonlinear, and simultaneous square system of equations.

ii. The short run equilibrium nests the traditional bottom-up solution, in the sense that letting the set of smoothing parameters (σ) tend to zero, we can get arbitrarily close to the bottom-up assumptions.

iii. The short run equilibrium is linearly homogeneous in the sense that increasing (1) all plant capacities, (2) yearly levels of electricity and heat demand, and (3) trade capacities by $\alpha\% >$

o, *has the implication that for all hours, markets, and geographical areas:*

(i) Equilibrium prices remain unchanged,

(ii) the shares of electricity demand covered by net exports remain unchanged, and

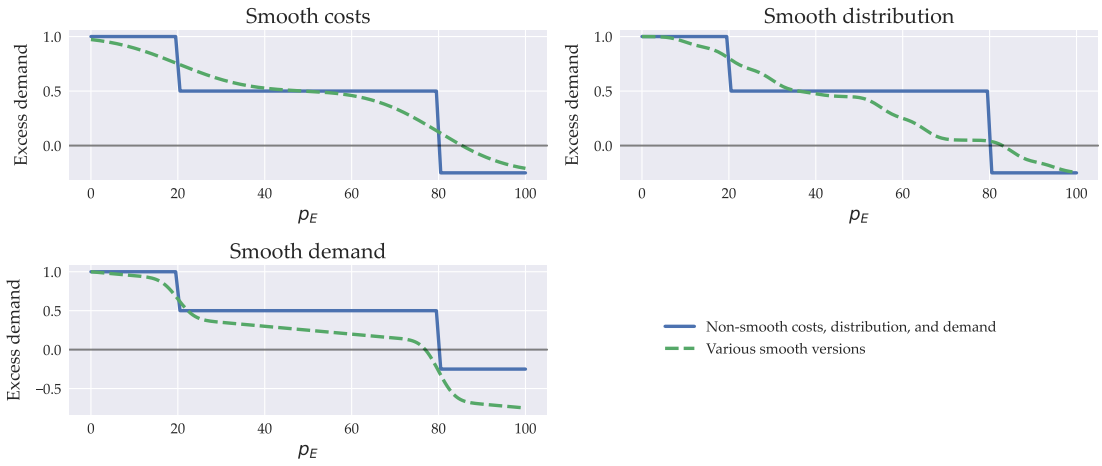
(iii) all plants increase production by α%,

Proof. See appendix F. □

The result in proposition 1 implies that (*i.*) the bottom-up model can be solved simultaneously with a conventional CGE model, and (*ii.*) that it can be seen as a smooth extension of traditional linear optimization systems. The implications of the third and final point, that the system is linearly homogenous in capacities, is not further explored in this paper. However, the linearity in capacities indicates that, for a given set of long-run assumptions e.g. on prices, it is likely that there exists a steady state investment strategy with a constant composition of plant types and capacities for trade in electricity. We leave this for future research.

With the definition of our short run equilibrium in place, we revisit the point that the smoothing approach is not appropriate in all bottom-up models. Specifically, the excess demand functions in (19)-(20) have to be sufficiently smooth to ensure reasonably fast convergence. This is problematic if the bottom-up model displays (i) perfectly inelastic demand, (ii) approximately constant marginal generation costs, and (iii) there are large intervals in the distribution of average generation costs that have zero density. Figure 4.6 illustrates this point by plotting one (solid line) excess demand function where (i)-(iii) are all in effect, and three plots (dashed lines) where one of the three conditions does not hold implying a smooth excess demand function. In the first (upper-left) the excess demand function becomes smooth because marginal costs are sufficiently smooth; in the second (upper-right) the marginal costs are approximately constant, but there are many more plants with different marginal costs; in the third (lower-left) demand is no longer perfectly inelastic.

Figure 4.6: Smooth vs. non-smooth bottom-up structures



Note: The blue solid line indicates a case with inelastic demand (condition i.) and two large plants with largely different marginal costs (condition iii.) that are approximately constant (condition ii.). The three dashed lines show the effect of dropping one of the three conditions.

5 Calibration and Model Fit

5.1 Calibration

The bottom-up model of the electricity and district heat systems is calibrated to the baseline year 2019. We select a set of statistics that ensure the model reproduces the yearly energy statistics that the CGE model is calibrated to as well. We note that while the model includes representation of foreign electricity markets, the objective of the model is to evaluate the effects of policies in Denmark. Thus, for our purposes, yearly statistics on the Danish electricity and heat systems are essential. It is straightforward to target a broader set of statistics in the calibration procedure, including information on foreign equilibrium prices and net exports. Specifically, we target moments from Danish Energy Agency (2021) yearly energy statistics listed in table 2. Furthermore, we calibrate the yearly average electricity price to the spot price on the day-ahead market on Nord Pool (Energy Data Service, 2021a; Energy Data Service, 2021b). The calibration procedure is based on reproducing the relevant statistics by adjusting a large set of technical coefficients as little as possible, measured as squared deviations from the technical data on Danish plants, cf. section 2.1. Appendix G elaborates on this.

Table 2: Calibration targets, 2019

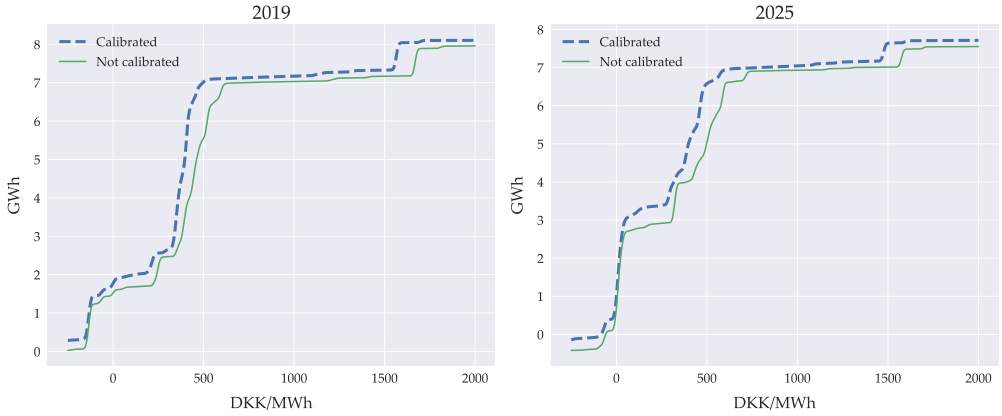
Variable	**Target** (PJ)
Coal (input)	33.51
Oil (input)	3.40
Natural gas (input)	26.98
Straw (input)	13.10
Wood (input)	68.43
Waste (input)	37.36
Bio gas (input)	8.63
Bio oil (input)	0.14
Electricity for district heat (input)	1.61
District heat from electricity (output)	1.97
Surplus heating (output)	4.33
Wind, PV, hydro (output)	61.67
Solar heat (output)	2.27
Gross production, electricity	106.29
Gross production, district heating	131.37
Net import (DK)	20.92
	Target
Average spot price (DK)	299.39 DKK/MWh

Note: Following the convention of the data available on plants, the category natural gas includes bio methane.

Reassuringly, the model is able to replicate the targets in table 2 with only minor adjustments to the technical data on Danish plants. Figure 5.1 confirms this by plotting the aggregate supply curve in 2019 and 2025 before and after the calibration. Only two small adjustments are needed: At prices around 300-400 DKK/MWh and again around 1500 DKK/MWh. In both instances, the supply curve is shifted to the left, implying a cheaper supply in these regions. The first change

is undertaken primarily to ensure that the average spot price is 299 DKK/MWh while the second shift ensures that a certain amount of oil (3.4 PJ) is used in the baseline year. Appendix G provides additional evidence on effects of the calibration.

Figure 5.1: The effect of calibration on the electricity supply function, non-storage plants



Note: The 'not calibrated' series indicates the aggregate supply from plants prior to the calibration procedure, i.e. when technical coefficients are given by the techno-economic data cf. section 2.1.

5.2 Baseline Model Features and Validation

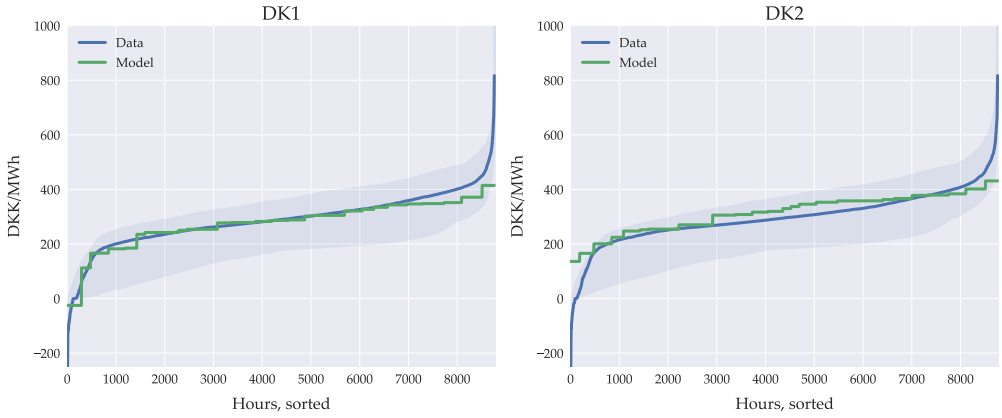
In this section we validate the model in terms of statistics the model has not been targeted to fit and outline the main novelties in the aggregate supply curve it generates. Appendix A.2 provides supplementary figures on the baseline scenario of the model.

An important validation of the model is the representation of intermittency and the variation of prices within a year. Figure 5.2 illustrates the model fit of intra-yearly prices. The shaded areas indicate the range between minimum and maximum hourly prices in the period of 2010-2019. The solid lines are the distribution of actual and simulated prices for 2019. For both of the two domestic regions (DK1 and DK2) the model provides a reasonable fit of the variation in intra-yearly prices. It should, however, be noted that the model is not capturing hours of the year where prices are at their very highest or lowest. While the thin tails are essential for certain types of analyses, e.g. to analyze capacity reserve requirements, they are not essential to the economic value of electricity producing technologies.[20]

Figure 5.3 plots the projected price distribution for 2020, 2030, and 2040. Over time, the model predicts a significant increase in the intra-yearly price variation, in particular due to hours with peak prices. This is mainly due to back-up capacity being retired, leaving only the most costly back-up plants (domestic or foreign) to balance demand and supply in hours with little intermittent energy production.

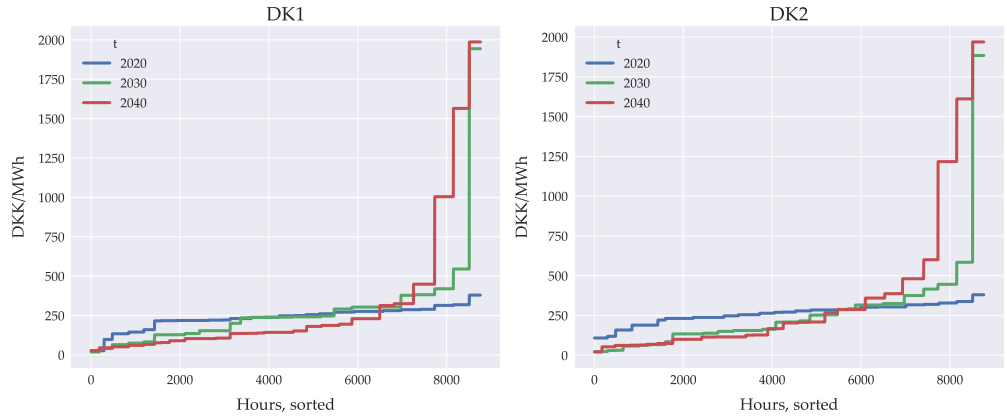
[20] In the companion paper Berg and Eskildsen (2021), we show that increasing the number of intra-yearly states K does not provide a better fit of the intra-yearly prices. Thus, replicating the tails of the distribution likely requires that we model short-run uncertainty and ramp-up constraints.

Figure 5.2: Model fit of intra-yearly prices



Source: Energy Data Service (2021a).

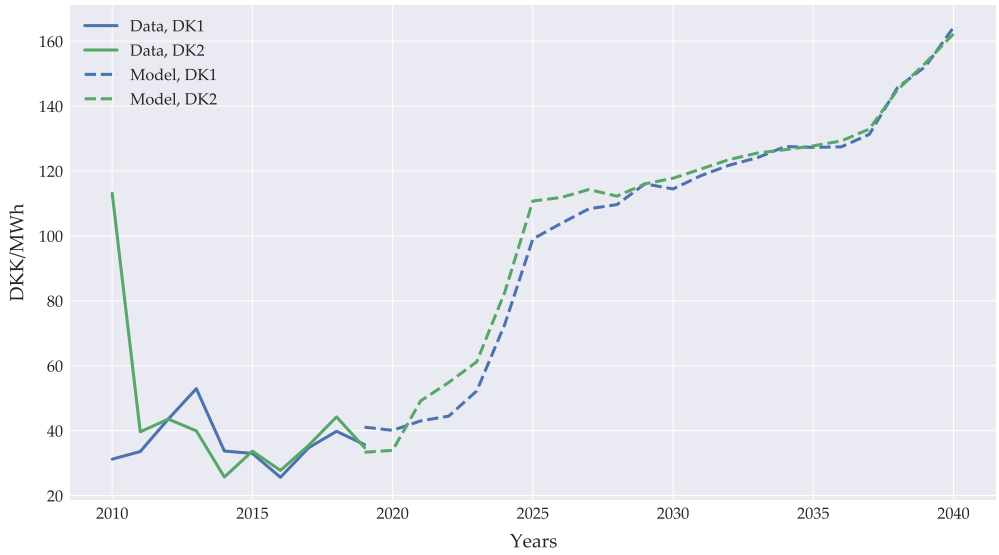
Figure 5.3: Intra-yearly price variation, model projection



The intra-yearly price variation in our model further implies a realistic so-called 'downlift' on wind power in Danish price regions. The downlift is defined as the difference between the yearly average price of electricity and the average price received for electricity produced by wind turbines. Figure 5.4 illustrates the data on prices between 2011-2019, and the model results for 2019-2040. Between 2011-2019 the downlift is roughly 30-40 DKK/MWh, similar to the level in our model in 2019 and 2020. Furthermore, note that the downlift increases drastically over time in Denmark. Thus, the model captures the fact that when more wind enters the system and dispatchable back-up capacity is retired, the value of wind power is driven down relative to the average plant.

Finally, we showcase the effects of deriving aggregate supply from the technology-rich bottom-up model. With the detailed information on plant heterogeneity, we can, for instance, consider which primary energy goods are used in the production of electricity and district heat for different price levels. Figure 5.5 illustrates the supply curve from figure 5.1 in terms of what primary energy sources are used in production. The left panel illustrates the composition in 2019: We note that intermittent renewable energy (WWS) is fully utilized when the prices turn positive. Between 0 and 300 DKK/MWh the supply of electricity slowly increases; around this threshold large natural

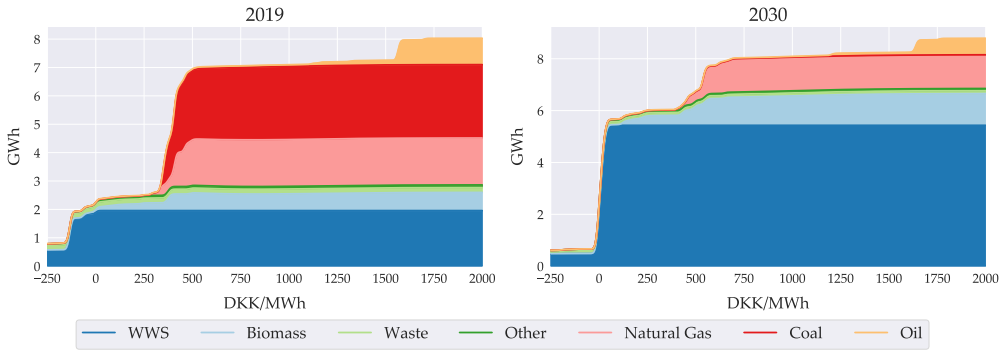
Figure 5.4: Downlift on wind, historic and model



Source: Energy Data Service (2021a), Energy Data Service (2021b).

gas and coal fired CHP plants kick in. Between roughly 500 DKK/MWh and 1600 DKK/MWh, however, the supply function is relatively flat. At this point, the final peak supply comes from plants that rely primarily on oil.[21] The right panel in figure 5.5 shows that by 2030, the share of intermittent renewable energy has increased dramatically while coal has been phased out in exchange for biomass fired plants.

Figure 5.5: Fuel composition of aggregate electricity supply



Note: 'WWS' is wind, water and sunlight technologies. The plot assumes a yearly average capacity factor for intermittent plants (primarily WWS).

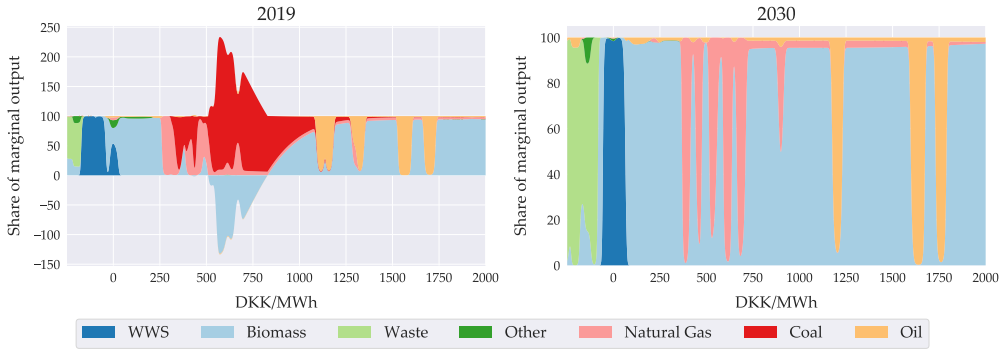
To get a clearer sense of what is happening on the margin of the aggregate supply curve, figure 5.6 illustrates the fuel composition of the marginal electricity supply for 2019 (left panel) and 2030 (right panel). When prices are negative, supply is mainly provided by a number of plants that receive subsidies (this is particularly waste incineration plants) or receive feed-in-tariffs (in particular WWS technologies).[22]

[21] See figure G.2 in appendix G.

[22] There is an additional effect from electrical heaters, in particular when prices are negative. This is not included in figure 5.6. See also figure G.1 in appendix G.

An interesting observation is that fuel shares on the margin are both above 100% and below 0% when prices are between 400-500 DKK/MWh. This result is driven by the fact that some large CHP plants can utilize either some type of biomass or coal and that fuels used for heat and electricity production are taxed differently. When prices of electricity are relatively low, these CHP plants primarily produce heat.[23] As coal used for heat production is taxed at a significantly higher rate than biomass, the CHP plants primarily use biomass in this case. When the price of electricity increases, the CHP plants lower their production of heat and increase production of electricity; as they pivot towards electricity production, the tax benefits from using biomass disappears, and it becomes optimal to use coal instead.

Figure 5.6: Fuel share of marginal electricity supply



Note: 'WWS' is wind, water and solar.

We emphasize that the advantage of using a technology-rich model of the electricity and district heat system, as compared to a conventional nested CES specification applied in CGE models, is that we capture these types of highly nonlinear and heterogeneous effects. The following section illustrates how this type of effects plays an important role when it comes to evaluating policies in an integrated model.

6 Carbon pricing in an Integrated Model

In Denmark greenhouse gas emissions are subject to either the allowance price in the EU ETS or a national CO_2-e tax.[24] Generally, electricity generation is covered by the EU ETS while district heat is regulated under the national CO_2-e tax. For combined power and heat plants, however, the fuel used to produce heat is subject to both the national CO_2 tax and the allowance price (Danish Council on Climate Change, 2018a). We note that, in our baseline scenario, the national CO_2 tax covers roughly 78-80% of Danish greenhouse gas emissions between 2019-2030.

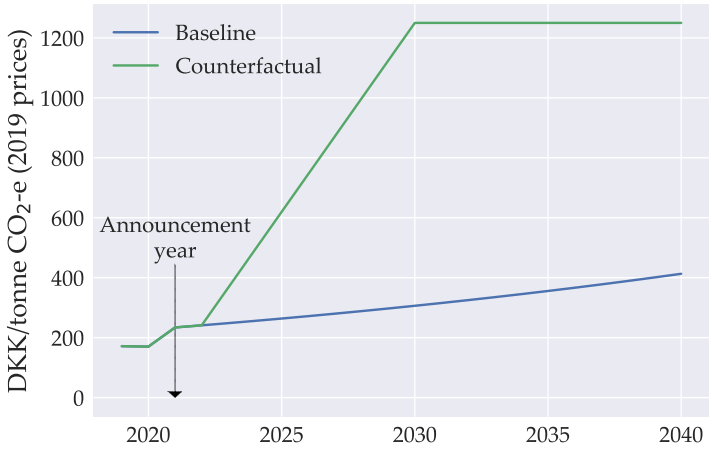
We simulate the effect of gradually increasing a domestic CO_2-e tax from 2023 to 2030. To highlight the importance of the bottom-up model, we compare the effects of the full integrated

[23] Recall that 'extraction' and 'back-pressure with bypass' plants both can adjust how much of its output goes to heat relative to electricity.

[24] We add the '-e' to indicate that the tax is levied on CO_2 equivalents.

model with a 'stand-alone CGE' model. We assume the CO_2-e tax increases linearly from 2023 to a level of 1250 DKK/tonne CO_2-e in 2030 (in 2019 prices). Furthermore, the new CO_2-e tax is announced in 2021 to capture anticipation effects. The counterfactual scenario is illustrated in figure 6.1.

Figure 6.1: National CO_2-e tax scenarios



Note: The CO_2-e tax is gradually increased from 2023-2030. However, the tax is announced in 2021.

To capture the effect of the national CO_2-e tax on the EU ETS allowance price and foreign emissions, we use a fixed leakage rate L for emissions covered by the EU ETS. In simple cap-and-trade systems the leakage rate is simply 100%; due to the latest reform, however, this is not the case for the EU ETS. As the time horizon for the bottom-up model is 2040, we use the average of the 2030 and 2050 estimates for the leakage rate, L=96.55%, from Beck and Kruse-Andersen (2020). Furthermore, assuming that allowance prices follow a Hotelling rule, the allowance price adjusts in the announcement year to ensure that $\Delta CO_2{}^{\text{Foreign}} = -L \times \Delta CO_2{}^{\text{Domestic}}$, where the endogenous variables $\Delta CO_2{}^{\text{Foreign}}$ and $\Delta CO_2{}^{\text{Domestic}}$ are the accumulated changes in, respectively, foreign and domestic CO_2-e emissions up to 2040.[25] If accumulated domestic emissions over the horizon fall (i.e $\Delta CO_2{}^{\text{Domestic}} < 0$) then foreign emissions increase by $-L \times \Delta CO_2{}^{\text{Domestic}} > 0$ because allowance prices adjust downwards. We note that it is only possible to include this mechanism in the integrated model, because foreign emissions are endogenous in the bottom-up model.

In the following we discuss the aggregate results, equilibrium effects in the district heat sector, the electricity sector, and finally carbon leakage.

Aggregate results

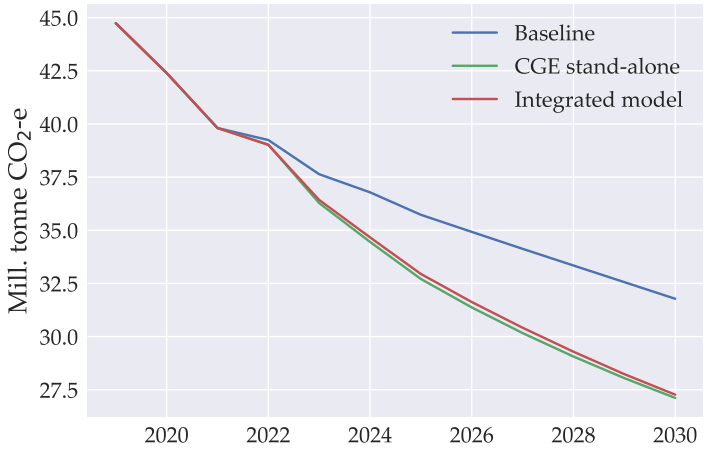
In the baseline scenario, without increases in the CO_2-e tax, domestic greenhouse gas emissions decrease by roughly 29% from 44.7 Mt in 2019 to 31.7 Mt by 2030.[26] Figure 6.2 illustrates that

[25] In models for the EU ETS, the path of the allowance price is typically given by a Hotelling rule (see e.g. Beck and Kruse-Andersen, 2020; Silbye and Sørensen, 2019; Perino and Willner, 2017).

[26] National aggregate greenhouse gas emissions are calculated under the UNFCCC principles.

the increase in the CO_2-e tax leads to 4.6 Mt reductions in emissions by 2030 in the 'stand-alone CGE' model. Surprisingly, this result is very similar in the integrated model. Furthermore, the underlying drivers of emission reductions are similar for the two models: The CO_2-e intensity of energy consumption, the energy intensity of production, and economic activity all decreases compared to the baseline (see appendix I).

Figure 6.2: Aggregate national CO_2-e emissions



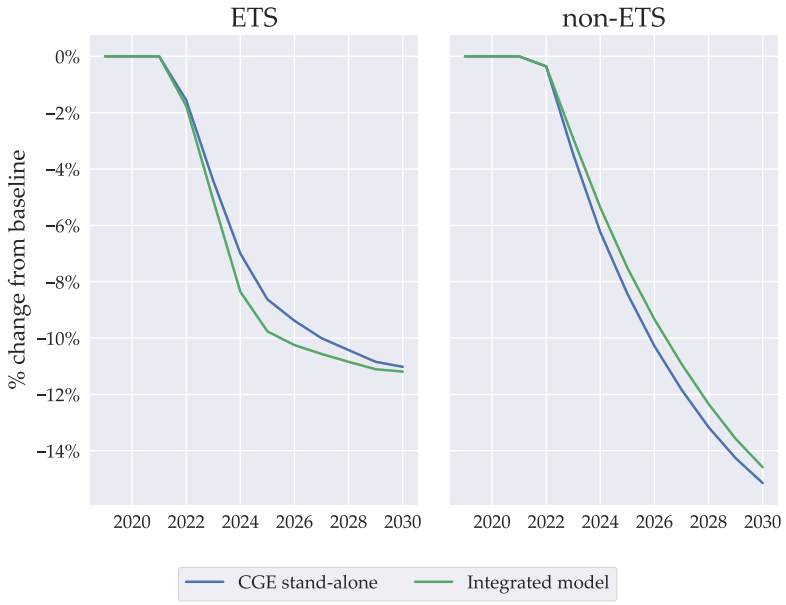
Note: National CO_2-e emissions are calculated under the UNFCCC principles.

While the aggregate results are almost identical, the emission reductions come from different sectors in the two counterfactuals. As illustrated in figure 6.3, the integrated model estimates larger reductions in the ETS and smaller ones in the non-ETS sector relative to the 'stand-alone CGE'.

For the non-ETS sector, there are two reasons why the integrated model predicts higher emissions: First, in the integrated model, the marginal costs of generation are simply not increased as much as the simple 'stand-alone' model predicts (the following subsections elaborate on this). Second, the types of plants that are no longer profitable are different. In both models, the costs of waste incineration increases; however, in the integrated model, waste incineration plants remain profitable (see figure I.3). In the ETS sector, the integrated model primarily predicts lower emissions due to the CHP plants in the bottom-up model. This creates a tight link between the CO_2-e tax and the costs of emission-intensive electricity production in the integrated model: Even though the CO_2-e tax is only levied on fuel used for heat production, the overall profitability of emission-intensive CHP plants decreases, which lowers their production of electricity as well. This also explains why the difference in emission reductions are largest in the early years (2023-2028), before large shares of emission-intensive CHP generation capacity are retired.

It turns out that the source of the CO_2-e reductions is important for other key variables. For instance, as the higher tax rate is charged on non-ETS emissions, the integrated model predicts significantly higher tax revenue than the 'stand-alone CGE' model. Figure 6.4 illustrates that by

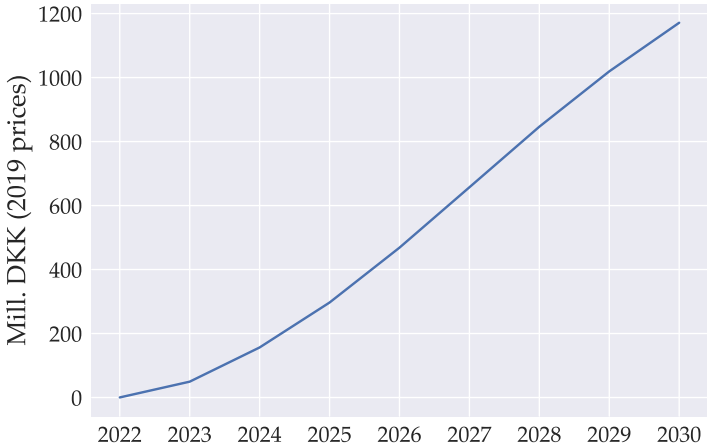
Figure 6.3: Change in national CO_2-e emissions from the baseline



Note: 'ETS' are CO_2-e emissions covered by the EU ETS and therefore not subject to the CO_2-e tax. Conversely, 'non-ETS' are emissions that are not included in the EU ETS but subject to the CO_2-e tax.

2030, the integrated model estimates an additional 1.2 billion DKK (2019 prices) in accumulated tax revenue compared to the 'stand-alone CGE' model.

Figure 6.4: Additional accumulated tax revenue in integrated model



Note: The additional accumulated tax revenue from the CO_2-e tax is referenced by the additional accumulated tax revenue in the 'stand-alone CGE' model.

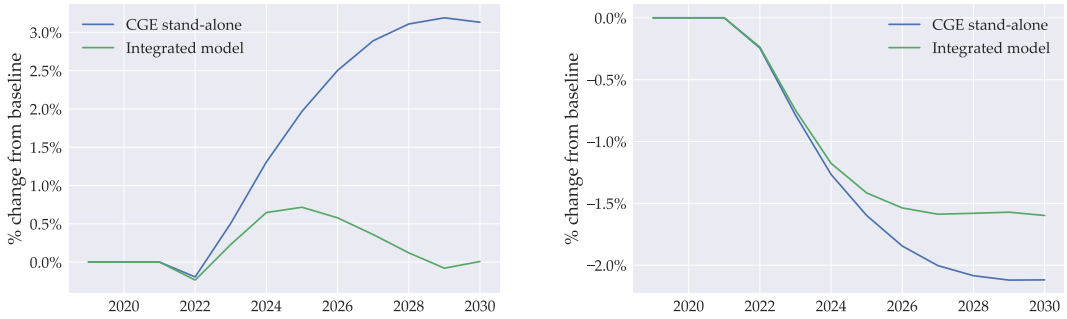
Equilibrium effects in the district heat sector

Figure 6.5 shows predicted changes in prices and consumption. In general, the figure illustrates that higher taxes on emissions in the district heat sector, naturally, leads to higher prices and lower consumption. This is partly due to the direct impact of higher production costs in the district heat

sector, but also because of a general decline in energy demand driven by a combination of lower economic activity and energy intensity. We note that as heat is not traded internationally, the only difference between consumption and production is a small distribution loss. Thus, production of district heat follows the decline in consumption.

Figure 6.5: Equilibrium effects in the domestic district heating sector

(a) District heat prices (b) District heat consumption



Note: Since district heat is not traded the only difference between consumption and production are distribution losses.

Panel (a) shows that district heat prices react differently in the two models.[27] The 'stand-alone CGE' model predicts a gradual increase of prices towards 2030 as a result of higher production costs. The integrated model, on the other hand, only predicts a small transitory increase in district heat prices. This is primarily driven by two technologies captured by the bottom-up model: Heat pumps that produce cheaper district heat when electricity prices decrease, and CHP plants that switch relatively easily from fossil fuels to other sources, e.g. biomass.

Panel (b) shows that the difference in prices can be traced in the level of consumption. In the integrated model, the level of demand stabilizes from around 2025 when the district heat price starts to decrease. However, due to a generally lower energy demand, the level of demand stabilizes at a lower level after 2025.

Equilibrium effects in the electricity sector

Next we turn to the equilibrium effects in the electricity sector. Figure 6.6 shows the changes in key variables for the two models. Panel (b) shows that consumption of electricity follows the general drop in energy demand and decreases by roughly 3% for both models. The drop in demand further results in lower prices on electricity cf. panel (a).[28] Furthermore, panel (c) shows that while consumption drops quite a lot, the domestic generation is barely affected in the integrated model. The reason for this is that the marginal supply of electricity primarily comes from foreign plants: Thus, when consumption decreases marginally, this primarily translates into lower imports

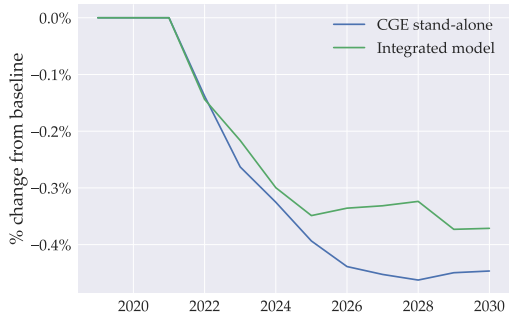
[27] We note that both models estimate a small price drop in 2022: This is an anticipation effect captured by both models: As they anticipate higher energy costs in 2023, they substitute capital for energy by investing in new machines in 2021. Since capital takes a year to become productive, district heat demand and prices decreases in 2022.

[28] Recall that the CO_2-e tax is not levied on electricity production. This explains why prices decrease for electricity, but increase for district heat.

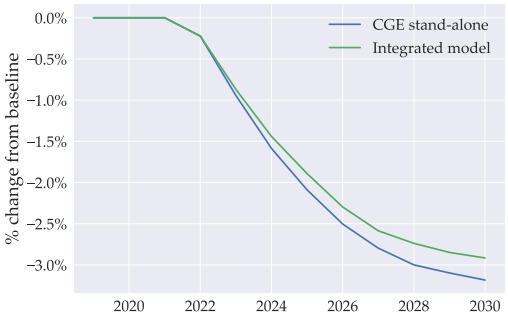
(see panel d)).

Figure 6.6: Equilibrium effects in the domestic electricity sector

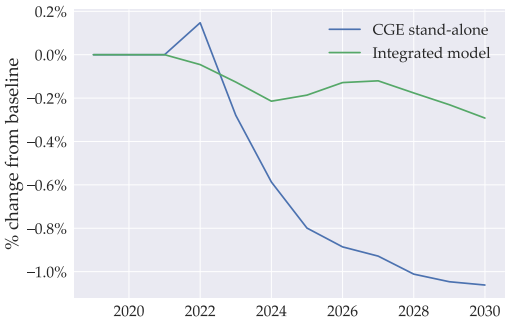
(a) Average yearly prices



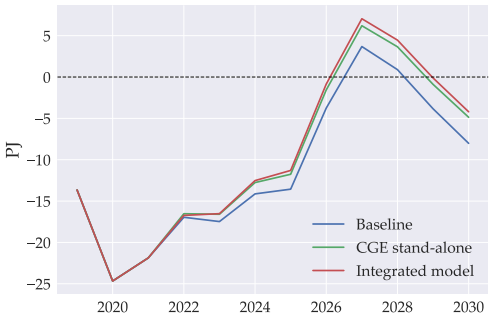
(b) Consumption



(c) Generation



(d) Net exports of electricity



Note: Net exports are shown in levels. The remaining variables are shown in percentage change from baseline values.

CO_2-e emissions in foreign electricity markets

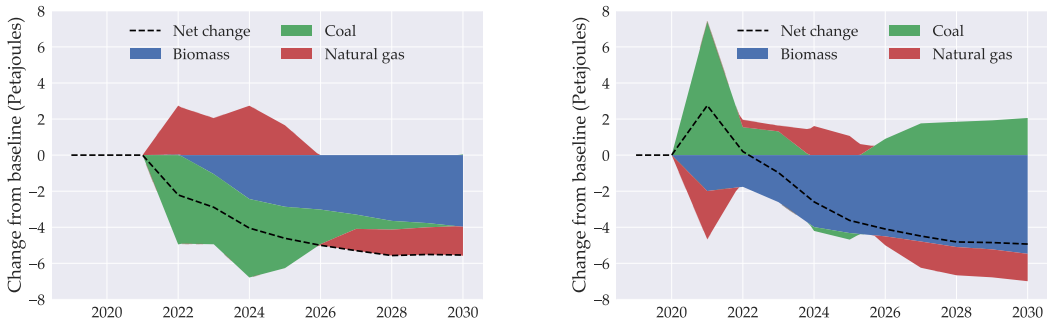
One of the advantages of the integrated model is that it allows us to examine the effect of the CO_2-e tax on greenhouse gas emissions in foreign electricity markets. Because domestic emissions drop, the allowance price in the announcement year drops by 0.67%. The small effect that Danish emissions has on the allowance price illustrates that Denmark is a small economy. However, it does have an effect on foreign emissions. To decompose the effects of the decrease in allowance prices, figure 6.7 shows changes in the fuel consumption of foreign plants when the allowance price is exogenous (panel a) and when the allowance price is endogenous (panel b).

When the allowance price is exogenous (panel a) the production costs of foreign plants are not directly affected by the shock. In this case, the main shock to foreign plants comes from lower levels of Danish import of electricity; this naturally translates into lower total fuel consumption. However, whereas coal and biomass consumption is lowered in this case, the use of natural gas increases from 2022-2026. This is once again a technology effect from the bottom-up model. In most hours of the year, the increased CO_2-e tax leads to a decreased Danish demand, which leads to the lower fuel consumption of coal and biomass. However, in some hours of the year, Danish heat pumps increase production, which leads to increased Danish demand for foreign electricity.

Naturally, the hours where heat pumps are active, the equilibrium prices are generally lower than on average. In these hours, the marginal foreign supply does not rely on coal or biomass, but rather on natural gas.

Figure 6.7: Fuel change in foreign markets in the integrated model

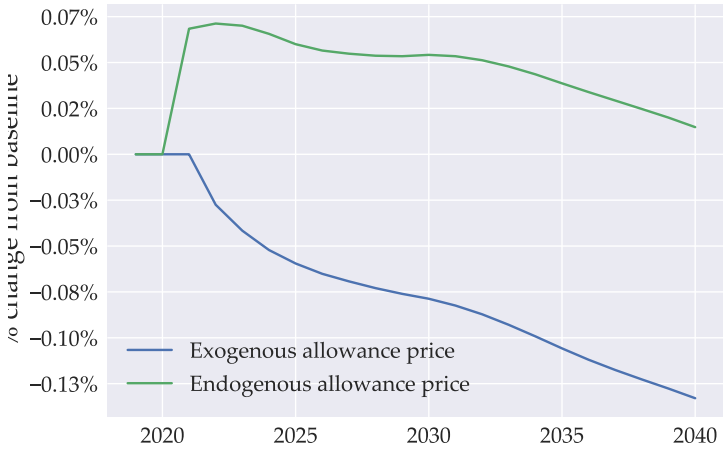
(a) Exogenous allowance price (b) Endogenous allowance price



Note: Net exports are shown in levels. The remaining variables are shown in percentage change from baseline values.

When the allowance price is endogenous (panel b), the reduction in Danish emissions covered by the EU ETS leads to a lower allowance price in general. This makes the more emission-intensive producers more competitive, which leads to an initial increase in coal and a reduction in consumption of biomass and natural gas. As depicted in figure 6.8, the result is that accumulated foreign emissions over the period 2021-2040 increase by 0.014% compared to the baseline result. While this is a small percentage, it is still an additional 4.8 Mt CO_2-e in total.[29]

Figure 6.8: Change in accumulated foreign emissions in the Integrated model



Note: Accumulated emissions are the emission from the announcement year (2021) to time t.

[29] Recall that we fixed the leakage rate in the EU ETS at 96.55%. The implication here is that Danish accumulated emissions from 2021-2040 covered by the EU ETS decrease by 4.8/0.9655 Mt CO_2-e.

7 Conclusion

Achieving net-zero greenhouse gas emissions requires large structural changes in the economy and depends crucially on the penetration of intermittent renewable energy. To estimate the effects of economic-, energy-, and climate policy initiatives for the green transition, we require analytical tools that capture essential technologies in the energy system as well as general equilibrium effects. To this end, we presented a technologically detailed bottom-up model for the electricity and district heat system that can be directly integrated with large-scale general equilibrium models. The bottom-up model solves for equilibria for electricity and district heat markets at the hourly level by balancing demand and supply from a large set of heterogeneous energy producing plants. The main contribution of the paper was methodological: We used smoothing techniques to specify the model as a square nonlinear system of equations instead of the conventional (mixed-integer) linear optimization problems. Our smoothing method relies on recasting the profit maximization problem of plants as smooth and convex optimization problems. We showed for all technologies except for storage technologies, the optimal dispatch decision can be summarized by closed-form expressions that are smooth functions of equilibrium prices. Thus, solving our bottom-up model boils down to a root-finding problem of identifying prices that ensure equilibrium in all states of the model. Beyond allowing for nonlinear elements, this allows us to efficiently solve both the bottom-up and general equilibrium model simultaneously using gradient-based methods.

To accurately capture the economic value and potential of intermittent, renewable energy, the bottom-up model uses hourly variation in generation capacity of plants, demand for electricity and heat, and capacity of transmission lines. Importantly, we showed that there are large hourly fluctuations in how much electricity can be transferred through transmission lines. To capture this, we estimated the variation for all relevant transmission lines in our model using a Tobit type I model. Furthermore, we introduced an intuitive and simple way of modeling hourly demand that encompasses habits and partial price-responsiveness. This approach allows us to match current hourly patterns of demand and simulate the effects of introducing load switching devices.

As an application of the framework, we applied the bottom-up model to the case of Denmark. We showed that calibrating the model to replicate yearly aggregates from the 2019 energy statistics for the Danish electricity and district heat systems only requires small adjustments to model parameters. We provided evidence on the type of nonlinearities that the technologically founded bottom-up model can capture, but the standard economic model is not likely to. Furthermore, we showed that the bottom-up model provides a realistic intra-yearly price variation in electricity prices, and that it produces a realistic downlift on wind power in Denmark. Thus, the bottom-up model accurately captures the economic value of intermittent technologies relative to dispatchable ones.

We coupled the bottom-up model with a large-scale CGE model for the Danish economy called GreenREFORM and simulated the effect of increasing the national carbon tax on domestic CO_2-e

emissions gradually towards 2030. To emphasize the effect the having an integrated model, we compared results to the case of a 'stand-alone' CGE model without the bottom-up model. We showed that the 'stand-alone' CGE model and the integrated model provides similar predictions in terms of emission reductions, but differ significantly when it comes to the effects in the electricity and district heat sectors. In general, the integrated model predicted smaller changes in prices and consumption of electricity and district heat. Also, the 'stand-alone' CGE model underestimated the additional cumulative tax revenue by 1.2 billion DKK between 2022-2030.

Finally, we note that investments in the bottom-up model are exogenous. They cover currently planned investments as well as a projection for future investments from Denmark's Energy and Climate Outlook. Future research will focus on endogenizing these investments.

References

Algan, Yann et al. (2014). *Solving and Simulating Models with Heterogeneous Agents and Aggregate Uncertainty*. Vol. 3. Handbook of Computational Economics. Chap. 6, pp. 277–324.

Ambec, S. and C. Crampes (2017). *Decarbonizing electricity generation with intermittent sources of energy*. Working paper, n. 15-603. Toulouse School of Economic, Toulouse, France.

Armington, Paul S. (1969). "A Theory of Demand for Products Distinguished by Place of Production". In: *Staff Papers (International Monetary Fund)* 16.1, pp. 159–178. ISSN: 00208027.

Bahar, Heymi and Jehan Sauvage (2013). "Cross-Border Trade in Electricity and the Development of Renewables-Based Electric Power". In: *OECD Trade and Environment Working Papers*.

Balyk, Olexandr et al. (2019). "TIMES-DK: Technology-rich multi-sectoral optimisation model of the Danish energy system". In: *Energy Strategy Reviews* 23, pp. 13–22. ISSN: 2211-467X.

Beck, Ulrik Richardt and Peter Kjær Kruse-Andersen (2020). "Endogenizing the Cap in a Cap-and-Trade System: Assessing the Agreement on EU ETS Phase 4". English. In: *Environmental and Resource Economics* 77, pp. 781–811. ISSN: 0924-6460.

Berg, Rasmus K. and Janek Eskildsen (2019). "Modelling the energy sector in a computable general equilibrium framework: A new approach to integrated bottom-up and top-down modelling".

Berg, Rasmus K. and Janek Eskildsen (2021). "The Value and Potential of Energy Storage".

Berg, Rasmus K., Jens S. Kirk, and Peter Stephensen (2020). *Strategy for integrating the energy sector model and the CGE model in GreenREFORM*. https://dreamgroup.dk/publications/2020/may/strategy-for-integrating-the-energy-sector-model-and-the-cge-model-in-greenreform/. Accessed: 29-07-2021.

Bistline, John et al. (July 2020). "Energy storage in long-term system models: a review of considerations, best practices, and research needs". In: *Progress in Energy* 2.3, p. 032001.

Bloess, Andreas, Wolf-Peter Schill, and Alexander Zerrahn (2018). "Power-to-heat for renewable energy integration: A review of technologies, modeling approaches and flexibility potentials". In: *Applied Energy* 212, pp. 1611–1626.

Bloomfield, P. (2004). *Fourier Analysis of Time Series: An Introduction*. Wiley Series in Probability and Statistics. Wiley. ISBN: 9780471653998.

Böhringer, Christoph (1998). "The synthesis of bottom-up and top-down in energy policy modeling". In: *Energy Economics* 20.3, pp. 233–248. ISSN: 0140-9883.

Böhringer, Christoph and Thomas F. Rutherford (2008). "Combining bottom-up and top-down". In: *Energy Economics* 30.2, pp. 574–596. ISSN: 0140-9883.

Böhringer, Christoph and Thomos F. Rutherford (2009). "Integrated assessment of energy policies: Decomposing top-down and bottom-up". In: *Journal of Economic Dynamics and Control* 33.9, pp. 1648–1661. ISSN: 0165-1889.

Brown, T. et al. (2018). "Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system". In: *Energy* 160, pp. 720–739. ISSN: 0360-5442.

Danish Council on Climate Change (2018a). *Fremtidens grønne afgifter på energiområdet*. Tech. rep. Copenhagen, Denmark: Danish Council on Climate Change.

Danish Council on Climate Change (2018b). *Status for Danmarks klimamålsætninger og forpligtelser 2018*. Tech. rep. Copenhagen, Denmark: Danish Council on Climate Change.

Danish Energy Agency (2018). *RamsesR docomentation*. https://ens.dk/sites/ens.dk/files/Analyser/ramsesr.pdf. Accessed: 03-09-2018.

Danish Energy Agency (2021). *Energistatistik 2019*. https://ens.dk/sites/ens.dk/files/Statistik/energistatistik2019_dk-webtilg.pdf. Accessed: 18-10-2021.

Delarue, E and J Morris (Mar. 2015). *Renewables Intermittency: Operational Limits and Implications for Long-Term Energy System Models*. Tech. rep. The MIT Joint Program on the Science and Policy of Global Change.

Drud, Arne (2019). *GAMS documentation 27, CONOPT manual*. https://www.gams.com/latest/docs/S_CONOPT.html. [Online; accessed 16-12-2021].

EA Energy Analyses (2018). *Balmorel Userguide*. https://ea-energianalyse.dk/papers/Balmorel_UserGuide.pdf. Accessed: 04-07-2019.

Energy Commission (Apr. 2016). *Baggrundsnotat med fakta om den danske energisektor i dag*. https://efkm.dk/media/8283/baggrundsnotat-med-fakta-om-energisektoren-i-dag.pdf. Downloaded 18 November 2018.

Energy Data Service (2018). *Electricity Balance Data*. https://www.energidataservice.dk/dataset/electricitybalance/resource_extract/498c68e3-d248-4965-b36f-3aa738130adc. [Online; accessed 02-10-2018].

Energy Data Service (2021a). *Elspot Prices*. https://www.energidataservice.dk/tso-electricity/elspotprices. [Online; accessed 19-10-2021].

Energy Data Service (2021b). *Production and Consumption - Settlement*. https://www.energidataservice.dk/tso-electricity/productionconsumptionsettlement. [Online; accessed 19-10-2021].

ENTSO-E (2017a). *Day-ahead Total Load Forecast* (Version 30 October 2020) [Data set]. https://transparency.entsoe.eu/load-domain/r2/totalLoadR2/show.

ENTSO-E (2017b). *Generation Forecasts for Wind and Solar* (Version 30 October 2020) [Data set]. https://transparency.entsoe.eu/generation/r2/dayAheadGenerationForecastWindAndSolar/show.

ENTSO-E (2021). *Completing The Map: Power System Needs in 2030 and 2040*. Tech. rep. Accessed: 17-10-2021. ENTSO-E, p. 70.

Grøn Energi & Ea Energianalyse (2016). *Energiforsyning 2030 - Baggrundsrapport*. Tech. rep. Kolding, Denmark: Grøn Energi.

Hirth, Lion, Falko Ueckerdt, and Ottmar Edenhofer (2015). "Integration costs revisited – An economic framework for wind and solar variability". In: *Renewable Energy* 74, pp. 925–939. ISSN: 0960-1481.

Hittinger, Eric and Roger Lueken (2015). "Is inexpensive natural gas hindering the grid energy storage industry?" In: *Energy Policy* 87, pp. 140–152. ISSN: 0301-4215.

Horridge, Mark et al. (2013). "Chapter 20 - Solution Software for Computable General Equilibrium Modeling". In: *Handbook of Computable General Equilibrium Modeling SET, Vols. 1A and 1B*. Ed. by Peter B. Dixon and Dale W. Jorgenson. Vol. 1. Handbook of Computable General Equilibrium Modeling. Elsevier, pp. 1331–1381.

IEA (Aug. 2021a). *Greenhouse Gas Emissions from Energy: Overview*. https://www.iea.org/reports/greenhouse-gas-emissions-from-energy-overview. Accessed: 2021-10-08.

IEA (2021b). *Net Zero by 2050*. Tech. rep. Lead authors and co-ordinators, Team, Bouckaert, Stéphanie and Pales, Araceli F. and Mcgalde, Christophe and Remme, Uwe and Wanner, Brent. IEA.

Jacobson, Mark Z. et al. (2017). "100% Clean and Renewable Wind, Water, and Sunlight All-Sector Energy Roadmaps for 139 Countries of the World". In: *Joule* 1.1, pp. 108–121. ISSN: 2542-4351.

Joskow, Paul L. (2011). "Comparing the Costs of Intermittent and Dispatchable Electricity Generating Technologies". In: *The American Economic Review* 101.3, pp. 238–241. ISSN: 00028282.

Krook-Riekkola, Anna et al. (2017). "Challenges in top-down and bottom-up soft-linking: Lessons from linking a Swedish energy system model with a CGE model". In: *Energy* 141, pp. 803–817. ISSN: 0360-5442.

Loulou et al. (2016). *Documentation for the TIMES Model*. https://iea-etsap.org/docs/Documentation_for_the_TIMES_Model-Part-II_July-2016.pdf. Accessed: 04-07-2019.

Lund, Henrik et al. (2021). "EnergyPLAN – Advanced analysis of smart energy systems". In: *Smart Energy* 1, p. 100007. ISSN: 2666-9552.

Mathiesen, Lars (1985). "Computation of economic equilibria by a sequence of linear complementarity problems". In: *Economic Equilibrium: Model Formulation and Solution*. Ed. by Alan S. Manne. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 144–162. ISBN: 978-3-642-00917-4.

Nord Pool (2021a). *Historical Market Data: External Nordic capacities* [Data sets]. https://www.nordpoolgroup.com/historical-market-data/. [Online, accessed: 01-09-2021].

Nord Pool (2021b). *Historical Market Data: Market coupling capacities* [Data sets]. https://www.nordpoolgroup.com/historical-market-data/. [Online, accessed: 01-09-2021].

Perino, Grischa and Maximilian Willner (2017). "EU-ETS Phase IV: allowance prices, design choices and the market stability reserve". In: *Climate Policy* 17.7, pp. 936–946. eprint: https://doi.org/10.1080/14693062.2017.1360173.

Reguant, Mar (2019). "The Efficiency and Sectoral Distributional Impacts of Large-Scale Renewable Energy Policies". In: *Journal of the Association of Environmental and Resource Economists* 6.S1, pp. 129–168.

Rutherford, Thomas F. (1995). "Extension of GAMS for complementarity problems arising in applied economic analysis". In: *Journal of Economic Dynamics and Control* 19.8, pp. 1299–1324. ISSN: 0165-1889.

Silbye, Frederik and Peter Birch Sørensen (May 2019). "National Climate Policies and the European Emissions Trading System". English. In: *Nordic Economic Policy Review* 2019.12, pp. 63–101. ISSN: 1904-4526.

Siskos, Pelopidas et al. (2018). "Implications of delaying transport decarbonisation in the EU: A systems analysis using the PRIMES model". In: *Energy Policy* 121, pp. 48–60. ISSN: 0301-4215.

Zerrahn, Alexander, Wolf-Peter Schill, and Claudia Kemfert (2018). "On the economics of electrical storage for variable renewable energy sources". In: *European Economic Review* 108, pp. 259–279. ISSN: 0014-2921.

Zhang, Runsen and Shinichiro Fujimori (Feb. 2020). "The role of transport electrification in global climate change mitigation scenarios". In: *Environmental Research Letters* 15.3, p. 034019.

Appendices

A Descriptive data, additional figures

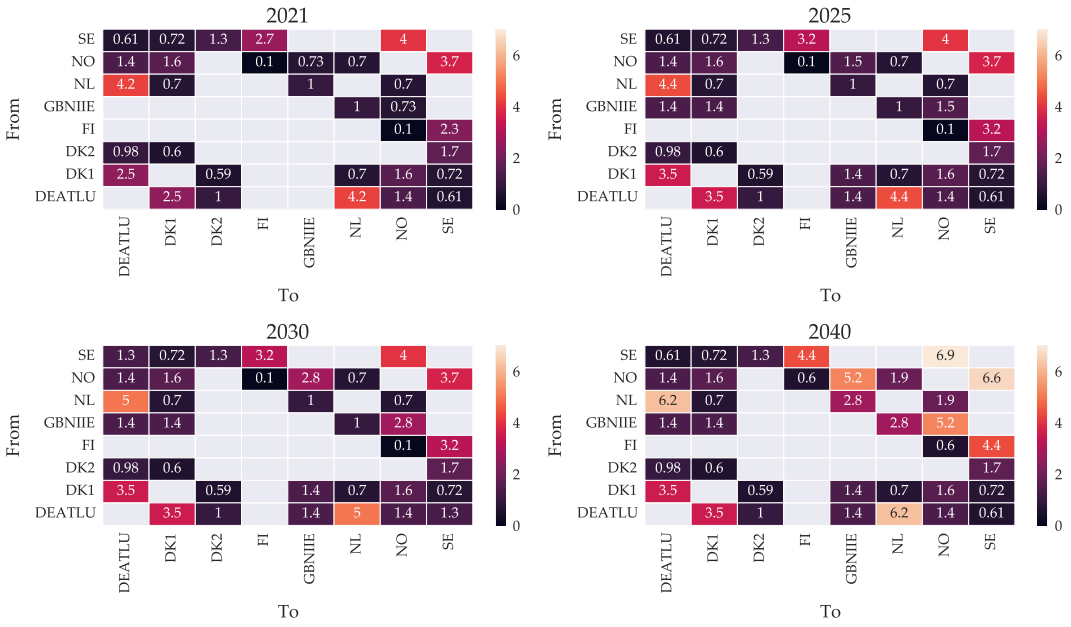
This appendix shows additional data on the bottom-up model for the Danish electricity and district heat systems. The first part plots some additional descriptive evidence on the input data used in the model, including how investments in technologies and transmission lines change over time. The second part plots some endogenous outcomes from the model in the baseline scenario.

A.1 Model data

Increasing transmission line capacities over time

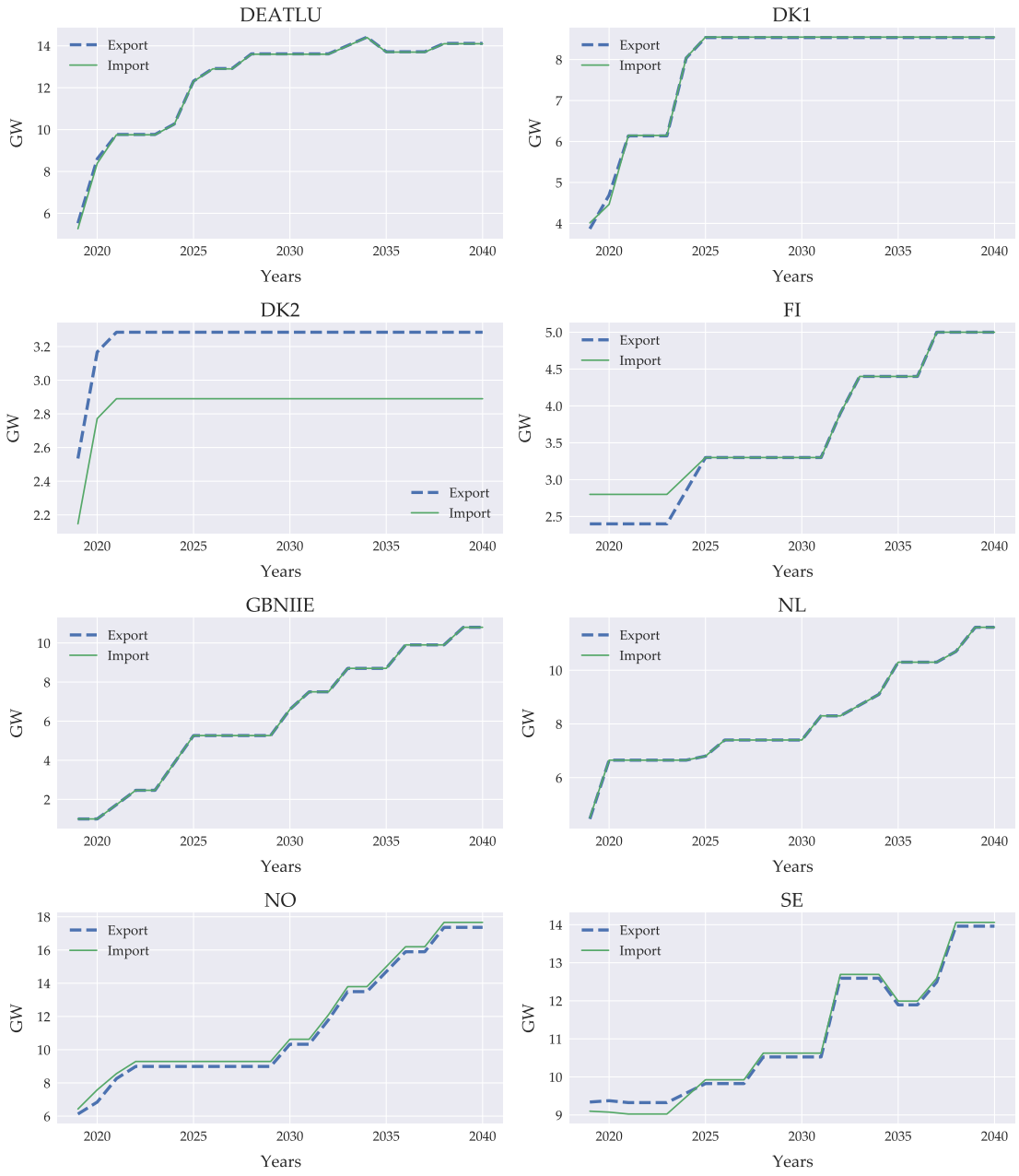
Figure A.2 plots the total transmission capacity on exports/imports for all geographical ares in the model. It shows that the trend of vastly increasing transmission capacities is common to all areas. Figure A.1 plots the aggregate information from figure A.2 on all geographical pairs for selected years.

Figure A.1: Hourly Transmission Capacities in GW



Note: The abbreviations denote the geographic areas Western Denmark (DK1), Eastern Denmark (DK2), Germany, Austria, and Luxembourg (DEATLU), Finland (FI), Great Britain, Northern Ireland, and Ireland (GBNIIE), Netherlands (NL), Norway (NO), and Sweden (SE).

Figure A.2: Each country's total access to hourly imports/exports

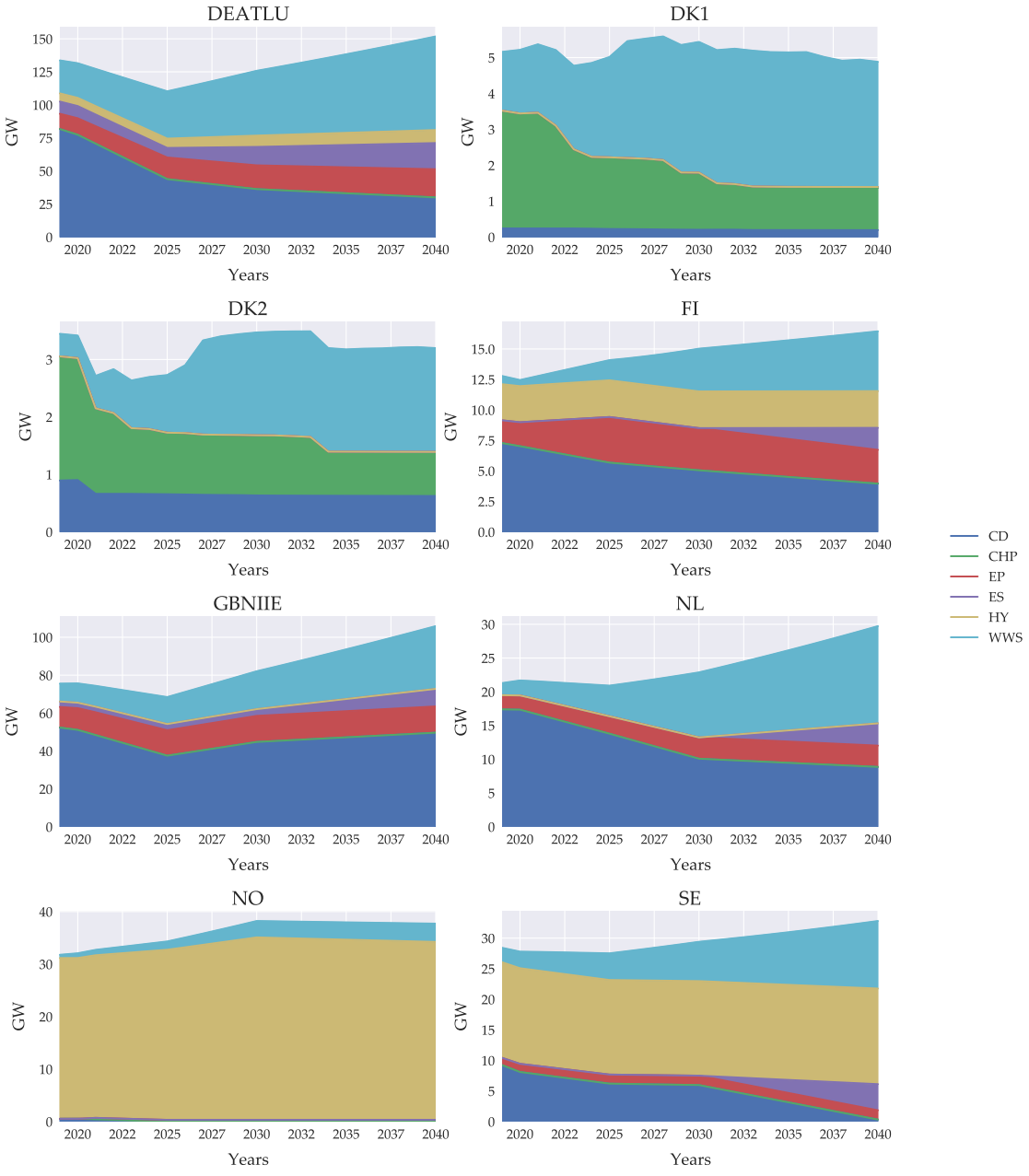


Note: The abbreviations denote the geographic areas Western Denmark (DK1), Eastern Denmark (DK2), Germany, Austria, and Luxembourg (DEATLU), Finland (FI), Great Britain, Northern Ireland, and Ireland (GBNIIE), Netherlands (NL), Norway (NO), and Sweden (SE).

Generation capacities on plant and fuel types

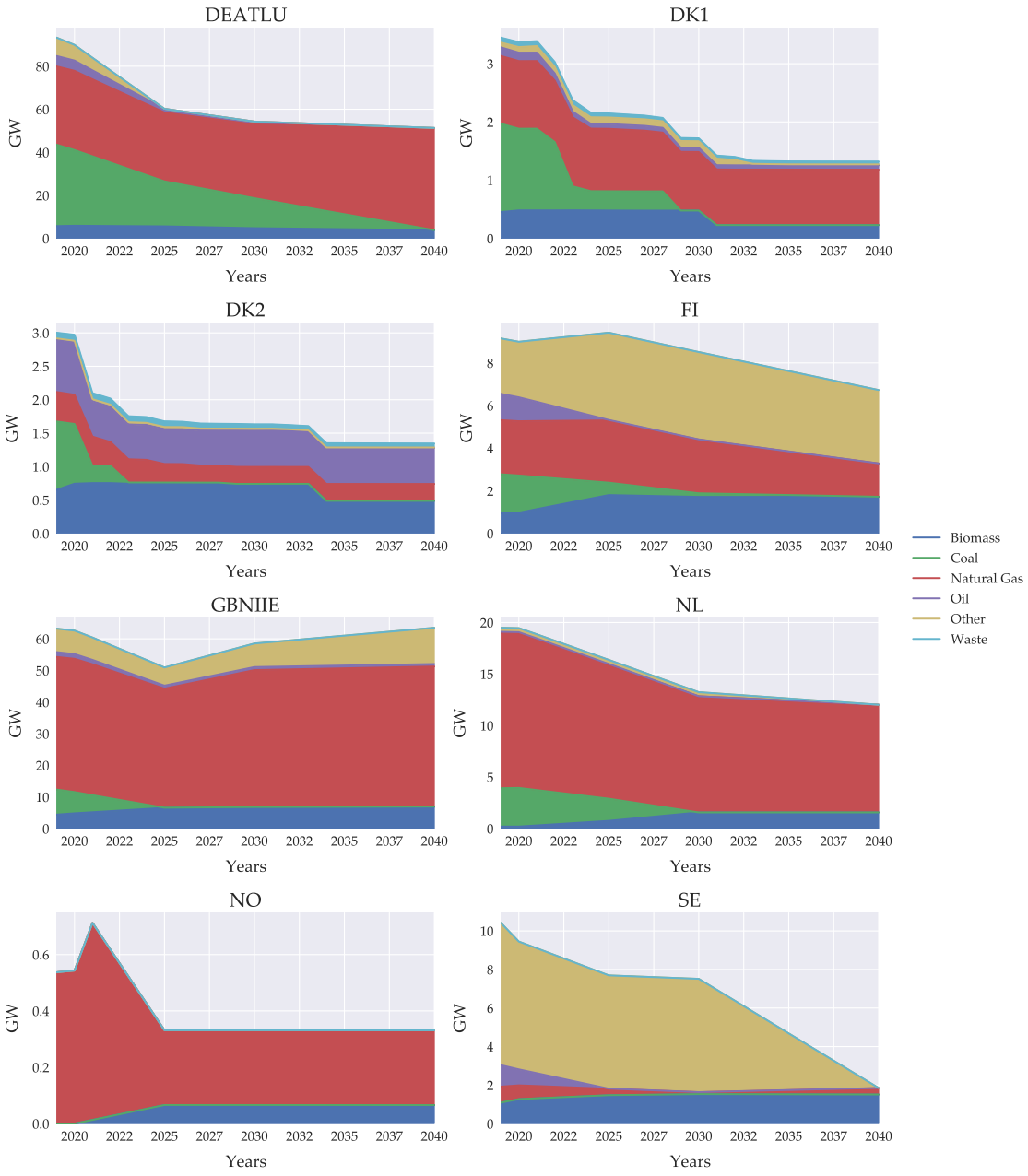
The figure A.3 shows how electricity generation capacity changes over time for all regions. Figure A.4 illustrates the same, but split onto fuel categories.

Figure A.3: Hourly average generation capacity for each country



Note: Abbreviations denote condensation plants (CD), Combined Heat and Power plants (CHP), Exogenous Production (EP), Electricity Storage (ES), Hydro plants with reservoir (HY), Wind, Water, and Sunlight (WWS).

Figure A.4: Hourly average generation capacity for each country



Note: The category "coal" covers coal and lignite, "oil" covers both fuel and gasoil, "Natural Gas" also covers synthetic natural gas, biomass covers various types of wood (pellets, chips, waste) as well as straw and peat. Finally, the category "other" covers biogas, biooil, hydrogen, and uranimum.

A.2 Baseline model results

The current appendix illustrates model results in the baseline scenario. Figure A.5 indicates that towards 2030, two price regimes more or less emerges: The Danish areas, DEATLU, and the Netherlands reach equilibrium prices of slightly below 300 DKK/MWh, while the other four regions have significantly lower prices between 175-225 DKK/MWh. Over time this price difference increases for the Danish areas.

Figure A.5: Average electricity prices

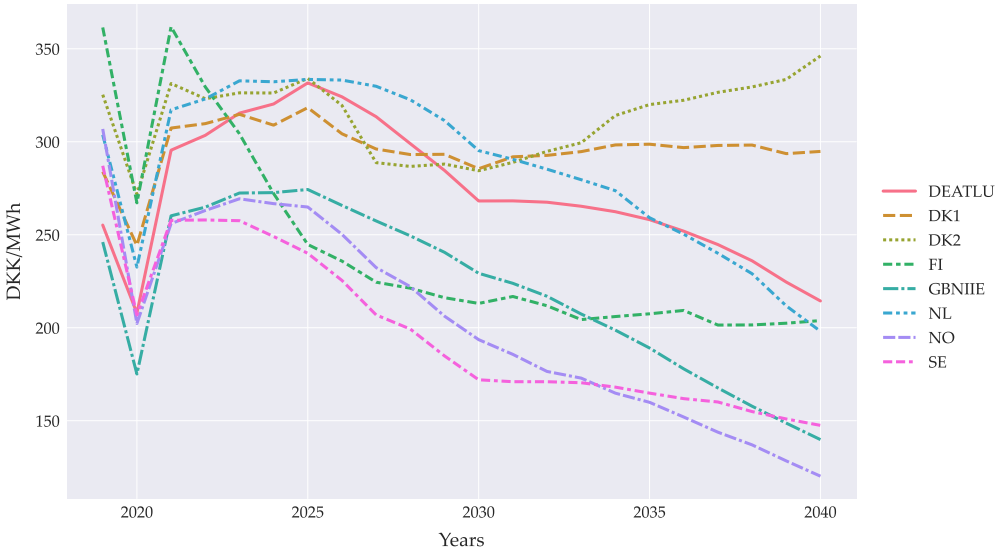


Figure A.6 indicates that net exports increase towards 2026-2027, in particular due to installation of additional wind capacity. From this point on, however, imports increase significantly. This is a consequence of the 'frozen policy' scenario applied for Danish capacity that projects less ambitious investments than neighboring countries that do not follow a similar 'frozen policy' assumption.

Figure A.6: Net exports of electricity, Denmark

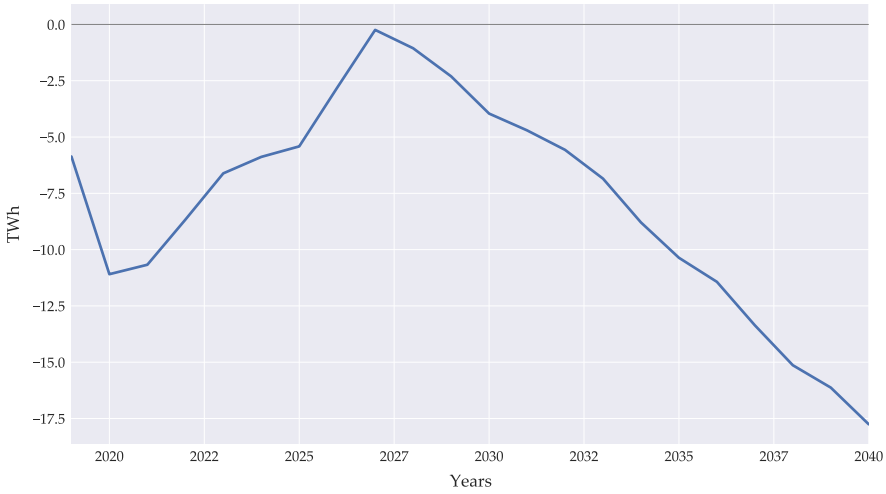


Figure A.7 shows that the consumption of all fuel decrease over time, in particular for years 2019-2030. Note that WWS energy is not included in this figure, which increases in similar magnitudes as other fuel consumption drops.

Figure A.7: Fuel use, Danish electricity and district heat systems

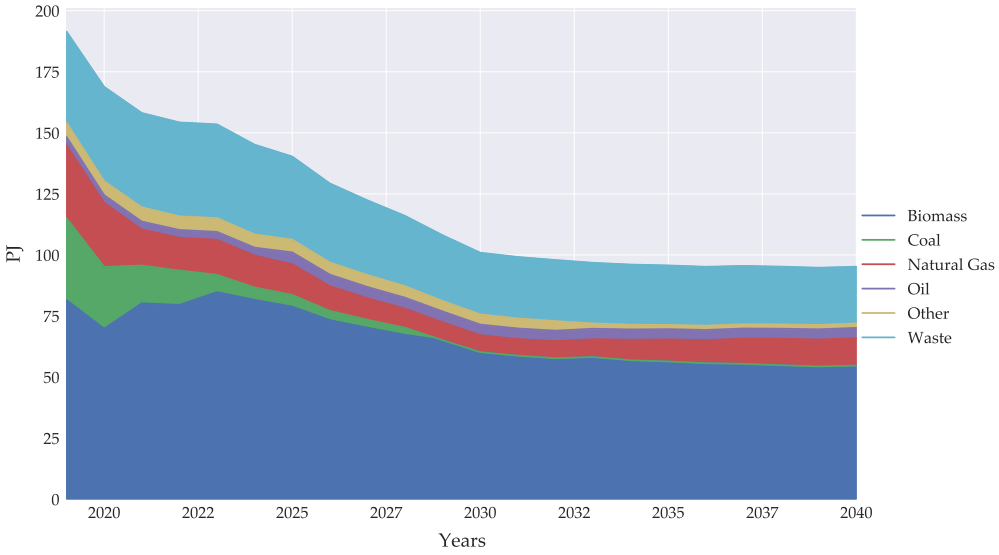


Figure A.8 shows how the price on imports/exports change over time. The decrease in export and import prices can largely be explained with the introduction of more cheap, renewable energy in neighboring countries, which also leads to lower prices in general (again, in neighboring countries), cf. figure A.5.

Figure A.8: Prices on Danish exports and imports of electricity

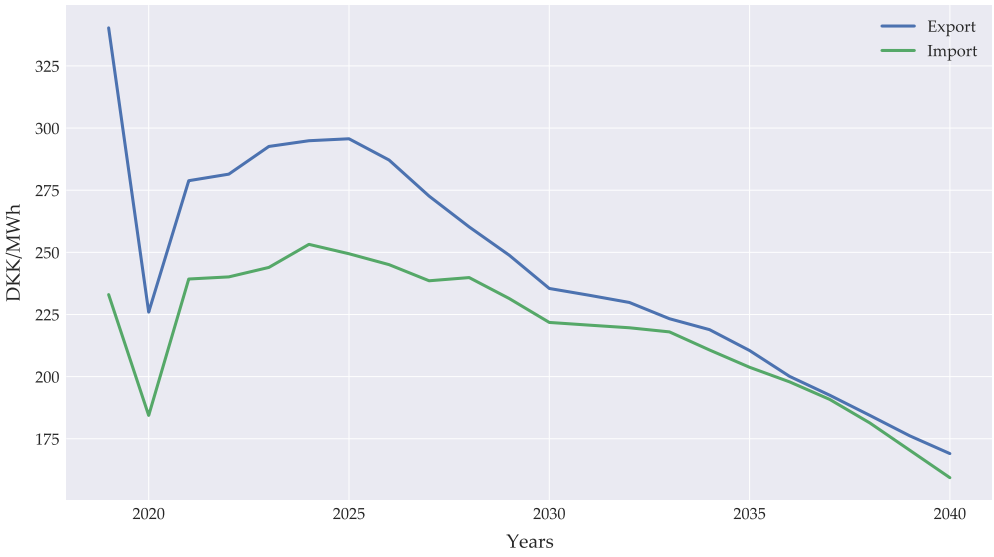
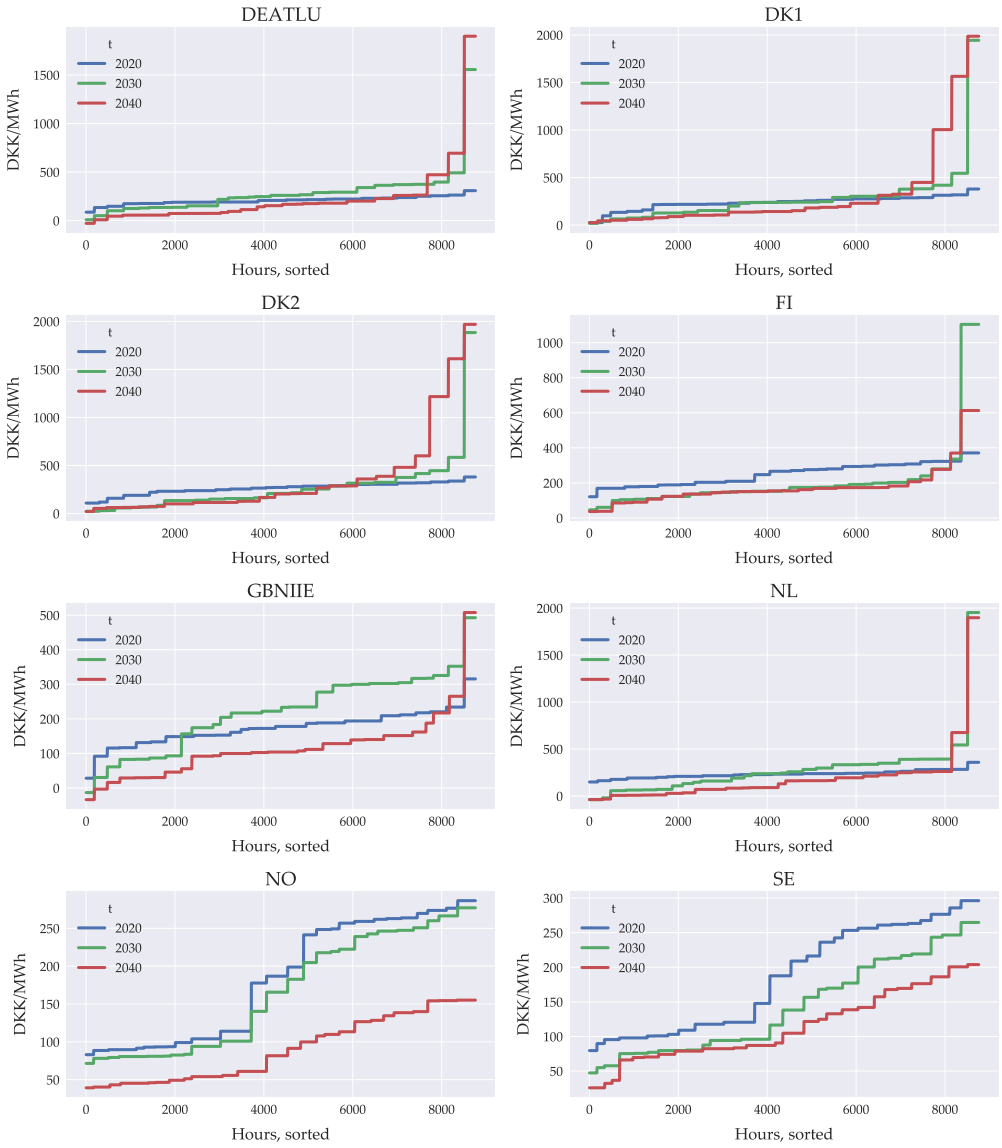


Figure A.9 shows how intra-yearly prices change over time for all price regions. We note that most countries experience far larger price fluctuations over time, with the exception of Norway and Sweden; our model predicts that they have sufficiently cheap WWS energy and storage to get lower prices in general.

Figure A.9: Intra-yearly price variation over time, all areas



B Technical information from bottom-up data

This appendix outlines relevant variables and discrete sets that define plants. This is a shortened version of the much more comprehensive technical appendices A-B in Berg and Eskildsen (2019).

Table 3: Plant specific variables from data

Variable:	Notation	
Year of commission	t_{o}	
Year of decommission	T	
Capacity, electricity production	q^E	
Capacity, heating production	q^H	
Inflow	W	(used if yearly production is known)
Storage capacity	$\bar{S}$	
Fuel efficiency	Eff	
Electricity-to heat ratio	C_b	(the exact use depends on plant type)
Transformation rate (E to H)	C_v	(used for CHP plants)
Probability of outage	$POutage$	
Tax efficiency	$TaxEff$	(used in taxation scheme)
Yearly maintenance costs	FOM	
Variable maintenance costs	VOM	
Included in CO2 quota sector	CO_2Cap	(value between 0 and 1)
Desulphurization	$Desulp$	
NO2 emission per GJ input	NO_2	
CH4 emission per GJ input	CH_4	
N2O emission per GJ input	N_2O	

Note: The *inflow* variable is used typically for plants of an intermittent nature, where the total production over a year is more or less certain. However, it is also used for hydro power plants that have an inflow of water from adjacent rivers. The variable *CO2Cap* can in principle be strictly between 0 and 1, indicating that a plant is only partly included in the quota sector.

Table 4: Discrete, plant-specific state space

Sets:	Notation	Cardinality	Examples of set elements
Fuel mix	$\mathscr{F}_F$	57	{ Avedøreværket, Coal, ... }
Electricity area	$\mathcal{G}_E$	15	{ DK-West, DK-East, Norway , ... }
District heat area	$\mathcal{G}_H$	34	{ Copenhagen, Aarhus, Natural Gas DK-west, ... }
Plant type	$\mathcal{T}$	15	{ Condensation plant, Electrical heater, ... }
Hourly variation	$\mathcal{V}$	50	{ Hornsrev2, Wind DK-West, ... }
Subsidy type	$\mathcal{S}$	46	{ Elpatron, Hornsrev2, Biogas, ... }
Bid type	$\mathcal{B}$	2	{ Marginal, Prioritized {

Note: A plant is characterized by a fuel mix type that has a cardinality of 57 (aggregated from original 127 types based). Each fuel mix type is in turn a linear combination of the basic fuel types (16 types). The bid type indicates whether or not the plant supplies electricity at its marginal costs or if it is prioritized electricity (bid is independent of market price).

C Details on the bottom-up supply

C.1 Bottom-up supply from standard plants

Interpreting smoothing as firm heterogeneity

We consider *standard plant* as defined in section 4.2. In the smoothing methodology applied in our model, the individual plant's supply function is smoothed using a flexible sigmoid function (F). For the standard plant, this entails a smoothed choice of electricity generation on the form

$$E_i = q_i F(p_i^E; c_i, \sigma). \tag{22}$$

Recall that the sigmoid function is assumed to be continuously differentiable, monotonically increasing and bounded such that $F : \mathbb{R} \to [0, 1]$.[30] Thus we can interpret the optimal dispatch function in (22) as the solution to a problem where plant i consists of a continuum of plants j (of unit mass), each with constant marginal costs of c_i^j. To see this, let $g(c_i^j; \sigma)$ denote the distribution of marginal costs and $G(c_i^j; \sigma)$ the corresponding cumulative distribution. The optimal dispatch function is then given by

$$\begin{aligned}\tilde{E}_i &= q_i \int^{p_i^E} g\left(c_i^j\right)\, dj \\ &= q_i G(p_i^E; c_i, \sigma),\end{aligned}$$

which is equivalent to the original formulation.

Value function in closed-form

Consider a plant with capacity q_i and costs c_i. Let $u_i \equiv Y_i / q_i \in [0, 1]$ denote the capacity utilization factor. The maximized profits in hour h are defined as:

$$\Pi_{i,h}^* = q_i \left[u_i^* p_h - C\left(u_i^*; c_i, \sigma\right)\right], \qquad u_i^* = F(p_h; c_i, \sigma).$$

The related first order conditions of the problem is defined as

$$p_h = \frac{\partial C}{\partial u_i} = F^{-1}\left(u_i; c_i, \sigma\right).$$

Integrating the two most-right-hand side parts we get

$$C(u_i) = \int_0^{u_i^*} F^{-1}\left(u; c_i, \sigma\right) du,$$

or similarly, using the change of variable $x = F(z)$ and the first order condition, we can write this

[30] We note that there are different conventions on what constitutes a *sigmoid* function, including what bounded interval it is defined on. An alternative convention of $\tilde{F} :\to \mathbb{R} \to [a, b]$ can easily be applied, with a few suitable transformations e.g. on the variable q_i.

as as function of the price

$$C = \int_{F^{-1}(0)}^{p_h} zf(z)dz,$$

where $f(z)$ is the density function to the cumulative distribution F. We note that with the interpretation that F represents a distribution, this can be written as:

$$C(u_i) = F(u_i)\mathbb{E}_F\left[x|X < p_h\right],$$

where $\mathbb{E}_F$ denotes the conditional expectation of x given that $x < p_h$ and x is distributed according to F. Using this, the value function can be written as

$$\Pi^*_{i,h} = q_i F(p_h; c_i, \sigma)\left(p_h - \mathbb{E}_F\left[x|X < p_h\right]\right)$$

Thus, the value function has an analytical solution if F and $\mathbb{E}_F$ has analytical representations. This is the case, for instance, with the normal distribution.

C.2 Bottom-up supply from combined power and heat plants

C.2.1 Deriving bottom-up supply from combined power and heat plants

As mentioned in the main section, there are three different type of CHP plants included in the data. In general the CHP plants are so-called joint production technologies. The implication is that the plant not only faces a problem of choosing its inputs correctly, it also faces the problem of optimally allocation this to either production of electricity and/or heat outputs.

The traditional way of modelling CHP plants in the bottom-up literature, is to adopt the following assumptions for production technology:

Assumption 5 (CHP bottom-up assumptions)**.** *For all types of CHP plants:*

i. Plants produce a composite good $\mathcal{Q}$ using a constant returns to scale technology (typically Leontief) of inputs $x \in \mathbb{R}^m_+$. This results in a constant marginal cost c_i.

ii. The scale of production of the composite good $\mathcal{Q}$ is bounded on $[0, 1]$, representing degree of capacity (q_i) that is utilized.

iii. Given the scale of production ($\mathcal{Q}(x)$), plants use a linear transformation function to (costlessly) produce electricity and heat:

$$q_i\mathcal{Q}(x) = \gamma^E_i E_i + \gamma^H_i H_i. \tag{23}$$

iv. The linear transformation function is bounded by minimum co-production constraints on

electricity (C_i^E) *or heat* (C_i^H)*:*

$$E_i \geq H_i C_i^E, \qquad\qquad H_i \geq E_i C_i^H. \tag{24}$$

Under assumption 5, consider the following two step maximization procedure: Given the choice of inputs x, the plant produces $\mathcal{Q}(x) \in [0,1]$. Given this, what is the optimal choice of output split? The marginal effect on profits from increasing E marginally by lowering H is then given by

$$\tilde{p}_E - \frac{\gamma_i^E}{\gamma_i^H}\tilde{p}_H, \qquad\qquad \tilde{p}_i^Y = p_Y - T_Y^i(\mathbf{t}), \text{ for } Y \in \{E,H\}$$

as long as this level is feasible, i.e. within the bounds in (24). Note that the relevant prices $\tilde{p}_E, \tilde{p}_H$ are output-prices corrected for potential *output*-specific taxation. With this linearity with bounds, the optimal output-split function is simply given by

$$H_i^* = \begin{cases} \dfrac{q_i \mathcal{Q}(x)}{\gamma_i^H + \gamma^E C_i^E}, & \tilde{p}_E < \dfrac{\gamma_i^E}{\gamma_i^H}\tilde{p}_H \\ \dfrac{q_i \mathcal{Q}(x)}{\gamma_i^H + \gamma_i^E / C_i^H}, & \tilde{p}_E \geq \dfrac{\gamma_i^E}{\gamma_i^H}\tilde{p}_H \end{cases} \qquad E_i^* = \begin{cases} \dfrac{q_i \mathcal{Q}(x)}{\gamma_i^E + \gamma^H C_i^H}, & \tilde{p}_E \geq \dfrac{\gamma_i^E}{\gamma_i^H}\tilde{p}_H \\ \dfrac{q_i \mathcal{Q}(x)}{\gamma_i^E + \gamma_i^H / C_i^E}, & \tilde{p}_E < \dfrac{\gamma_i^E}{\gamma_i^H}\tilde{p}_H \end{cases}$$

Given these functions, the optimal choice of $\mathcal{Q}$ is given by

$$\max_{\mathcal{Q} \in [0,1]} \tilde{p}_E E_i^* + \tilde{p}_H H_i^* - c_i \mathcal{Q} q_i.$$

This is similarly linear but bounded in $\mathcal{Q}$, thus again yielding a function on the form

$$\mathcal{Q}^* = \begin{cases} 1, & p_{EH} \geq c_i \\ 0, & p_{EH} < c_i \end{cases} \tag{25}$$

where p_{EH} depends on the relative prices as

$$p_{EH} = \begin{cases} \dfrac{\tilde{p}_H}{\gamma_i^H + \gamma_i^E C_i^E} + \dfrac{\tilde{p}_E}{\gamma_i^E + \gamma_i^H / C_i^E}, & \tilde{p}_E < \dfrac{\gamma_i^E}{\gamma_i^H}\tilde{p}_H \\ \dfrac{\tilde{p}_H}{\gamma_i^H + \gamma_i^E / C_i^H} + \dfrac{\tilde{p}_E}{\gamma_i^E + \gamma_i^H C_i^H}, & \tilde{p}_E \geq \dfrac{\gamma_i^E}{\gamma_i^H}\tilde{p}_H \end{cases} \tag{26}$$

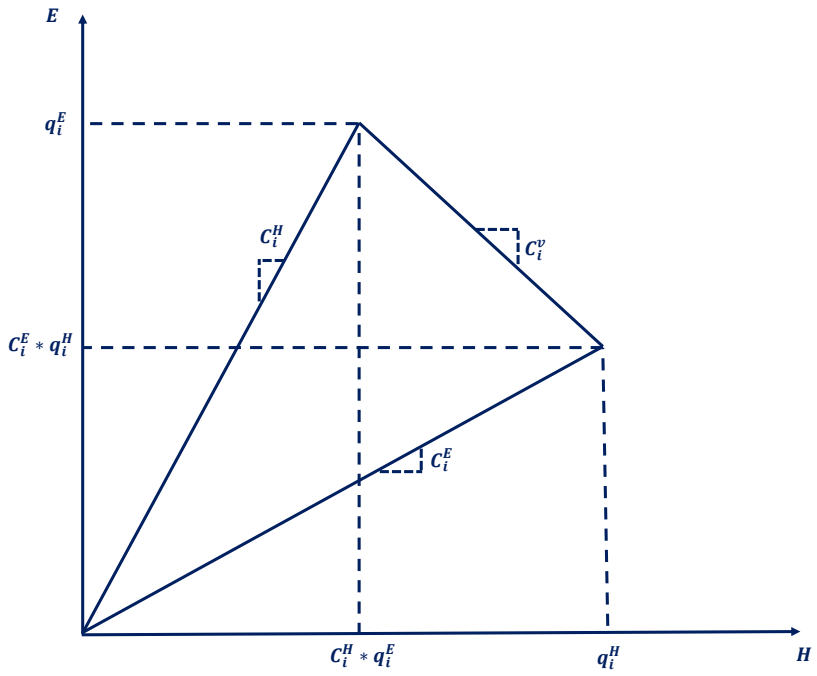
To see how this corresponds to the solution in the main text in equations (6)-(7) consider figure C.1 below copied from the main text for convenience.

Define the transformation function in (23) as

$$q_i^E \mathcal{Q} = E_i + C_i^v H_i,$$

using that $q_i = q_i^E, \gamma_i^E = 1, \gamma_i^H = C_i^v$. The minimum co-production constraints in (24) are defined as in the main text. It is straightforward to verify that the solution in (25)-(26) can be written on the form presented in the main text in (6)-(7).

Figure C.1: Production possibility set for CHP plants



C.2.2 A normalized constant elasticity of transformation assumption leads to flexible sigmoid share functions

The bottom-up approach outlined in the previous section, shows how the solution essentially boils down to a combination of two corner solutions: (1) Is it profitable to use the plant or not (scale decision)? And (2) Is it most profitable to maximize production of electricity or heat? One way to approach this is to approximate the discrete $\{0, 1\}$ choices with a form of flexible sigmoid function. Here we show that a simple change of the bottom-up assumptions implies replacing the discrete choice with a type of flexible sigmoid function.

Assumption 6 (CHP with normalized constant elasticity of transformation (NCET) assumptions)**.** *Consider the four assumptions in assumption 5. Replace the linearly bounded transformation function in iii.-iv. with a normalized constant elasticity transformation:*

$$q_i \mathcal{Q} = \left[\left(\theta_E^i\right)^{\frac{\sigma-1}{\sigma}} \left(E_i - \underline{E}\right)^{\frac{\sigma-1}{\sigma}} + \left(\theta_H^i\right)^{\frac{\sigma-1}{\sigma}} \left(H_i - \underline{H}\right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \qquad \sigma < -1. \tag{27}$$

One set of coefficients that nests the linear assumptions in assumption 5 for $\sigma \to -\infty$ is:

$$\theta_E = \frac{q_i^E}{q_i^E - q_i^H C_i^E}, \qquad \theta_H = \frac{q_i^E}{q_i^H - q_i^E C_i^H}, \qquad q_i = q_i^E$$
$$\underline{E} = q_i^H \mathcal{Q} C_i^E, \qquad \underline{H} = q_i^E \mathcal{Q} C_i^H.$$

As in the bottom-up case, we consider the optimization problem of the CHP plant in two steps. First, assume that $\mathcal{Q}(x)$ is given. The problem of optimally choosing production of E, H is under

assumption 6 given by:

$$\max_{E_i,H_i} \ \tilde{p}_E E_i + \tilde{p}_H^i H_i,$$
$$\text{s.t. } q_i^E \mathcal{Q} = \left[\left(\frac{q_i^E}{q_i^E - q_i^H C_i^E} \right)^{\frac{\sigma-1}{\sigma}} \left(E_i - q_i^H \mathcal{Q} C_i^E \right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{q_i^H}{q_i^H - q_i^E C_i^H} \right)^{\frac{\sigma-1}{\sigma}} \left(H_i - q_i^E \mathcal{Q} C_i^H \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}.$$

This formulation gives a marginal rate of transformation of

$$-\frac{\partial E}{\partial H} = \left(\frac{\theta_H}{\theta_E} \right)^{\frac{\sigma-1}{\sigma}} \left(\frac{E_i - \underline{E}}{H_i - \underline{H}} \right)^{\frac{1}{\sigma}}.$$

Thus the optimal output-split is given by

$$\tilde{p}_E \left| \frac{\partial E}{\partial H} \right| = \tilde{p}_H.$$

After some standard CET manipulations we get the solution

$$H_i^* = \underline{H} + \frac{q_i^E \mathcal{Q}}{\theta_H} \frac{1}{\left[1 + \left(\frac{\tilde{p}_H \theta_E}{\tilde{p}_E \theta_H} \right)^{\sigma-1} \right]^{\frac{\sigma}{\sigma-1}}}.$$
$$E_i^* = \underline{E} + \frac{q_i^E \mathcal{Q}}{\theta_E} \frac{1}{\left[1 + \left(\frac{\tilde{p}_H \theta_E}{\tilde{p}_E \theta_H} \right)^{1-\sigma} \right]^{\frac{\sigma}{\sigma-1}}}.$$

Substituting back for $\theta_E, \theta_H, \underline{E}, \underline{H}$ this yields:

$$H_i^* = \mathcal{Q} \left[q_i^H \frac{1}{\left[1 + \left(\frac{\tilde{p}_H (q_i^H - q_i^E C_i^H)}{\tilde{p}_E (q_i^E - q_i^H C_i^E)} \right)^{\sigma-1} \right]^{\frac{\sigma}{\sigma-1}}} + q_i^E C_i^H \left(1 - \frac{1}{\left[1 + \left(\frac{\tilde{p}_H (q_i^H - q_i^E C_i^H)}{\tilde{p}_E (q_i^E - q_i^H C_i^E)} \right)^{\sigma-1} \right]^{\frac{\sigma}{\sigma-1}}} \right) \right] \tag{28}$$

$$E_i^* = \mathcal{Q} \left[q_E^H \frac{1}{\left[1 + \left(\frac{\tilde{p}_H (q_i^H - q_i^E C_i^H)}{\tilde{p}_E (q_i^E - q_i^H C_i^E)} \right)^{1-\sigma} \right]^{\frac{\sigma}{\sigma-1}}} + q_i^H C_i^E \left(1 - \frac{1}{\left[1 + \left(\frac{\tilde{p}_H (q_i^H - q_i^E C_i^H)}{\tilde{p}_E (q_i^E - q_i^H C_i^E)} \right)^{1-\sigma} \right]^{\frac{\sigma}{\sigma-1}}} \right) \right] \tag{29}$$

From the solution in (28)-(29) we note the following:

- Both H_i^* and E_i^* is linear in the scale decision $\mathcal{Q}$. Thus the optimal choice of $\mathcal{Q}$ is still a corner decision, that depends on whether or not the plant is profitable when active. Replacing the assumption that $\mathcal{Q}$ is produced with constant marginal costs, with the assumption of marginal costs outlined for standard plants in assumption 1, the optimal solution of $\mathcal{Q}$ becomes a flexible sigmoid type as well.

- Consider the two functions that depend on the relative prices $\tilde{p}_H/\tilde{p}_E$. Both of these are flexible sigmoid functions that maps smoothly onto $[0,1]$.[31] Thus we can write the functions as

$$H_i^* = \mathcal{Q}\left[q_i^H F_i^H(\tilde{p}_H/\tilde{p}_E;\sigma) + q_i^E C_i^H\left(1 - F_i^H(\tilde{p}_H/\tilde{p}_E;\sigma)\right)\right]$$
$$E_i^* = \mathcal{Q}\left[q_i^E F_i^E(\tilde{p}_H/\tilde{p}_E;\sigma) + q_i^H C_i^E\left(1 - F_i^E(\tilde{p}_H/\tilde{p}_E;\sigma)\right)\right].$$

  Note that in the limit of $\sigma \to -\infty$ this yields

$$\lim_{\sigma\to-\infty} F_i^H = \begin{cases} 1, & \tilde{p}_H(q_i^H - q_i^E C_i^H) > \tilde{p}_E(q_i^E - q_i^H C_i^E) \\ 0, & \tilde{p}_H(q_i^H - q_i^E C_i^H) < \tilde{p}_E(q_i^E - q_i^H C_i^E) \end{cases}$$

- Finally, note that by construction, the minimum co-production constraints, capacities and linear transformation rate C_i^v relates as follows:

$$C_i^v\left(q_i^H - C_i^H q_i^E\right) = q_i^E - C_i^E q_i^H.$$

  Using this, note that the threshold for when this share function jumps from 0 to 1, corresponds to the same limit as in the bottom-up case, i.e. when

$$\lim_{\sigma\to-\infty} F_i^H = \begin{cases} 1, & \tilde{p}_H > \tilde{p}_E C_i^v \\ 0, & \tilde{p}_H < \tilde{p}_E C_i^v \end{cases}$$

D Trade and equilibrium concepts

D.1 Derivation of the smooth net exports function assuming increasing marginal costs of trade

Consider the bottom-up net export function in the main section 4.3. Recall that we based the net export function on three assumptions: (i) Homogeneity of electricity across geographic areas, (ii) competitive markets, and (iii) trade constraints from transmission lines. Note that instead of formulating this in terms of $NX_{i,j}$, we can interpret this as a trade cost function $\tilde{C}_{i,j}$ of trading between area i and j. Thus we can formulate a trade cost function that is equivalent to the net export formulation in the main section:

$$\tilde{C}_{i,j} = \begin{cases} \infty, & \text{if } NX_{i,j} > T_{i,j}, \text{ or } NX_{i,j} < -T_{j,i} \\ 0, & \text{else} \end{cases}. \tag{30}$$

[31] In section 4.1 we defined the flexible sigmoid function, as one with a parameter $\sigma > 0$, where the smoothing 'dies out' when $\sigma \to 0$. Technically, the functions presented here needs a simple transformation before it obeys this criterion, as the smoothing 'dies out' here when $\sigma \to -\infty$. Letting $\gamma \equiv -1/\sigma$, this is indeed the case.

Assume now instead that trade costs are not exactly zero within the bounds $[-T_{j,i}, T_{i,j}]$, but rather gradually approached these limits. For instance, consider the class of *flexible sigmoid functions* defined in section 4.1. Note that with the definition $F : \mathbb{R} \to [0,1]$ where F is monotonically increasing and continuously differentiable, the inverse function $g \equiv F^{-1} : [0,1] \to \mathbb{R}$ exists and is similarly continuously differentiable on $[0,1]$.[32] Furthermore, we know that

$$\begin{aligned}\frac{\partial g}{\partial x}(x;c,\sigma) &> 0,\\ \lim_{x\to 0^+} g(x;c,\sigma) &= -\infty\\ \lim_{x\to 1^-} g(x;c,\sigma) &= \infty.\end{aligned}$$

With this in place consider the following equilibrium concept:

Problem 1 (Cost-minimizing equilibrium)**.**
Consider the electricity and heat markets in hour h and year t. Let $(\mathbf{p}_{g_E}, \mathbf{p}_{g_H})$ denote the vectors of all hourly prices in a year in electricity area g_E and heat area g_H. Let $E^d_{g_E}$ denote the demand of electricity in area g_E and similarly for H^d. Finally let $(\mathcal{I}_{g_E}, \mathcal{I}_{g_H})$ denote the set of individual plants producing electricity or heat in the relevant area (g_E, g_H). The trade-cost minimizing equilibrium is defined from the problem:

$$\min \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} C_{k,j}(NX_{k,j}) \tag{31a}$$

$$s.t. \quad \sum_{g_E \in \mathcal{G}_E} E^d(\mathbf{p}_{g_E}) = \sum_{g_E \in \mathcal{G}_E} \sum_{i \in \mathcal{I}_{g_E}} E_i\left(\mathbf{p}_{g_E}, \mathbf{p}_{g_H}\right) \tag{31b}$$

$$E^d(\mathbf{p}_{g_E}) = \sum_{i \in \mathcal{I}_{g_E}} E_i\left(\mathbf{p}_{g_E}, \mathbf{p}_{g_H}\right) - \sum_{g_k \in (\mathcal{G}_E \setminus g_E)} NX_{g_E, g_K}, \qquad \forall g_E \in \mathcal{G}_E \tag{31c}$$

$$H^d(\mathbf{p}_{g_H}) = \sum_{i \in \mathcal{I}_{g_H}} H_i\left(\mathbf{p}_{g_E}, \mathbf{p}_{g_H}\right), \qquad \forall g_H \in \mathcal{G}_H, \tag{31d}$$

where the cost function is defined such that

$$\frac{\partial C_{k,j}}{\partial NX_{k,j}} = g\left(\frac{NX_{i,j} + T_{j,i}}{T_{i,j} + T_{j,i}};\ \sigma\right) \tag{32}$$

The problem in 1 minimizes trade costs, subject to the sum of supply equals sum of demand (31b) and where each electricity and heat area is in equilibrium (31c)-(31d). Finally, the assumption that marginal costs follow the function g (defined as the inverse of any flexible sigmoid function F) ensures that the trade cost function is strictly convex in $NX_{k,j}$.

This leads us to the following proposition:

Proposition 2 (Trade-cost minimizing equilibrium)**.**
Consider the equilibrium defined by the solution to problem 1. Assume that the sigmoid function F

[32] The differentiable part follows from the inverse function theorem: F is strictly monotonically increasing for $\sigma > 0$ implying that $f'(p) \neq 0$ for all $p \in \mathbb{R}$.

and its inverse g both have analytical representations. Let n denote the number of electricity areas $g_E \in \mathcal{G}_E$. Let $m = \sum_{i=1}^{n-1} i$. Let $\tilde{p}_E \in \mathbb{R}^{n-1}$ denote the vector of all electricity prices in one hour except one. Let $\mathbf{NX} \in \mathbb{R}^m$ denote the vector of all net export terms in one hour. We then have:

i. The global minimum of problem 1 exists.

ii. The global minimum always obeys transmission line constraints.

iii. There exists a mapping $\tilde{F}^p : \mathbb{R}^{n-1} \to \mathbb{R}^m$ from a vector of country specific electricity prices $(\tilde{p}_E)$ to a vector of all net export elements $\mathbf{NX}$.

iv. The mapping $\tilde{F}^p$ is unique (injective) and has an analytical representation. In particular, if the marginal trade cost function is the inverse of a flexible sigmoid function as in problem 1, the trade-cost minimizing solution is on the the form

$$NX_{i,j} = -T_{j,i} + \left(T_{i,j} + T_{j,i}\right) F\left(p_{E,j} - p_{E,i}; 0, \sigma\right).$$

Initially, note that the sum of the n equilibrium constraints in (31c) corresponds to the total equilibrium condition in (31b), thus we can without loss of generality disregard the n'th equilibrium constraint from the minimization problem. Letting λ_E^i denote the shadow-value of the i'th equilibrium constraint, the first order conditions for problem 1 is thus given by

$$\begin{aligned} \text{For } j < n: \quad & \frac{\partial C_{k,j}}{\partial NX_{k,j}} = \lambda_E^j - \lambda_E^k, \\ \text{For } j = n: \quad & \frac{\partial C_{k,j}}{\partial NX_{k,j}} = -\lambda_E^k. \end{aligned}$$

Letting $\partial\mathbf{C}/\partial\mathbf{NX}$ denote the marginal trade cost function applied on each element of a vector input, and λ denote the vector shadow values, the first order conditions can be summed up on the form

$$\frac{\partial \mathbf{C}}{\partial \mathbf{NX}} = A\lambda,$$

where A is an $(m \times n-1)$ coefficient matrix. Similarly, the equilibrium constraints in (31c) can be written as a linear system:

$$B\mathbf{NX} = \zeta,$$

where B is an $(n-1 \times m)$ coefficient matrix, and ζ is an $n-1$ vector of constants (given prices). Both A and B are of rank $n-1$. In Berg and Eskildsen (2019) we show that the following holds:

i. The objective function and constraints are twice continuously differentiable.

ii. Gradients of constraints are linearly independent, i.e. the Jacobian matrix B is of full (row)

rank. In particular we have that the null space of B

$$Null(B) = span\Big\{(Z)\Big\},$$

where Z is a basis matrix of dimensions $m \times (m - (n - 1))$.

iii. The matrix $\mathcal{H} \equiv Z' \left(\nabla^2_{\mathbf{NX},\mathbf{NX}}\mathcal{L}\right) Z$ is positive definite. In particular $\nabla^2_{\mathbf{NX},\mathbf{NX}}\mathcal{L}$ is a $(m \times m)$ diagonal matrix with elements $\lambda_i > 0$. Thus $\mathcal{H}$ is defined in its diagonalized form, with eigenvalues λ_i and eigenvectors in Z.

This ensures that the first order conditions are necessary and sufficient for the global minimum. To show that this minimum always obeys the transmission line constraints, we refer to appendix H of Berg and Eskildsen (2019). Next, note that with competitive markets, the marginal trade cost will act as a price-wedge between electricity areas, such that:

$$p_i = p_j - \frac{\partial C_{i,j}}{\partial NX_{i,j}}, \qquad \forall (i,j) : T_{i,j} > 0. \tag{33}$$

Using this in the first order conditions we have:

$$NX_{i,j} = \left(\frac{\partial C_{i,j}}{\partial NX_{i,j}}\right)^{-1} (p_j - p_i).$$

Using that F and g are each other's inverse functions, this condition yields the net export function from the main section (equation (14)).

D.2 Alternatives to modelling trade in zero event cases

Consider the case where trade capacities fluctuate from hour to hour, in such a way that transmission line capacities are sometimes zero in one direction, but not in the other one: $T_{i,j,h}T_{j,i,h} = 0$ and $T_{i,j,h} + T_{j,i,h} > 0$. In the main section 4.3 we argue that one caveat of the proposed method is that net exports can go in the 'wrong direction': If imports are shut down, net exports can be positive even though foreign prices are lower than domestic. One immediate way of alleviating this problem is to lower the smoothing parameter σ towards zero. However, in the instances where σ either needs to remain relatively large due to computational constraints or due to the fact that trade in electricity as a function of prices is estimated to be relatively smooth, the following proposes a simple alternative.

Consider the two-parameter flexible sigmoid function g defined as the product of two sigmoid functions:[33]

$$g(x; \gamma, \sigma_1, \sigma_2) = F_1(x; \gamma, \sigma_1) F_2(x; \gamma, \sigma_2). \tag{34}$$

Note that the function g immediately inherits the characteristics outlined in definition 1:

[33] Note that by induction the product of any number of flexible sigmoid functions is a flexible sigmoid function.

i. g is monotonically strictly increasing and continuously differentiable for $\sigma_1, \sigma_2 > 0$,

ii. g approaches the discrete-choice function in the limit:

$$\lim_{\sigma_1,\sigma_2 \to 0} g(x;\gamma,\sigma_1,\sigma_2) = \begin{cases} 0, & x < \gamma, \\ 1, & x \geq \gamma \end{cases}.$$

iii. g's supremum (infimum) is 1 (0) and attained when $x \to \infty$ $(-\infty)$.

In cases where one of the transmission lines are shut down, we can apply the two-parameter flexible sigmoid function, where one parameter σ_1 is relatively low, in order to keep net exports close to zero, and the other σ_2 is relatively large in order to keep the function smooth.

E Short run demand for energy goods

E.1 Demand response and numerical stability

To increase numerical stability of the model, it can be beneficial to include a small, but price-responsive part in the demand function. Instead of the formulation in the main section, we can include a simple linear term such that hourly demand is defined as

$$x_h^d(\boldsymbol{p}^x) = x_t^d \frac{\gamma(p_t - p_h) + z_h\left[\phi + (1-\phi)L_h\left(\boldsymbol{p}^x\right)\right]}{\sum_j \gamma(p_t - p_h) + z_j\left[\phi + (1-\phi)L_j\left(\boldsymbol{p}^x\right)\right]}, \tag{35}$$

where $\gamma > 0$ is a small, but positive price response. We note that in cases where the model builds on relatively few, aggregated plants, there might be intervals of prices, where the supply function is completely flat. For a gradient-based solver, this may lead to very slow numerical convergence. Including a small price response on the demand side ensures that excess demand is price-responsive for all price levels.

E.2 Estimation of habits in electricity demand in Denmark

Assuming $\phi = 1$ and allowing habits $z_{h,t}$ to depend on year t as well, implies that (18) simplifies to

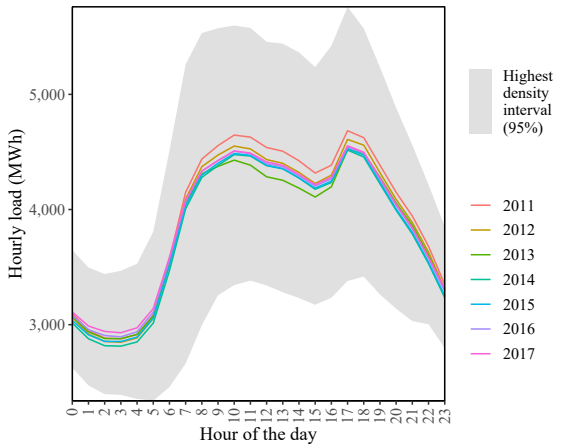
$$x_h = x_t z_{h,t}, \tag{36}$$

The specification in (36) assumes that the hourly load is perfectly inelastic, where $z_{h,t}$ can be interpreted as purely habitual consumption patterns repeating itself as cycles (seasonality): consumers have a preference for allocating their consumption at specific hours of the year due to e.g. labor/leisure time and holiday effects[34].

[34] Remember we only assume a perfectly inelastic demand on the very short-run (hourly frequency), since the yearly consumption level x_t is price dependent in the top-down model

To illustrate the idea of $z_{h,t}$ the average daily consumption cycle of electricity per year is plotted in figure E.1. The average daily cycle is almost identical across year and the 95% highest density interval suggests that the daily cycle is also persistent. Furthermore, the pattern is very much in line with expectation that consumers demand electricity in conjunction with their daily task without consideration for prices. Finally, note from (36) that seasonalities are assumed to be

Figure E.1: Average daily electricity consumption cycle for Denmark



Note: Hourly load for DK-West and DK-East has been aggregated.
Source: Energy Data Service (2018)

multiplicative and that $z_{h,t}$ is allowed to depend on year t. One reason for this is e.g. that the seventh hour of a Monday can fall at different hours, h, within year t assuming demand is the same in this state over time. It should be noted that such a distinction seems unnecessary in the bottom-up model for the forecast to be realistic, however, it might be important in the estimation of the demand.

In a regression framework we evaluate two different assumptions on the structure of the irregular component, $\epsilon_{h,t}$:

$$e_{h,t} = E_t z_{h,t} + \epsilon^A_{h,t} \tag{37}$$

$$e_{h,t} = E_t z_{h,t} \epsilon^M_{h,t} \tag{38}$$

In (37) the error term is assumed additive and in (38) the error term is multiplicative. From equation (37) and (38) we get two regression models:

$$y^1_{h,t} = z_{h,t} + \eta^1_{h,t}, \qquad y^1_{h,t} \equiv \frac{e_{h,t}}{E_t}, \qquad \eta^1_{h,t} \equiv \frac{\epsilon^A_{h,t}}{E_t} \qquad \text{(Model 1)}$$

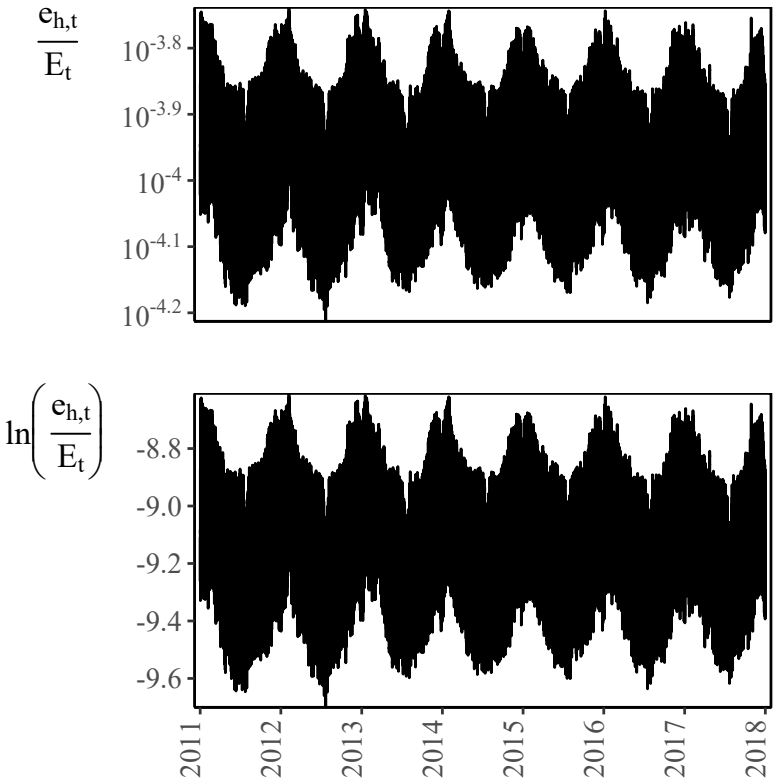
$$y^2_{h,t} = \ln z_{h,t} + \eta^2_{h,t}, \qquad y^2_{h,t} \equiv \ln\left(\frac{e_{h,t}}{E_t}\right), \qquad \eta^2_{h,t} \equiv \ln \epsilon^M_{h,t} \qquad \text{(Model 2)}$$

Model 1 is a run off of the additive model in (37) and Model 2 follows from the multiplicative model in (38). Model 1 and Model 2 have the advantage of estimating the function $z_{h,t}$ directly[35].

[35]We disregarded the option of estimating the logarithmic level of the hourly electricity load given by $\ln e_{h,t} =$

The hourly load, $e_{h,t}$, is collected from Energy Data Service, 2018 for the period 2011-2017. The time-series of the transformed dependent variables in Model 1 and Model 2 are illustrated in figure E.2.

Figure E.2: Time-series data for Model 1 and Model 2



Note: Hourly load for DK-West and DK-East has been aggregated.
Source: Energy Data Service (2018)

Remembering that we want to evaluate how well (36) explains or rather predicts electricity demand, we split the data into a training and a test set. The training data consists of 2011-2016. We test the forecast on data from 2017. The benchmark for choosing between models is the mean-squared-error (MSE) on the test set, defined as

$$\text{MSE}_i = \frac{1}{8760}\sum_{h}^{8760}\left(e_{h,t} - \tilde{y}^i_{h,t}\right)^2, \qquad i = \{1,2,3\} \tag{39}$$

$$\tilde{y}^1_{h,t} \equiv E_t\hat{y}^1_{h,t} = \hat{e}_{h,t} \tag{40}$$

$$\tilde{y}^3_{h,t} \equiv E_t exp\left(\hat{y}^2_{h,t}\right) = \hat{e}_{h,t}, \tag{41}$$

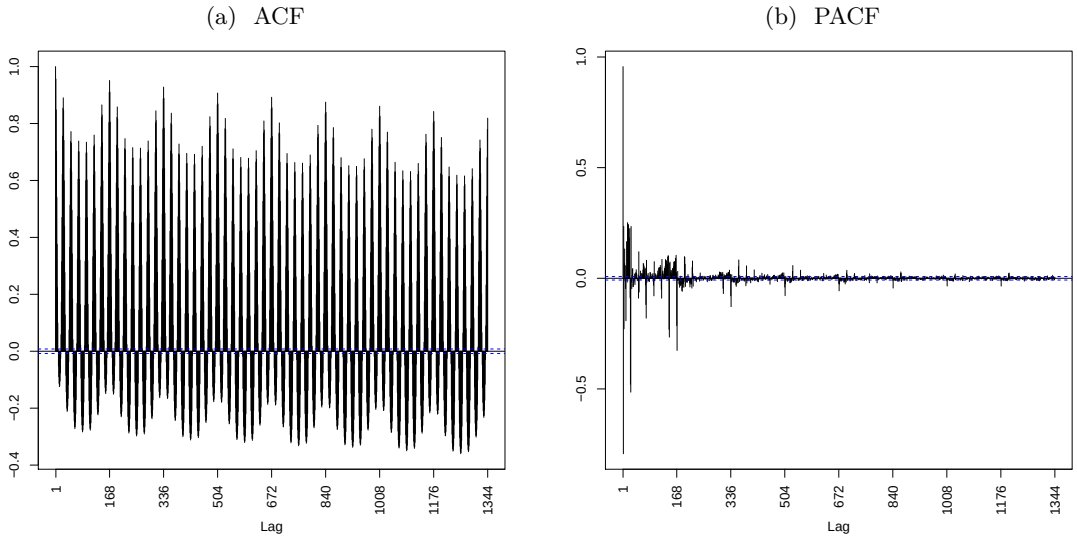
where $t = 2017$. The transformations in (40) and (41) ensures that the MSE is measured in the units of the hourly load $e_{h,t}$ for both models.

Before estimation we have computed the auto-correlation function (ACF) and the partial auto-

$\ln E_t + \ln z_{h,t} + \ln \epsilon^M_{h,t}$. We are only interested in the allocation shares captured by $z_{h,t}$ as the yearly load level is coming from the top-down model in the integrated model.

correlation function (PACF) of the transformed time-series to examine stationarity. These are illustrate in figure E.3 and suggest that the process is stationary although with considerable seasonalities.

Figure E.3: ACF and PACF for $y^2_{h,t} = e_{h,t}/E_t$



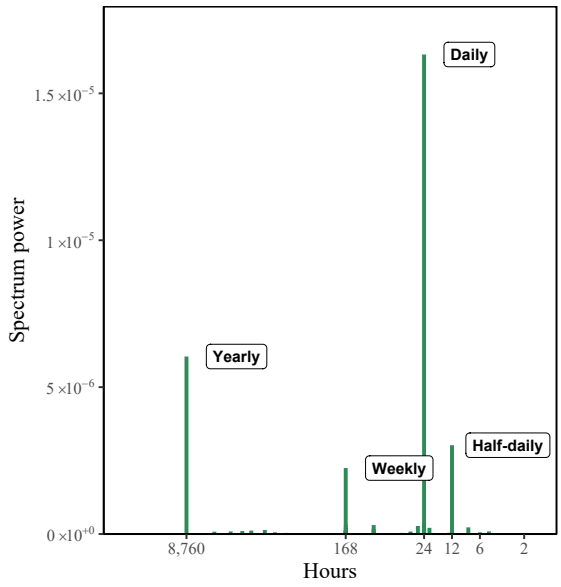
Note: Although not shown here, the ACF and PACF for $y^1_{h,t} = \ln(e_{h,t}/E_t)$ show identical patterns.
Source: Energy Data Service (2018)

The ACF and PACF for $e_{h,t}/E_t$ (Model 1) is plotted in figure E.3 with a lag of 8 weeks (1.344 hours). The ACF in figure E.3a shows a high degree of auto-correlation, however, it is diminishing rather quickly. Similarly the PACF in figure E.3b converges quickly to zero. These two observations indicate that the time-series are stationary[36]. The ACF also shows clear seasonalities. The ACF is peaking every 24 hours and 168 hours, suggesting a daily and weekly seasonality, respectively. Looking at the time-series in figure E.2 also suggest a yearly cycle.

The patterns of seasonality are exactly what the function $z_{h,t}$ is trying to capture. A more formal examination of these seasonalities can carried out by means of a periodogram (Bloomfield, 2004). The basic idea of the periodogram is to measure the spectral density (or spectrum power) running on a certain frequency where the frequency is the inverse of the seasonality, e.g. $1/24$ for the daily frequency. The periodogram for $e_{h,t}/E_t$ is shown in figure E.4, and suggests there are clear consumption cycles within every 12 hour (half-daily seasonality), 24 hour (daily seasonality), 168 hour (weekly seasonality), and every 8760 hour (yearly seasonality)[37]. These findings are consistent with the ACF and PACF in figure E.3 with the exception of the cycle every 12 hour; for this reason the 12^{th} hour is not included in $z_{h,t}$[38].

[36] Similarly patterns hold for $\ln\left(e_{h,t}/E_t\right)$ although not shown here.
[37] Although not shown, the result for $\ln(e_{h,t}/E_t)$ is similar.
[38] We have also examined sub-series plots of the 12^{th} hour and there does not seem to be any consistent cycle running at this frequency.

Figure E.4: Periodogram of $e_{h,t}/E_t$



In the modelling of $z_{h,t}$ we evaluate two different sets of regressors based on the findings of seasonality. The first set includes harmonic regressors following Bloomfield, 2004. The intuition is that a wave can be approximated using trigonometric functions. A sinusoidal function, $x(h)$, on a certain frequency, $f = 1/s$, where s is the seasonality period in units of time (e.g. $s = 24$) is given by

$$x(h) = A\cos 2\pi\left(\frac{h}{s} + \phi\right) \tag{42}$$

$$= \beta_1 \cos\left(2\pi\frac{h}{s}\right) + \beta_2 \sin\left(2\pi\frac{h}{s}\right), \qquad \beta_1 \equiv A\cos 2\pi\phi, \quad \beta_2 \equiv -A\sin 2\pi\phi \tag{43}$$

In (43) A is the amplitude, ϕ is the phase, h is the indicator for hour, and, finally, β_1 and β_2 are the parameters to be estimated. To approximate a complex cycle such as in figure E.1 we include multiple fourier terms, ω_s, for the same seasonality although with the limitation that $\omega_s \in [1; s/2]$. Consequently, the harmonic regressor set is given by

$$\begin{aligned} y^i_{h,t} = \beta_0 + \overbrace{\sum_{s}\sum_{\omega_s}^{\Omega_s}\left[\beta_{1\omega_s}\cos\left(\frac{2\pi\omega_s h}{s}\right) + \beta_{2\omega_s}\sin\left(\frac{2\pi\omega_s h}{s}\right)\right]}^{\text{Fourier series}} \\ + \beta_3 h + \text{WeekDay} + \eta^i_{h,t} \\ \forall \quad s = \{24, 168, 8670\} \text{ and } \omega_s \in [1; \Omega_s] \text{ with } \Omega_s \leq \frac{s}{2} \end{aligned} \tag{X-Harmonic}$$

A dummy for the day of the week (excluding Monday), WeekDay, is included in (X-Dummy) because the weekly cycle is relatively complex and the set of dummies aids the fourier terms in capturing the weekly seasonality[39]. The second specification consist of a set of dummy variables

[39] Another way of formulating it is that $\Omega_1 68 = 84$ is too few regressors to fully capture the weekly cycle.

given by:

$$y^i_{h,t} = \beta_0 + \underbrace{\text{HourOfDay}}_{K=23} + \underbrace{\text{WeekDay}}_{K=6} + \underbrace{\text{Week}}_{K=51} + \underbrace{\text{HourOfDay} \times \text{WeekDay}}_{K=23\times 7} + \underbrace{\text{WeekDay} \times \text{Week}}_{K=6\times 52}, \tag{X-Dummy}$$

where K indicates the number of regressors for each variable. The intercept β_0 is level of y^i_h on Monday in the first week of the 2011 at hour 0.

Model 1 and Model 2 are estimated using OLS, Ridge, and Lasso since this is an prediction exercise. The objective of lasso and ridge is to minimize the prediction error subject to a tuning parameter:

$$\min_{\beta}(y_t - x_t\beta)^2, \qquad \sum_{k=1}^{K} \beta_k^2 \leq \lambda \tag{Ridge-Problem}$$

$$\min_{\beta}(y_t - x_t\beta)^2, \qquad \sum_{k=1}^{K} |\beta| \leq \lambda \tag{Lasso-Problem}$$

where K is the length of the vector β. The tuning parameter, λ, is the "penalty" on parameters sizes (i.e. the shrinkage effect). When $\lambda = 0$ Ridge is equivalent to OLS. The attractive feature of Ridge and Lasso is the possibility of choosing λ optimally by minimising the MSE of the model. In other words, it chooses the optimal balance between bias and variance of the model's prediction error:

$$\begin{aligned} \hat{\lambda} &= \underset{\lambda \in \Lambda}{\text{argmin}}\, \text{MSE}(\hat{\eta}(\lambda)) \\ &st. \\ \text{MSE}(\hat{\eta}(\lambda)) &\equiv \mathbb{E}[(\hat{\eta}(\lambda))^2] = \mathbb{E}[(y - \hat{y}(\lambda))^2] \\ &= \text{Bias}(\hat{y}(\lambda))^2 + \sigma^2 + Var(\hat{y}(\lambda)), \end{aligned} \tag{44}$$

where Λ is the set of tuning parameters and σ^2 is the variance of y. The bias-variance trade-off can also be interpreted as a problem of over- and under-fitting the model on the training data. If the bias is high the model is under-fitted and the model can miss important relations on the training data important for predicting the test data. Conversely, if the variance is high the model is over-fitted and can use what essentially is random noise in the training data to predict the hourly load on the test data. The optimal choice $\hat{\lambda}$ balances these two trade-offs and it is implemented by means of cross-validation where 525 observations are left out in each estimation, implying the model is estimated roughlu $(2016 - 2011 + 1) * 8760/525 \approx 100$ times. The tuning set consists of 102 elements and is defined as $\Lambda = 0 \cup \left\{n \in \{0, 0.1, 0.2, ..., 10,\} \ : \ f(n) = 10^{-n}\right\}$.

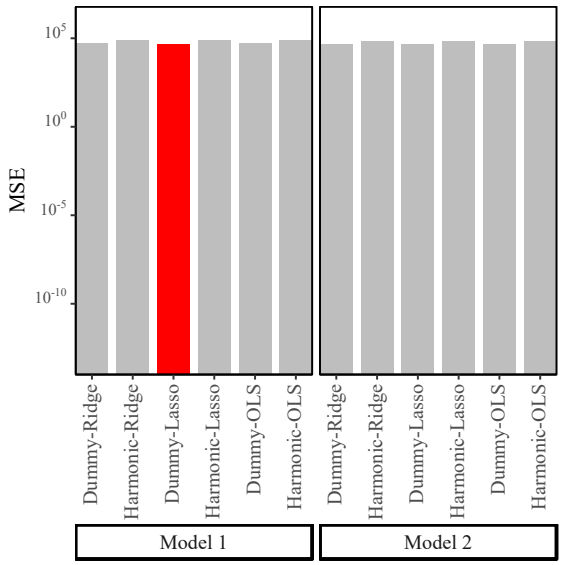
Finally, we have two different regression models (Model 1 and Model 2) with two sets of different regressors (Dummy and Harmonic) each estimated using three different methods (OLS, Ridge, and Lasso), implyinh we evaluate $2 * 2 * 3 = 12$ different models in total.

It should be noted that the error term of all models are auto-correlated and heteroschedastic suggesting that the models are not fully specified. For instance, we are not accounting for temperature, although this has been shown to be important in determining electricity demand.

Temperature is not included in $z_{h,t}$ (although it easily could be be) because it is a difficult variable to forecast. Importantly, the error term does not contain a unit root.[40]

The prediction performances of the 12 different models are presented in figure E.5. The specifications performs virtually the same albeit the dummy regressors generally outperforms the harmonic regressors. The best prediction is Model 1 with the set of dummy regressors using Lasso. This specification explains a little over 92% of the test data indicating the the assumption of $\phi_t = 1$ is not an unreasonable assumption.

Figure E.5: Prediction performance on test data (2017)



Note: The specification with the lowest MSE is indicated by red.

F A CNS-compatible, Linearly Homogeneous bottom up model

First, let $\phi = 1$ in (18). This yields the inelastic short-run demand form. Appendix C.1 and C.2.2 show that under assumptions 1-3 supply from all plants is continuously differentiable and that it nests the relevant bottom up assumptions. Appendix D shows a similar result for the trade function in definition 2.

To see that the short run equilibrium is linearly homogeneous, note that increasing yearly demand levels, plant capacities, and trade capacities by $\alpha\%$, increases all components in the equilibrium condition by $\alpha\%$ and thus leaves prices unchanged.

[40] The ACF and PACF of all estimations suggest a stationary series since lag-terms are converging to zero. Both plots indicates seasonality at the 24th and 168th lag suggesting that seasonality at these two frequencies has not been completely accounted for.

G Calibration

The model is calibrated to fit a the yearly moments for 2019 displayed in table 2. To this end, the model is adjusted in a number of ways. As a first step we make three adjustments. Table 5 outlines the relevant targets and parameters affected here. First, we target the overall price level by including a correction factor in the marginal cost function for all dispatchable domestic plants. Second, we include a marginal cost correction variable for all foreign dispatchable plants. Given that the total energy consumption throughout a year enters exogenously in the demand function (see section 4.4), we do not need to include a calibration parameter to fit energy consumption; it is always guaranteed to be at the correct level. Thus, we only have to target either gross production of electricity or the total net imports in order to fit both in our model. In calibrating to this target there are a number of alternatives to the marginal cost component of foreign plants. These include adjustment of trade capacities, trade smoothing parameters and the price elasticity of domestic demand for electricity. However, including a constant marginal cost component across all foreign plants achieves the target very easily. Reassuringly, the required adjustment to marginal costs of foreign plant is generally very small (less than 5 DKK/MWh). Third, we adjust the production capacities for a set of plants. This includes intermittent technologies that have (near) zero marginal costs as wind, solar and hydro.[41]

Table 5: Calibration target and parameters, step 1

Target	**Adjustment of parameter**
Average spot price (DK):	Marginal cost parameter added to all dispatchable Danish plants
Sum of net import (DK):	Marginal cost parameter added to all dispatchable foreign plants
Output targets (DK)	
Electricity, intermittent plants	⋮
Heat, intermittent plants	⋮
Surplus heating	Correction of yearly average capacity for relevant plant/fuel type
Electricity + heat, waste	⋮
Electricity + heat, bio gas	⋮

With the first three adjustments described above the model is relatively close to the targets for all relevant endogenous variables. To fit the rest of the targets we add correction parameters to a number of technical parameters. These include (1) efficiency corrections, (2) marginal cost corrections, (3) production capacity corrections, and (4) fuel mix coefficient matrix corrections. We allow all four types of corrections to be on a fuel type specific level in order to ensure that we reach the fuel input targets in table 2. In contrast to a structural estimation, the parameters we

[41] We also allow for the adjustment of production capacity for plants using biogas and waste as primary input factors. The reason for this is the same as for intermittent technologies. The supply is basically perfectly inelastic at a very low price and therefore fixed at capacity levels. Thus, the only way to affect the supply from these firms is basically through an adjustment of the capacity variable.

unfix here are clearly not identified from the targets above. We have approximately 3500 technical parameters that are endogenized and only eight fuel input targets to hit. Our reasoning behind this approach is that technical parameters should more or less be identified directly from our bottom-up data. As we *a priori* have no information on which of the technical parameters best explain the discrepancy between the model solution and the targets in table 2, we formulate a criterium function to be minimized, which penalizes large deviations from the original bottom-up data. This is done subject to the model being in equilibrium and that all targets in table 2 are met. The criterium function used here is given by

$$\min \mathcal{Q} \equiv \sum_{i \in \mathscr{F}_B} \sum_{j \in \mathscr{F}_F} \left(C_{i,j}^{MC}\right)^2 + w_{FM} \left(C_{i,j}^{FM} - FM_{i,j}^{data}\right)^2 + w_{Eff} \left(C_{i,j}^{Eff} - 1\right)^2 + w_{Cap} \left(C_{i,j}^{cap} - 1\right)^2,$$

where $\mathscr{F}_B$ are the set of basic fuel types and $\mathscr{F}_F$ are actual fuel types formed as a linear combination of the basic fuel types. The marginal cost correction $C_{i,j}^{MC}$ is added to all plants' marginal costs and is therefore measured in DKK/MWh. C^{FM} is the corrected fuel-mix matrix. It consists of variables between zero and one with entry (i, j) denoting the amount of basic fuel i that is used for a plant of fuel-type j. We only allow for non-zero values in the fuel-matrix to be adjusted. To remedy that the corrections in $C_{i,j}^{FM}$ are not measured in DKK as the marginal cost corrections, we include a weight w_{FM}. A large value of w_{FM} indicates that we should change the marginal costs of the plant, rather than coefficients in the fuel-mix matrix given by the bottom-up data. $C_{i,j}^{Eff}$ is a multiplicative correction term on plants' fuel efficiency. $C_{i,j}^{Cap}$ is a similar multiplicative correction term on plants' production capacity.

To illustrate the effect of the calibration, we plot the domestic supply functions split into technology and fuel types. Figure G.1 illustrates the supply of electricity in an average hourly state for Denmark. The plot shows that there are no plant type that is changed dramatically by our calibration. The most prominent change is the capacity for electrical heaters that is adjusted by in order to target the observed amount in table 2. Beyond this, small adjustments to the marginal costs are made for condensation plants (CD) and combined heat and power plants (CHP).

Figure G.2 shows a somewhat similar result when it comes to supply split onto fuel types. Notably, the category of fuels denoted 'other' is, in relative terms, altered quite a lot. This covers the categories bio gas, bio oil, hydrogen, and uranium. In the Danish case, this is specifically the use of bio gas that is prominent (there is no use of hydrogen nor uranium). Note, however, that even though the relative shift is large, the change in absolute terms is very small (less than 1% of total supply).

Figure G.1: Danish supply on plant types

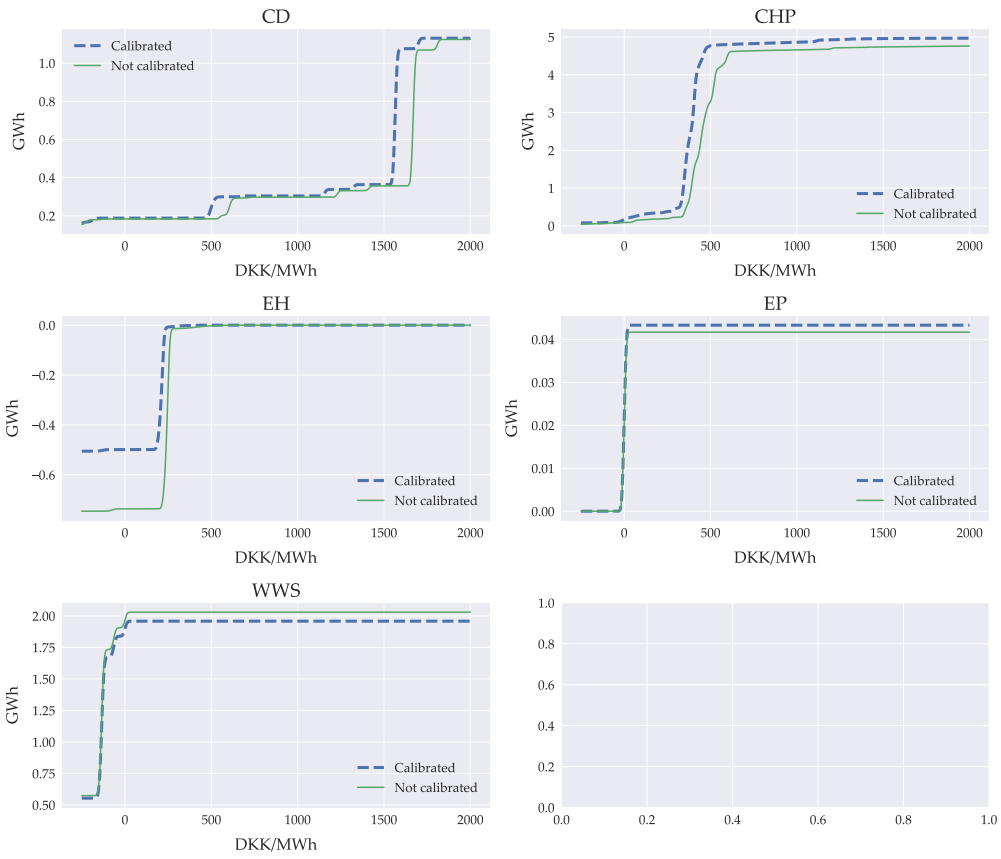
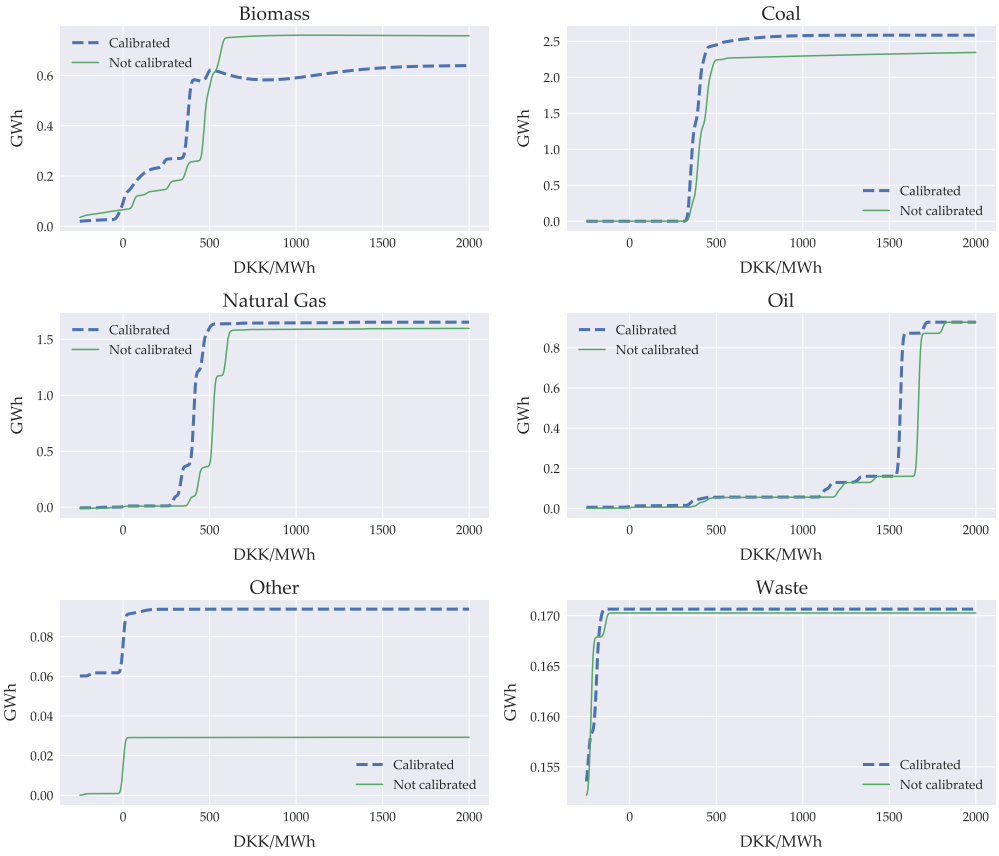


Figure G.2: Danish supply on fuel types



H Hourly variation in trade capacities

This appendix outlines the estimation of hourly variation in transmission lines' capacity between neighboring countries trading in electricity.

Let $T_{i,j,h,t}$ denote the observed net transfer export capacity from country i to country j in hour h and year t. We do not have data on the theoretical TTC; in stead we define the yearly TTC $\left(\overline{T}_{i,j,t}\right)$ as the maximum observed NTC within that given year, i.e. $\overline{T}_{i,j,t} \coloneqq \max\left\{T_{i,j,1,t}, ..., T_{i,j,h,t}, ..., T_{i,j,H,t}\right\}$. Note that $T_{i,j,h,t} \in [0, \overline{T}_{i,j,t}]$. Rather than estimating the absolute NTC we estimate the percentage reduction in the NTC compared to the TTC. We define the percentage reduction as $\Delta T_{i,j}(h,t) \coloneqq T_{i,j}(h,t)/\overline{T}_{i,j}(t) - 1$. The main advantage of this transformation is that $\Delta T_{i,j,h,t} \in [-1, 0]$ so that the upper bound is not year-specific but constant.

Our approach to predicting hourly variation in transmission capacities is the following: Remember that $\Delta T_{i,j}$ is truncated from below at -1 and from above at 0. We can describe $\Delta T_{i,j}$ by a double truncated Tobit model. We furthermore assume that the observed $\Delta T_{i,j,h,t}$ can be explained by two orthogonal latent variables. The first variable is a discrete random variable Z with three potential outcomes $\{-1, \zeta, 0\}$. The corresponding probabilities of an event of X is

$$\begin{aligned} p_j &\coloneqq \Pr(Z = j), \qquad j = \{-1, \zeta, 0\} \\ \sum_j p_j &= 1. \end{aligned}$$

We think of Z as being determined by exogenous forces outside of the model framework. For instance, the transmission line can exogenously fail or the TSO have to repair a transmission line in a given hour and therefore has to close it. To be specific, $Z = -1$ desribes an outage in the transmission line, implying $\Delta T_{i,j,h,t} = -1$. p_{-1} therefore denotes the probability of an exogenous failure to the transmission line so that trade capacities are zero. Event $Z = 0$ implies there are no exogenous technical constraints on the system and $\Delta T_{i,j,h,t} = 0$. Finally, if $Z = \zeta$ then $\Delta T_{i,j,h,t} \in (-1, 0)$. p_m therefore denotes the probability of a capacity reduction of $\Delta T_{i,j} \in (-1, 0)$. For instance, the TSO may choose to lower trade capacities due to planned outages in the connected domestic area. Besides the random variable Z, a second latent variable may predict the outcome of $\Delta T_{i,j}$. We assume this second force is described by a latent variable $\Delta T^*_{i,j,h,t}$. Hence, we formulate the model:

$$\Delta T_{i,j,h,t} = \begin{cases} 0 & \text{if} \quad Z = 0 \quad \text{or} \quad Z = \zeta \text{ and } \Delta T^*_{i,j,h,t} \geq 0 \\ \Delta T^*_{i,j,h,t} & \text{if} \quad Z = \zeta \text{ and } \Delta T^*_{i,j,h,t} \in (-1, 0) \\ -1 & \text{if} \quad Z = -1 \quad \text{or} \quad Z = \zeta \text{ and } \Delta T^*_{i,j,h,t} \leq -1 \end{cases} \tag{45}$$

Note that $\Delta T^*_{i,j}$ is only predicting $\Delta T_{i,j}$ whenever the random variable Z takes the event M.

The latent variable $\Delta T^*_{i,j,h,t}$ is given by the linear model:

$$\Delta T^*_{i,j,h,t} = \beta_0 + \sum_g \sum_k \beta_{g,k} X_{g,k,h,t} + \eta_{i,j}(h,t), \qquad \eta_{i,j,h,t} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2), \quad g = \{i,j\} \tag{46}$$

Our goal here is to estimate the hourly variation in transmission capacity in a 'normal' year. A priori we assume there are no yearly time trends in $\Delta T_{i,j,h,t}$ because it would violate the assumption of a normal year existing. It is possible, however, that some years are outliers due to exogenous forces. In the transmission line from DK1 to the Netherlands we find that in its first operating year (2018), the transmission line is closed much of the year. We account this 'new tech teething problem' by including a year specific fixed effect in the opening year in (46).

When choosing the explanatory variables in (46) (indexed by k) we are restricted by variables that are available to us in the techno-economic data. To be consistent with the model framework, we need to simulate the hourly variation in transmission capacities based on the variables included in the model. Here we choose forecasted hourly load (denoted by D) and forecasted generation from on-shore wind (denoted by W). Hence $k = \{D, M\}$ leaving us with four parameters to estimate in (46). All explanatory variables are measured in TWh.

Given Z and $\Delta^*_{i,j}$ are orthogonal, we can derive the density $f\left(\Delta T_{i,j} | X_i, X_j\right)$ as

$$f\left(\Delta T_{i,j} | X_i, X_j\right) = \begin{cases} p_0 + p_\zeta \Pr\left(\Delta T^*_{i,j} \geq 0 | X_i, X_j\right) & \text{if } \Delta T_{i,h} = 0 \\ p_\zeta \Pr\left(\Delta T_{i,j} | X_i, X_j\right) & \text{if } \Delta T_{i,h} \in (-1, 0) \\ p_{-1} + p_\zeta \Pr\left(\Delta T^*_{i,j} \leq -1 | X_i, X_j\right) & \text{if } \Delta T_{i,h} = -1 \end{cases} \tag{47}$$

With (47) the log-likelihood function (ℓ) is proportional to:

$$\begin{aligned} \ell(\beta, \sigma, p_{-1}, p_0) \propto & \sum_{\Delta T_{i,j}=0} \ln\left\{p_0 + (1 - p_{-1} - p_0)\left(1 - \Phi\left(\frac{X\beta}{\sigma}\right)\right)\right\} \\ & + \sum_{\Delta T_{i,j} \in (-1,0)} \ln\left\{(1 - p_{-1} - p_0)\frac{1}{\sigma}\phi\left(\frac{\Delta T_{i,j} - X\beta}{\sigma}\right)\right\} \\ & + \sum_{\Delta T_{i,j}=-1} \ln\left\{p_{-1} + (1 - p_{-1} - p_0)\left(1 - \Phi\left(\frac{-1 - X\beta}{\sigma}\right)\right)\right\} \end{aligned}$$

where we used that $p_\zeta = 1 - p_{-1} - p_0$. From this, we can formulate the Maximum Likehood (ML) estimation. It is important to note that $1 - p_{-1} - p_0 \geq 0$, implying a constraint on the ML problem. Formally, the constrained ML problem is defined as:

$$\begin{aligned} \max_{\beta, \sigma, p_{-1}, p_0} \ell(\beta, \sigma, p_{-1}, p_0) \propto & \sum_{\Delta T_{i,j}=0} \ln\left\{p_0 + (1 - p_{-1} - p_0)\left(1 - \Phi\left(\frac{X\beta}{\sigma}\right)\right)\right\} \\ & + \sum_{\Delta T_{i,j} \in (-1,0)} \ln\left\{(1 - p_{-1} - p_0)\frac{1}{\sigma}\phi\left(\frac{\Delta T_{i,j} - X\beta}{\sigma}\right)\right\} \\ & + \sum_{\Delta T_{i,j}=-1} \ln\left\{p_{-1} + (1 - p_{-1} - p_0)\left(1 - \Phi\left(\frac{-1 - X\beta}{\sigma}\right)\right)\right\} \end{aligned} \tag{48a}$$

subject to

$$1 - p_{-1} - p_0 \geq 0 \tag{48b}$$

For starting values of β and σ we use estimates from a OLS estimation of (46). p_{-1} and p_0 are initialized with their sample moments, i.e.

$$p_{-1}^0 = \frac{\sum_t \sum_h \mathbb{1}\left\{\Delta T_{i,j,h,t} = -1\right\}}{N}$$
$$p_0^0 = \frac{\sum_t \sum_h \mathbb{1}\left\{\Delta T_{i,j,h,t} = 0\right\}}{N},$$

where N is the totalt number of observations. Instead of optimizing the constrained ML problem we solve an unconstrained ML problem by disregarding (48b) and check whether $1 - p_{-1} - p_0 \geq 0$ is binding at any time. If the constraint is never binding - both on the path to the optimum and in the optimum - the solution to the unconstrained ML problem is the same as the one for the constrained ML problem. With the chosen starting values we find that the constraint in (48b) is never binding.

H.1 Data

Data on the explanatory variables are collected form ENTSO-E's transparency platform.[42] Descriptive statistics are shown in table 6. Note that data on explanatory variables are collected for each of the electricity markets listed in the first column of table 6. Data on hourly NTC between country pairs is collected from Nord Pool.[43] The data is on bidding-zone level (i.e. DK1, DK2, NO1 etc.) but we aggregate the data to match the electricity markets in the model. This leaves us with data on the 22 transmissions lines listed in the first column in 7. Descriptive statistics for the NTC is also shown in table 7. The 'whole samle' columns describe the number of observations when data on the dependent variable and explanatory variables are combined. We have tried collecting data for 2015-2020 and if we had data for every hours from both data sources in the sample period, the maximum number of observations would be 52,608 (accounting for the leap years 2016 and 2020). However, not all transmission connections have data for the entire sample period. For instance, data for the transmission line from DK1 to Germany (DEATLU) is only available for 2019-2020. Therefore, the maximum number of observations for this particular transmission line is 35,064. This is listed in the second last column in table 7. The last column defines the percentage of hours where are missing data for. For the most part, this share is relatively low except for the connections between Denmark and Germany (DEATLU), which is at 6.78%.

Finally, we note that we do not have data on all market couplings represented in the model. At illustrated in figure 2.1, there are 32 transmission lines in the model with only 22 connections present in table 7; and not all of the 22 connections is represented in the model. In stead, we

[42] https://transparency.entsoe.eu/.
[43] https://www.nordpoolgroup.com/historical-market-data/.

construct a unbalanced panel data for the missing connections. For instance, the we do not have data on the transmission line from Norway (NO) to Finland (FI). We then take all exporting transmission lines from Norway (these are the ones connected to DK1, The Netherlands, and Sweden) and all importing capacity to Finland (the only one comes from Sweden). We then have an unbalanced panel data set with four different transmission lines. The rest of the estimation remains the same.

H.2 Results

The estimated parameters for each of the transmission lines are reported in table 8. We are not directly interested in the value of the estimated parameters in table 8, but rather how well the model is fitting the data. The fit is illustrated in figure H.1 and, in general, the model provides a good fit of the variation in hourly transmission capacity. Note that we do not report the fit of the estimation for all connections but only a subset.[44]

With the estimated parameters, and the conviction that they prediction hourly transmission capacities well, we use the estimates to simulate the variation in the hourly transmission capacity consistent with our model framework. We simply due this by replacing the data on the explanatory variables in (46) with the corresponding variables in the techno-economic data.

[44] The remaining are available upon request.

Table 6: Descriptive statistics for explanatory variables

Electricity market	Forecasted load (D)					Forecasted wind generation (W)				
	mean	median	std	min	max	mean	median	std	min	max
Belgium	9.44	9.34	1.32	6.5	13.33	0.44	0.31	0.39	0.0	1.84
DEATLU	248.34	249.45	48.69	0.0	345.63	35.87	27.16	28.65	0.0	166.02
DK1	2.27	2.24	0.46	0.0	3.54	0.88	0.67	0.72	0.0	3.49
DK2	1.49	1.50	0.31	0.0	2.42	0.18	0.13	0.16	0.0	0.81
Finland	9.33	9.19	1.61	0.0	15.17	0.50	0.37	0.42	0.0	1.98
France	53.10	51.80	12.63	0.0	95.15	2.99	2.24	2.35	0.0	14.59
GBNIIE	65.15	65.01	15.43	0.0	111.78	6.47	5.64	4.16	0.0	21.13
Holland	48.76	47.62	9.86	0.0	81.87	3.93	2.97	3.38	0.0	23.38
Norway	15.11	14.75	3.20	0.0	29.16	0.51	0.38	0.44	0.0	4.03
Sweden	15.56	15.30	3.49	0.0	26.42	2.11	1.82	1.37	0.0	9.20

Source: ENTSO-E (2017a) and ENTSO-E (2017b).
Note: All variables are measured in TWh.

Table 7: Descriptive statistics for net transmission capacity (NTC)

Market coupling	NTC[i]					Whole sample[ii]		
	mean	median	std	min	max	T	Max T[iii]	Missing obs.[iv]
Belgium->GBNIIE	9.27	10.26	1.92	0.0	20.48	16,818	17,544	4.14%
DEATLU->DK1	1.23	1.5	0.44	0.0	3.0	32,686	35,064	6.78%
DEATLU->DK2	0.59	0.6	0.32	0.0	1.2	32,686	35,064	6.78%
DK1->DEATLU	0.47	0.32	0.49	0.0	1.78	32,686	35,064	6.78%
DK1->DK2	0.57	0.59	0.08	0.0	1.18	52,458	52,608	0.29%
DK1->Norway	1.23	1.29	0.36	0.0	3.06	52,458	52,608	0.29%
DK1->Sweden	0.56	0.72	0.21	0.0	1.43	52,458	52,608	0.29%
DK2->DEATLU	0.57	0.58	0.31	0.0	1.17	32,684	35,064	6.79%
DK2->DK1	0.59	0.6	0.08	0.0	1.2	52,458	52,608	0.29%
DK2->Sweden	1.28	1.65	0.5	0.0	2.1	52,458	52,608	0.29%
Finland->Sweden	2.19	2.27	0.3	0.24	4.64	52,602	52,608	0.01%
France->GBNIIE	12.29	16.28	8.05	0.0	40.47	52,590	52,608	0.03%
GBNIIE->Belgium	9.18	10.24	1.99	0.0	20.48	16,818	17,544	4.14%
GBNIIE->France	21.06	20.44	10.26	0.0	78.55	52,590	52,608	0.03%
Holland->Norway	0.6	0.72	0.21	0.0	1.45	25,077	26,304	4.66%
Norway->DK1	1.23	1.29	0.34	0.0	3.26	52,458	52,608	0.29%
Norway->Holland	0.57	0.72	0.23	0.0	1.45	25,077	26,304	4.66%
Norway->Sweden	2.66	2.78	0.68	0.65	6.96	52,602	52,608	0.01%
Sweden->DK1	0.57	0.68	0.19	0.0	1.36	52,458	52,608	0.29%
Sweden->DK2	1.15	1.3	0.31	0.1	2.6	52,458	52,608	0.29%
Sweden->Finland	2.64	2.71	0.23	0.78	5.48	52,602	52,608	0.01%
Sweden->Norway	2.84	3.0	0.76	0.27	7.89	52,602	52,608	0.01%

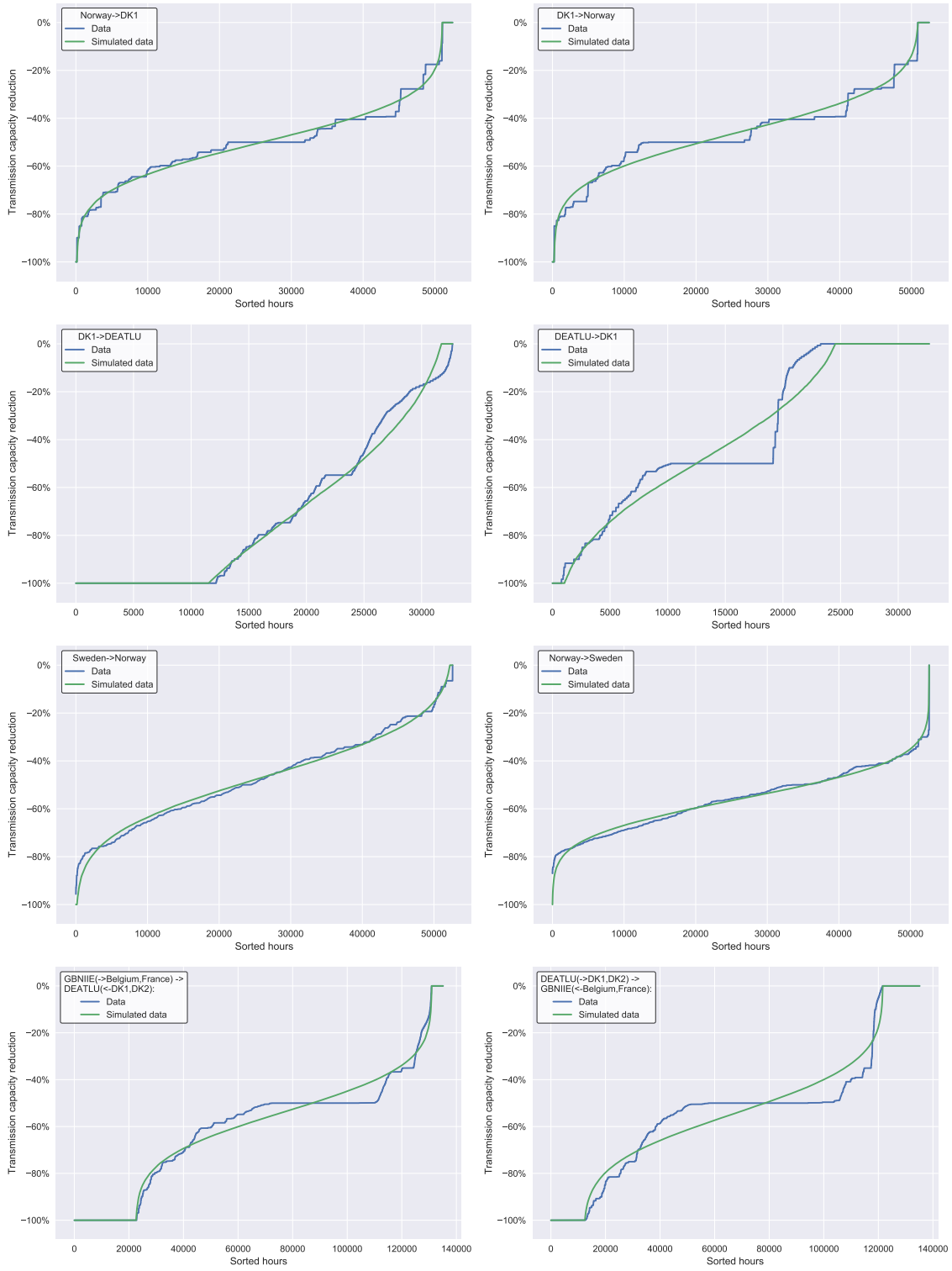
Source: Nord Pool (2021b) and Nord Pool (2021a).
[i] Net Transmission Capacity (NTC). All statistics are measured in TWh.
[ii] The whole sample is comprised of both the data from NordPool and ENTSO-E's Transparency Platform for which we have data from both sources in a given hour.
[iii] Max N defines the maximum number of observations if we had observations for every hour in the sample period.
[iv] The number of missing observations is the ratio between column 'N' and 'Max N'.

Table 8: Estimation results

Market coupling (i->j)	Parameters[i]								Obs.[ii]	N[iii]	Export samples[iv]	Import samples[v]
	β_0	$\beta_{i,D}$	$\beta_{i,W}$	$\beta_{j,D}$	$\beta_{j,W}$	p_{-1}	p_0	σ				
DEATLU->DK1	-0.469	-0.124	1.387	-58.679	-0.706	0.021	0.000	0.127	32,686	1		
DEATLU->DK2	-0.501	0.049	0.199	-7.333	-60.534	0.148	0.129	0.029	32,686	1		
DK1->DEATLU	-0.577	56.805	-83.321	-0.797	-1.613	0.093	0.000	0.353	32,686	1		
DK1->DK2	-0.503	-29.377	-1.730	28.650	29.212	0.000	0.000	0.041	52,458	1		
DK1->Norway	-0.622	-64.704	26.449	13.982	-102.566	0.003	0.000	0.119	52,458	1		
DK1->Sweden	-0.596	8.481	-8.689	16.223	-47.643	0.049	0.095	0.222	52,458	1		
DK2->DEATLU	-0.498	-10.811	-72.275	0.056	0.228	0.153	0.129	0.032	32,684	1		
DK2->DK1	-0.489	16.469	34.459	-19.403	-2.925	0.015	0.000	0.035	52,458	1		
DK2->Sweden	-0.635	135.270	-188.259	9.294	-17.501	0.000	0.127	0.300	52,458	1		
Finland->Sweden	-0.572	1.284	-8.149	2.546	-0.202	0.000	0.000	0.074	52,602	1		
Holland->Norway	-0.574	2.181	0.044	-5.818	13.260	0.077	0.000	0.083	25,077	1		
Norway->DK1	-0.600	10.061	-57.113	-54.122	8.867	0.000	0.000	0.147	52,458	1		
Norway->Holland	-0.589	0.877	-0.316	0.307	-0.690	0.093	0.000	0.102	25,077	1		
Norway->Sweden	-0.741	17.268	-35.836	-6.373	0.894	0.000	0.000	0.101	52,602	1		
Sweden->DK1	-0.563	10.032	-6.712	-67.000	7.763	0.036	0.000	0.113	52,458	1		
Sweden->DK2	-0.574	4.900	-3.784	-30.892	-34.634	0.000	0.000	0.119	52,458	1		
Sweden->Finland	-0.530	1.659	-0.547	-1.652	7.041	0.000	0.000	0.041	52,602	1		
Sweden->Norway	-0.756	4.871	-1.172	17.070	-132.279	0.000	0.000	0.170	52,602	1		
DEATLU->GBNIIE	-0.471	-0.437	0.820	-2.185	5.696	0.090	0.031	0.130	134,780	4	DEATLU->DK1,DK2	GBNIIE<-Belgium,France
DEATLU->Holland	-0.543	-0.138	1.350	-0.701	-1.220	0.046	0.093	0.156	90,449	3	DEATLU->DK1,DK2	Holland<-Norway
DEATLU->Norway	-0.638	0.216	1.566	8.715	-71.897	0.039	0.021	0.139	195,509	5	DEATLU->DK1,DK2	Norway<-DK1,Holland,Sweden
DEATLU->Sweden	-0.582	-0.116	1.232	6.330	-10.069	0.025	0.096	0.162	275,492	6	DEATLU->DK1,DK2	Sweden<-DK1,DK2,Finland,Norway
DK1->GBNIIE	-0.508	-2.229	4.946	-0.089	-2.394	0.073	0.038	0.159	259,468	6	DK1->DEATLU,DK2,Norway,Sweden	GBNIIE<-Belgium,France
DK1->Holland	-0.504	-2.439	-4.652	-0.083	-3.251	0.065	0.046	0.159	215,137	5	DK1->DEATLU,DK2,Norway,Sweden	Holland<-Norway
Finland->Norway	-0.645	0.996	-3.814	7.230	-8.362	0.012	0.000	0.133	182,739	4	Finland->Sweden	Norway<-DK1,Holland,Sweden
GBNIIE->DEATLU	-0.492	-1.337	-5.932	-0.177	-0.303	0.159	0.031	0.166	134,778	4	GBNIIE->Belgium,France	DEATLU<-DK1,DK2
GBNIIE->DK1	-0.532	-0.307	1.277	-1.862	-11.043	0.030	0.000	0.119	259,468	6	GBNIIE->Belgium,France	DK1<-DEATLU,DK2,Norway,Sweden
GBNIIE->Holland	-0.436	-0.052	-6.961	-3.154	-1.529	0.072	0.000	0.136	94,485	3	GBNIIE->Belgium,France	Holland<-Norway
GBNIIE->Norway	-0.498	0.965	-2.117	-3.233	-19.546	0.025	0.000	0.154	199,545	5	GBNIIE->Belgium,France	Norway<-DK1,Holland,Sweden
Holland->DEATLU	-0.518	-0.225	-3.689	-0.017	-0.465	0.211	0.047	0.144	90,447	3	Holland->Norway	DEATLU<-DK1,DK2
Holland->DK1	-0.565	-0.401	1.700	-1.599	0.062	0.003	0.001	0.109	215,137	5	Holland->Norway	DK1<-DEATLU,DK2,Norway,Sweden
Holland->GBNIIE	-0.444	-2.289	2.470	-1.405	5.526	0.089	0.000	0.129	94,485	3	Holland->Norway	GBNIIE<-Belgium,France
Norway->DEATLU	-0.631	4.210	-47.377	0.414	-0.108	0.100	0.022	0.127	195,507	5	Norway->DK1,Holland,Sweden	DEATLU<-DK1,DK2
Norway->Finland	-0.646	4.263	15.515	0.584	-5.809	0.014	0.000	0.099	182,739	4	Norway->DK1,Holland,Sweden	Finland<-Sweden
Norway->GBNIIE	-0.567	-2.218	-5.011	0.525	5.492	0.044	0.000	0.123	199,545	5	Norway->DK1,Holland,Sweden	GBNIIE<-Belgium,France
Sweden->DEATLU	-0.604	5.803	-9.509	0.308	-0.558	0.069	0.015	0.137	275,490	6	Sweden->DK1,DK2,Finland,Norway	DEATLU<-DK1,DK2

[i] β_0 is the intercept, β_W is the coefficient related to forecasted on-shore wind generation and β_D is the coefficient related to forecasted demand.
[ii] Obs. are the total number of observations.
[iii] N is the number of crossectional samples included in the estimation, i.e. the sum of export and import samples.
[iv] Export samples are only shown for market couplings that we do not have data on. They are the the samples of transmission lines going from country i to neighboring partners in trade included in the estimation.
[v] Import samples are only shown for market couplings that we do not have data on. They are the samples of transmission lines going from neighboring partners in trade to country j included in the estimation.

Figure H.1: Model fit of estimated variation in transmission capacities



I Simulation results

This appendix includes additional results from the counterfactual scenario.

Aggregate results

To investigate the underlying mechanisms in the decrease of CO_2-e emissions illustrated in figure 6.2 in the main text, we decompose the change into four effects. By definition emissions are given by

$$CO_2\text{-e} = \underbrace{\frac{CO_2\text{-e}}{\text{Energy consumption}}}_{\equiv\mu} \times \underbrace{\frac{\text{Energy consumption}}{\text{GDP}}}_{\equiv\eta} \times \underbrace{\frac{\text{GDP}}{N}}_{\equiv y} \times N, \tag{49}$$

where μ is the CO_2-e intensity of energy consumption (or the average CO_2-e emission coefficient), η is the energy intensity of GDP (or the average energy inefficiency), y is GDP per capita and N is the total population size. According to (49), we can decompose the log-difference in CO_2-e emission from 2019-2030 as

$$\Delta CO_2\text{-e} = \Delta\mu + \Delta\eta + \Delta y + \Delta N,$$

where $\Delta x = \ln(x_{2030}) - \ln(x_{2019})$. The results of the decomposition are show in figure I.1. For clarity, percentages changes are show above or below the bars. Note, however, that percentages changes do not sum to the total change in CO_2-e emissions.

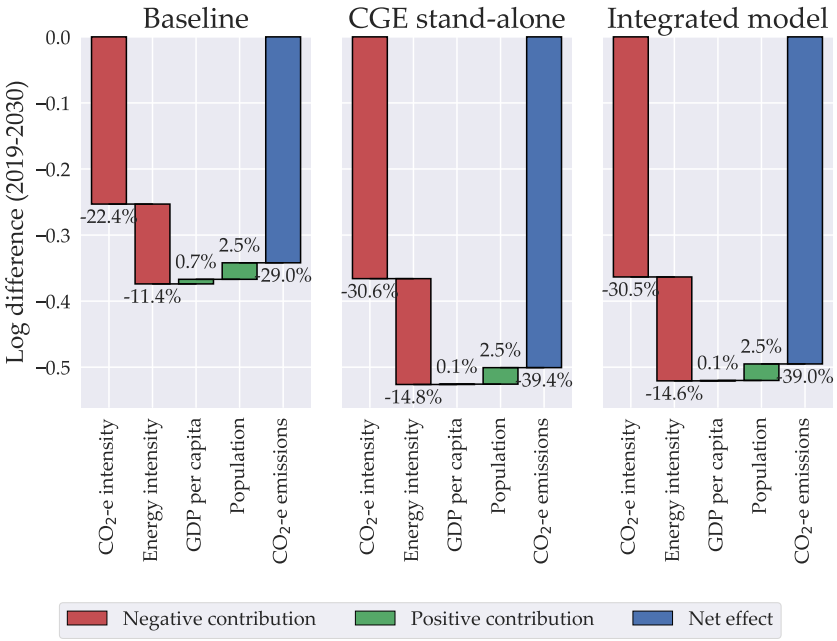
In the left panel of figure I.1, the CO_2-e reduction of 29% in the baseline scenario is equivalent to -0.35 in log-differences. The main reason for this decrease is a decrease in the CO_2-e intensity of the fuels used. Hence, in the baseline, the economy is gradually shifting to cleaner fuels and the average CO_2-e intensity drops by 22.4%. Another factor explaining the drop in emissions is a drop in energy intensity. These two effects more than offset the positive contribution from income and population growth.

When examining the effect of a higher CO_2-e tac, the additional decrease in emissions from the baseline is explained by a further decrease in CO_2-e intensity and energy inefficiency. The results from the 'stand-alone CGE' model and the integrated model are virtually the same. Both models also estimates a small negative scale effect as GDP per capita expected grows less than in the baseline scenario.[45]

That the two models predicts the same relative decrease in economic activity is also evident in aggregate GDP, consumption and investments. This is illustrated figure I.2.

[45] Note that population growth is exogenously given in the model, impying that the CO_2-e emission from a changing population does not change in the counterfactual.

Figure I.1: Decomposition of change in aggregate CO_2-e emission from 2015-2030



Note: Bars are measured in log-differences from 2019-2030. Numbers above the bars is the percentage change from 2019-2030.

Figure I.2: Relative decline in GDP, consumption and investments from baseline

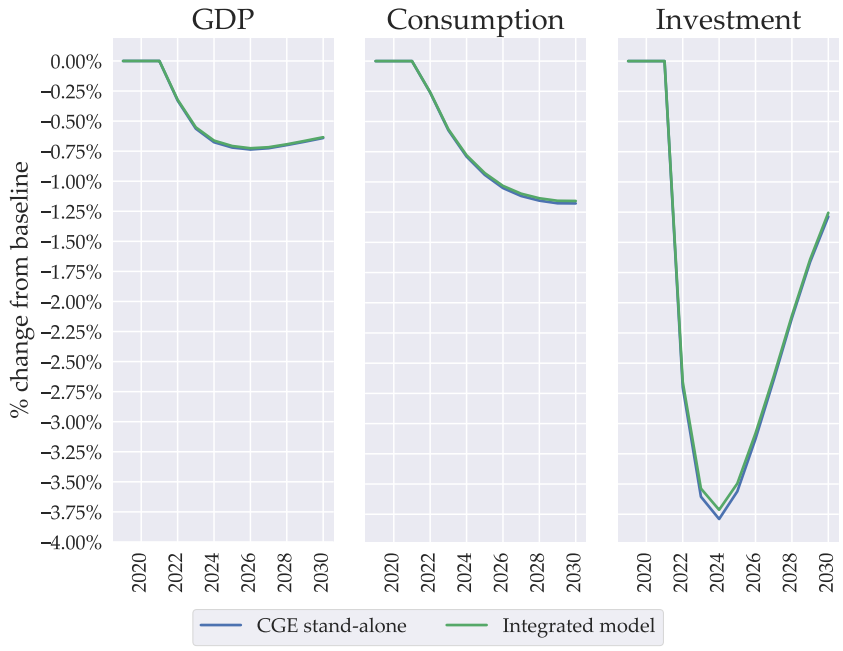


Figure I.3: Fuel change in non-ETS sectors from the baseline

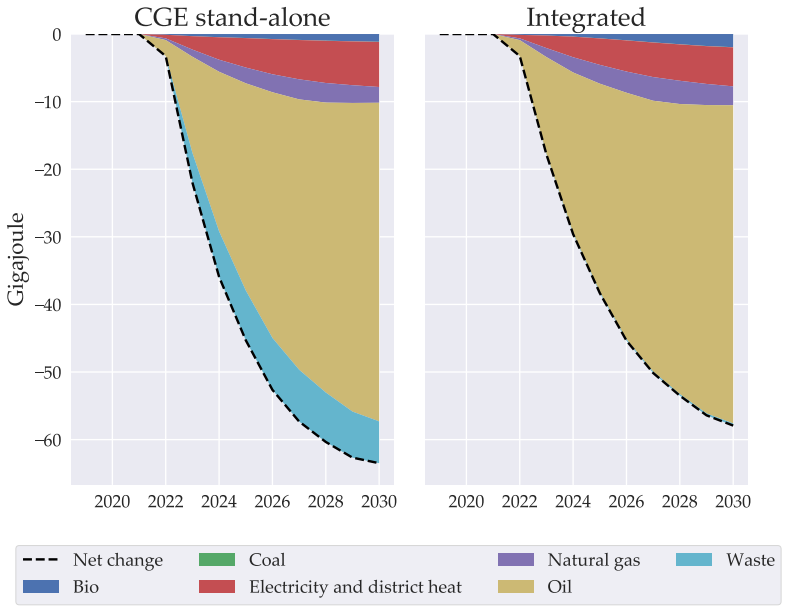
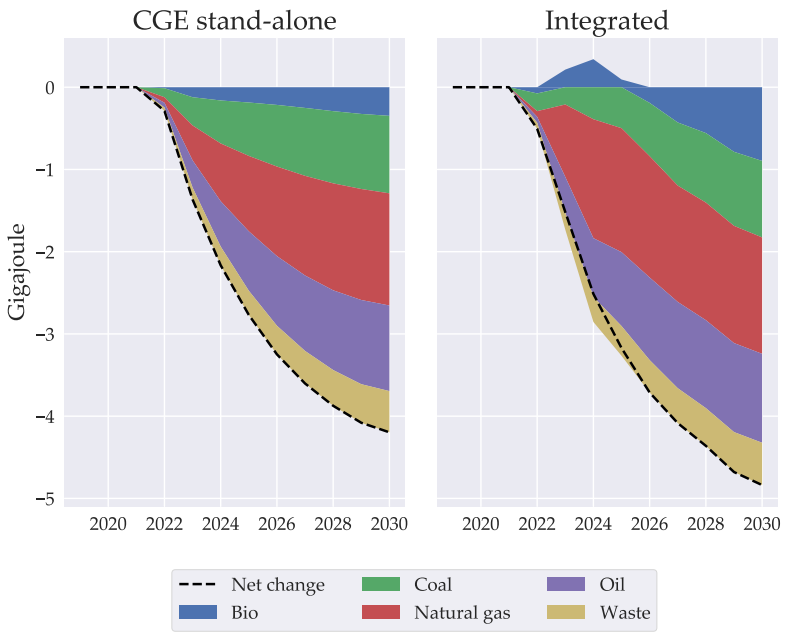


Figure I.4: Fuel change in ETS sectors from the baseline



Part 2

The Value and Potential of Electricity Storage

Disclaimer: *The current part extends the work from my Master's thesis. Some text may be repeated from Berg and Eskildsen (2019).*

The Value and Potential of Electricity Storage*

Rasmus K. Berg† Janek B. Eskildsen‡

December 29, 2021

Abstract

Reductions in greenhouse gas emissions from electricity generation are largely achieved through the expansion of intermittent renewable energy. A key technology expected to mitigate the economic consequences of intermittency and increase the share of renewables is electricity storage. In this paper we develop a convex optimization representation of electricity storage technologies that is particularly easy to solve numerically and nests a traditional linear optimization framework. In a structural estimation using data on Norwegian hydroelectric power plants, we show that our framework is able to replicate observed patters in power generation. Importantly, the estimated parameters suggest that our framework is a significant improvement on the traditional linear optimization approach.

To asses the value and potential of electricity storage, we include our representation of electricity storage in a detailed model for Northwestern European electricity and district heat systems. The model accounts for existing generation technologies, access to trade in electricity, and potential of demand flexibility technologies. We focus on Denmark, which has a high share of capacity based on intermittent energy sources. Compared to previous estimates, we find that electricity storage lowers both average electricity prices and greenhouse gas emissions significantly. This happens even though the considered storage technologies are net consumers of energy. Furthermore, although electricity storage increases the market value of technologies based on intermittent energy sources, electricity storage is, in itself, unlikely to become profitable before 2030 in a Danish context. This is because other strategic substitutes for mitigating intermittency - such as trade in electricity and dispatchable back-up capacity - reduce the scope to profit from large price variations. We emphasize that our conclusions likely depend on the valuation of price stability and emission reductions as well as ambitions for electrification of economic activity.

Keywords: Energy storage; Intermittency; Environment and trade; Bottom-up modeling; Renewable energy integration.

JEL classification codes: C6; F1; Q4.

*The authors wish to thank Peter Kjær Kruse-Andersen, Peter Birch Sørensen, and Peter Stephensen for helpful comments and feedback.

†University of Copenhagen. Oster Farimagsgade 5, 1353 Copenhagen K, Denmark. E-mail: rasmus.kehlet.berg@econ.ku.dk.

‡University of Copenhagen. Oster Farimagsgade 5, 1353 Copenhagen K, Denmark. E-mail: jbe@econ.ku.dk

1 Introduction

A common feature in roadmaps to ensure reductions in greenhouse gas emissions is a massive expansion in generation capacity of intermittent renewable energy sources (Jacobson et al. 2017; IEA 2021). In many places the trend of increased use of intermittent renewables is already well underway. For instance, the electricity capacity share of wind, water, and solar technologies (WWS) increased from 28% to 42% in the EU-27 between 2010-2019 (Eurostat 2021). However, as economies rely more and more on intermittent energy it becomes increasingly challenging to retain the stability of the electricity grid by continuously balancing demand and supply.

The value of electricity storage derives from this stability challenge. While electricity storage can provide a number of services (e.g. operating reserves, Bistline et al. 2020; Denholm, Jorgenson, et al. 2013; Denholm and Hand 2011), its primary value comes from energy arbitrage, i.e. the ability to move excess electricity in off-peak hours of demand to peak hours (de Sisternes et al. 2016). A substantial body of literature has been dedicated to evaluating the value and effects of electricity storage, and the jury is still out: Several recent papers conclude that large-scale deployment of electricity storage is not cost-effective, may lead to increased emissions and consumer prices, and may even be detrimental to social welfare. However, as e.g. Linn and Shih (2019) and Bistline et al. (2020) point out, the conclusions depend crucially on a number of conditions, including the generation capacity shares of existing electricity generating technologies and accurately accounting for strategic substitutes to electricity storage in mitigating intermittency.

We contribute to this body of literature by evaluating the effects of electricity storage in Denmark in a new model of the Northwestern European electricity systems developed in Berg and Eskildsen (2021). First, to describe the supply from storage technologies, we develop a convex optimization representation of a traditional linear optimization framework using smoothing techniques. We show that our approach results in a square system of nonlinear equations that uniquely identifies the global optimum. The system of equations is convenient because it nests the conventional linear framework when the smoothing parameters approach zero, and because it is particularly easy to solve numerically, as it can be topologically ordered. In a structural estimation using data on Norwegian hydroelectric power generation, we show that our model for electricity storage is able to produce very similar patterns in generation to what we observe in the data. Importantly, we note that the estimated smoothing parameters are quite large, suggesting that the traditional linear optimization framework overestimates how flexible storage technologies are. Second, we use the proposed framework for energy storage in the model of the Northwestern European electricity system to assess the value and potential of electricity storage in Denmark. In this context, we estimate that large-scale electricity storage from lithium-ion batteries leads to considerable reductions in price volatility, average electricity prices, and greenhouse gas emissions. Furthermore, technologies based on intermittent sources such as wind energy are complements to electricity storage in the sense that the value relative to other technologies is increasing in installed storage capacity. We further conclude that electricity storage is not privately cost-effective with

the current price volatility in Denmark, but is likely to be beyond 2030.

The assessment of the value and potential of electricity storage requires a technologically detailed description of the electricity system. It is important to include a realistic representation of complementary technologies, such as WWS technologies, as well as strategic substitutes to storage in mitigating intermittency, such as flexible demand, trade in electricity, and dispatchable backup technologies (Cebulla et al. 2017; Bistline et al. 2020). In our analysis, we account for all of these aspects. First, while electricity storage shifts supply on the temporal dimension, trade can move supply on a spatial dimension and therefore presents another method for mitigating intermittency. To incorporate trade in electricity, the model features eight electricity markets connected via transmissions lines with technical capacity constraints; we solve for equilibria prices in all eight connected regions. Second, to capture the potential of flexible demand technologies, hourly demand is split into two components: An inelastic, habit-driven share of demand and a price-responsive one. This results in multinomial-logit-like hourly-to-yearly demand ratios with key parameters capturing (i) a share of demand that is habitual (i.e. price inelastic demand), (ii) a share of demand that observes and reacts to short run prices, and (iii) the price responsiveness of elastic demand. Third and finally, to capture a realistic generation mix, market supply is derived from aggregation of individual plants that feature heterogeneity in technology type, fuel inputs, efficiency, and generation capacity, to name a few.

Another essential aspect of evaluating the value and potential of electricity storage is capturing the temporal resolution (Diaz et al. 2019). However, a technologically detailed model with a full hourly resolution that covers all years from 2019-2040 quickly becomes computationally intractable. To reduce the dimensionality of the model, we follow the common approach of aggregating hours within the year to representative states. In our model of the electricity and district heat systems, we draw on exogenous hourly variation in demand, generation capacities of intermittent plants, transmission line capacities, and water inflow into the reservoirs of hydro power plants. To capture seasonal energy storage, we start by estimating three seasonal breaks in the hourly data using conditional inference trees (Hothorn et al. 2006). Subsequently, we apply a customized K-means algorithm to identify representative states within each season. This establishes a non-chronological state space of K clusters of data within each year. The challenge of the non-chronological aggregation is to adjust the dynamic optimization of electricity storage technologies and, in particular, to capture the serial charge constraints.[1] In this regard, we also make a methodological contribution; we develop an approximation to the hourly model that can be solved over non-chronological states that provides relatively accurate results for as low as 25 intra-yearly states.

In the literature on storage modeling, there are largely two approaches to the issue of chronology: The first approach is to reestablish chronology. For instance, the design days approach solves

[1] The serial charge constraints are the combination of the law of motion of stored energy and storage capacity constraints. Essentially, it states that for a battery to discharge x GWh of electricity in a given hour, it has to be sufficiently charged at that point in time so as not to violate the storage capacity constraints. This depends on the history of charge/discharge decisions.

for a number of independent days with 24 chronologically ordered hours (Krajačić et al. 2011; Nahmmacher et al. 2016; de Sisternes et al. 2016; Poncelet et al. 2017). As the days are independent, however, they do not capture medium and long run storage across days. To remedy this issue, various papers rely on variations on the multi-grid solution where a subset of variables related to the level of stored energy is defined over a more coarse time grid that accounts for the dynamic between design days (e.g. Renaldi and Friedrich 2017; Kotzur et al. 2018; Gabrielli et al. 2018).[2] To obtain a reasonable approximation of a model with full hourly resolution, Gabrielli et al. (2018) use around 25 design days (roughly 600 intra-yearly states). The second approach is to adjust the storage model to the non-chronological state space. Wogrin et al. (2016) develop a 'system states framework' for electricity storage plants that only solves for the optimal charge/discharge decisions for each cluster state. However, to capture serial charge constraints in the non-chronological state space, Wogrin et al. (2016) propose computing the hourly storage level for all hours where storage constraints may be violated. While Wogrin et al. (2016) show that this approach produces a good representation of seasonal storage plants' behavior with 100 intra-yearly states, this way of checking for implied hourly constraints adds significantly to the computational burden of deriving supply from storage technologies. We note that as our model includes a greater level of detail along other dimensions, we require much larger reductions in the number of intra-yearly states.

To avoid computing the implied levels of stored energy on an hourly level, we define an approximation that only evaluates the storage constraints in cluster-state averages. In this Markov representation, a storage plant decides how much to charge/discharge in all states taking into account (i) the average level of stored energy when entering the state and (ii) which cluster states it is likely to transition to in the future. To remedy that this approach only relies on cluster-state averages, we propose extending the framework and solving for charge/discharge decisions for all cluster-tuple states (k, j), where k indicates the current cluster state and j indicates the next one. This extension implies that the storage plant can choose to charge in some states (k, j) if the price in state j is relatively high and discharge in other states (k, l) if the price in state l is relatively low.

We emphasize that our proposed method is only an approximation, as it still only solves for cluster-tuple state averages, but that, in our applications at least, the approximation provides very similar results to the full hourly version of the model. To give some idea of the computational advantages of our model, we solve the entire model, establishing equilibria on a number of European electricity markets and Danish district heat markets from 2019-2040, in roughly the same time as it takes to solve the "stylized electric power system over a time horizon of seven weeks" (10-15 seconds, Wogrin et al. 2016).[3]

Finally, our paper relates to the literature that evaluates the value of electricity storage, and

[2] Gabrielli et al. (2018) proposes defining a sequence of design days and only solving for the binary variables of the mixed integer linear-programming problem between days; Kotzur et al. (2018) define intra- and inter-period states to capture e.g. seasonal variation.

[3] We solve the model using CONOPT 4.15 in GAMS on an Intel(R) Core(TM) i7-6600U 2.60GHz with 2 cores and 8GB RAM. As is generally the case with nonlinear systems, the solution time is highly sensitive to the starting point. With completely naive starting values of constant prices, the solution time remains within a few minutes.

its effects on prices, emissions, and renewable energy. A number of papers look at these issues by investigating the electricity market in Texas. de Sisternes et al. (2016) look at a 'greenfield' experiment that selects a long run optimal generation mix.[4] They find that short run energy storage can increase penetration of capacity based on intermittent energy, reduce investments in dispatchable capacity, and increase utilization of all installed capacity. However, these benefits do not outweigh the projected costs of the technology. Similarly, Sioshansi (2011) shows that the value of wind energy can increase significantly with electricity storage (in line with earlier findings, e.g. Barton and Infield 2004), but finds a net social welfare loss from introducing storage in the first place. One reason for this is that average electricity prices increase with the introduction of storage. Relatedly, Carson and Novan (2013) and Linn and Shih (2019) show that introducing electricity storage increases CO_2 and SO_2 emissions, and lowers NO_x emissions from electricity generation. The results are driven by low penetration rates of WWS technologies in combination with emission-intensive baseload plants. The implication is that electricity storage only moves energy from high price states to low price states where the marginal supply comes from the emission-intensive baseload plants. A similar rationale explains why Helm and Mier (2021) reach the conclusion that second best subsidies on storage may be negative.[5] Compared to these papers, we provide results in a model with greater spatial resolution, greater detail in the generation mix, and in particular, with higher penetration rates of WWS technologies. We find that electricity storage significantly lowers CO_2, SO_2, and NO_x emissions, and that it also lowers average electricity prices significantly. Furthermore, we show that in 2030 large-scale electricity storage becomes profitable at a cost of storage of around 250$/kWh compared to 160$/kWh in Linn and Shih (2019).

While our description of the electricity and district heat system is detailed in many aspects, we abstract from modeling certain essential parts of the system. First, we do not include a balancing market for electricity. Second, we disregard the issue of market power by assuming perfectly competitive markets. Third, we do not consider unit commitment or ramp-up constraints of power plants. While this is in line with many other larger models (see e.g. Balyk et al. 2019), this simplification overestimates the flexibility of the system which in turn will tend to underestimate the potential of storage (Diaz et al. 2019). Fourth, trajectories for investments in new generation capacity and transmission lines are currently taken as given. Future research will investigate the effects of endogenizing investments in new capacity.

The remainder of the paper proceeds as follows: Section 2 presents our convex optimization framework for storage technologies and provides a structural estimation of parameters in the model. Section 3 shows how we aggregate the model non-chronologically, and how to adjust the storage optimization framework accordingly. Section 4 provides evidence on the baseline scenario of the model for the European electricity and Danish district heat systems. Section 5 investigates the potential of storage and its effect on key outcomes such as emissions and prices. Section 6 concludes.

[4] A 'greenfield' experiment assumes a clean slate and does not rely on the current status of the energy systems.

[5] The second best-optimum assumes that we cannot use efficient Pigouvian taxes. The theoretical model builds on an assumption of three energy sources: Clean intermittent, dirty dispatchable, and storage. As long as the marginal supply is "dirty", subsidies should remain negative.

2 A Convex Optimization Framework for Storage Technologies

The model for electricity and district heat systems solves for a decentralized, competitive equilibrium on eight electricity markets that are connected by transmission lines and two district heat markets. The supply of electricity and district heat is derived from individual plants' optimization. The model includes a range of technologies including simple condensation plants, a number of combined heat and power (CHP) plants, and heat pumps to name a few. Hourly demand is driven by habits and, for a share of consumers, by prices. This results in multinomial logit-like hourly demand functions, expressed in shares of yearly demand, that can simulate the effect of using flexible demand technologies. In short, the model solves for equilibria for all states within a year (indexed by h), for each year (t), on all markets (g) by identifying a vector of prices $\boldsymbol{P}_{t,g} = (p_{t,1,g}, ..., p_{t,H,g})$ that solves

$$\sum_{i \in \mathcal{I}_g} Y^i_{t,h}\left(\boldsymbol{P}_{t,g}\right) = Y^d_{t,h,g}\left(\boldsymbol{P}_{t,g}\right) + \sum_{g'} NX^{g,g'}_{t,h}\left(p_{t,h,g}, p_{t,h,g'}\right),$$

where $Y^i_{t,h}$ is the optimal generation decision for plant i, $\mathcal{I}_g$ indicates the set of plants that belong to market g, $Y^d_{t,h,g}$ is the demand function, and $NX^{g,g'}$ defines net exports from market g to g'.[6] In general, the model is based on a convex optimization framework such that maximizing profits/utility yields continuously differentiable decision rules $(Y^i_{t,h}, Y^d_{t,h,g}, NX^{g,g'}_{t,h})$. This paper focuses on the dynamic optimization problem facing plants with capacity for energy storage. A companion paper elaborates on behavior of non-storage technologies, trade in electricity, and short-run demand (Berg and Eskildsen 2021). Furthermore, the generation capacity of plants is currently exogenous; future research will implement endogenous investments in new generation capacity. Finally, the following sections deal with modeling storage technologies in general; in sections 4-5 we return to the model for Northwestern European electricity systems.

2.1 The Storage Model

The model includes a discrete set of plants with storage capacity, including hydroelectric power plants with pumped storage or reservoirs with inflow from adjacent rivers, batteries, and seasonal heat storage. All plants can be represented by the same analytical framework, but parameter values differ according to e.g. storage inefficiency and ability provide seasonal storage.

Formally, $\eta \in [0, 1]$ denotes the hourly rate of self-discharge, γ_d and γ_c the efficiencies of discharging and charging, respectively, and $\overline{c}$ the marginal short run cost of operation and maintenance. Furthermore, let Y_h denote the hourly generation with $\underline{Y}$ and $\overline{Y}$ denoting lower and upper generation constraints, respectively. Similarly, S_h is the level of stored energy ultimo hour h with 0 and $\overline{S}$ denoting the lower and upper bounds on energy storage. Given an exogenous path of

[6] We only include trade across geographical markets for electricity that are connected by transmission lines.

hourly energy inflow, W_h, and prices, p_h, the traditional linear programming problem of a storage plant can be defined as

$$\max_{\{Y_h\}_{h=1}^{H},\{S_h\}_{h=1}^{H-1}} \sum_{h=1}^{H} \beta^h \left[(p_h - c)\, Y_h\right] \tag{1a}$$

$$\text{s.t.} S_h = \eta S_{h-1} + \gamma\left(W_h - Y_h\right), \qquad \gamma(x) = \begin{cases} x\gamma_c, & x > 0 \\ x/\gamma_d, & x \leq 0 \end{cases} \tag{1b}$$

$$Y_h \in [\underline{Y}, \overline{Y}], \tag{1c}$$

$$S_h \in [0, \overline{S}], \tag{1d}$$

$$S_0 = S_H \geq 0, \text{ given}, \tag{1e}$$

where β is an hourly discount factor, (1b) is the law of motion[7], and the periodicity constraint in (1e) is a standard transversality condition (see e.g. Renaldi and Friedrich 2017; Zerrahn et al. 2018).[8]

Traditionally, problems as the one in (1) are solved as linear programs or mixed-integer linear programming problems if binary variables are needed to e.g. model ramp up constraints/costs or differences in charge and discharge efficiencies. We offer a different approach based on a convex optimization framework:

Assumption 1 (Convex and chronological optimization)**.**
The storage plant solves:

$$\max_{\{Y_h\}_{h=1}^{H},\{S_h\}_{h=1}^{H-1}} \sum_{h=1}^{8760} \beta^h \left[p_h Y_h - C(Y_h, S_h)\right], \tag{2a}$$

$$S_h = \eta S_{h-1} + \zeta\left(W_h - Y_h\right) \tag{2b}$$

$$S_0 = S_H \geq 0, \text{ given}, \tag{2c}$$

where

i. marginal costs are increasing in both Y_h, S_h in a separable manner such that

$$\frac{\partial C}{\partial Y_h} = g_y\left(\frac{Y_h - \underline{Y}}{\overline{Y} - \underline{Y}};\ \bar{c}, \Sigma_y\right) - \kappa_y, \qquad \kappa_y \equiv g_y\left(-\frac{\underline{Y}}{\overline{Y} - \underline{Y}}\right) \tag{2d}$$

$$\frac{\partial C}{\partial S_h} = g_s\left(\frac{S_h}{\overline{S}};\ 0, \Sigma_s\right) \tag{2e}$$

where $g_x : [0,1] \to \mathbb{R}$ for $x = y, s$ are inverse flexible sigmoid type functions with characteristic parameters $(\bar{c}, \Sigma_y)$ and $(0, \Sigma_s)$ respectively (see definition 1), and κ_y serves the purpose of centering the function such that $\partial C / \partial Y_h = \bar{c}$ when $Y_h = 0$;

ii. The function $\zeta : \mathbb{R} \to \mathbb{R}$ is monotonically increasing with characteristic parameter $\sigma_\zeta \in \mathbb{R}_+$.

[7] Also referred to as the serial charge constraint.

[8] As the model is solved for several years (2019-2040), the reservoir constraint could be that $S_{H,t} = S_{0,t+1}$. This would, however, increase the complexity of the model significantly as the computation of equilibria in one year now depends on equilibria in other years.

For $\sigma_\zeta > 0$ the function is continuously differentiable, and

$$\lim_{\sigma_\zeta \to 0} \zeta\left(W_h - Y_h; \sigma_\zeta\right) = \begin{cases} x\gamma_c, & x \geq 0 \\ x/\gamma_d, & x < 0 \end{cases} \tag{2f}$$

iii. The values $\gamma_c - 1/\gamma_d$ and σ_ζ are sufficiently small such that for all $Y \in [\underline{Y}, \overline{Y}]$.[9]

$$\frac{\partial^2 C}{\partial Y^2} \geq -\frac{\zeta''(W-Y)}{\zeta'(W-Y)} \frac{\partial C}{\partial Y} \tag{2g}$$

Assumption 1 is convenient for a number of reasons: First, we note that letting the smoothing parameters $\Sigma_y, \Sigma_s, \sigma_\zeta \to 0$, the problem in (2) is equivalent to the original problem in (1). Second, for positive values of smoothing parameters, the optimization problem is globally concave on the convex set $(Y_h, S_h) \in [\underline{Y}, \overline{Y}] \times [0, \overline{S}]$. This ensures that the first order conditions uniquely identify the global optimum (see appendix B). Third, the set of equations that identify the optimal storage/dispatch behavior is particularly easy to solve as it can be topologically ordered. To see this, consider the set of first order conditions:

$$Y_h = \underline{Y} + \left(\overline{Y} - \underline{Y}\right)(g_y)^{-1}\Big(p_h - \overline{c} - \zeta'\lambda_h + \kappa_y;\ \Sigma_y\Big) \tag{3a}$$

$$\lambda_h = \beta\eta\lambda_{h+1} - g_s\left(\frac{S_h}{\overline{S}};\ \Sigma_s\right) \tag{3b}$$

$$S_h = \eta S_{h-1} + \zeta\left(W_h - Y_h\right), \tag{3c}$$

where λ_h measures the shadow value of stored energy. To solve the system of equations, form an initial guess on λ_1 and recall that we know the initial reservoir level S_0. Now, given λ_1, we can compute Y_1 from (3a), then S_1 from (3c), and λ_2 from (3b). In other words, the problem can be boiled down to identifying one variable, λ_1, such that solving forward as explained above yields a terminal level $S_H = S_0$.[10]

Assumption 1 can alternatively be interpreted as a penalty approach to handling generation and storage capacity constraints: Whenever the constraints are approached the related marginal costs of generation/storage increase.[11] Consider for instance the shadow value of storage in equation (3b). Whenever the level of storage is nowhere near the capacity constraints, the term g_s evaluates to approximately zero, and the optimal path of the shadow value of storage simply reflects impatience (β) and efficiency of the storage technology (η). If, however, storage approaches its capacity $\overline{S}$, the term g_s increases,which translates into a lower value of storage λ_h. This incentivizes the plant to increase generation in the given hour.

[9] Appendix B.1 shows that the set of permissible parameter values for which the condition holds is non-empty. For a specific type of smooth piecewise linear function, there exists a $\sigma_\zeta > 0$ for all values of $\gamma_c, \gamma_d > 0$ where the condition holds.

[10] Put differently, again: Given λ_1 the system is a directed acyclical graph.

[11] See appendix B.2 for a clarification on why the convex optimization problem can be seen as a penalty function method with specific penalty terms.

2.2 Estimation of storage model

To evaluate how well the convex optimization framework captures relevant incentives for storage plant managers, we consider the case of Norwegian hydroelectric power plants with reservoirs and adjacent rivers that provide an exogenous inflow of water. While the empirical sections 4-5 focus on the case of Denmark, electricity storage is currently not employed in a significant scale. We focus on the neighboring country Norway due to data availability and its historical position as the largest provider of hydroelectric power in the EU.

We collect data on production, spot prices, and reservoir levels for 2019. Hourly hydroelectric generation and spot prices are retrieved from ENTSO-E (2017a) and ENTSO-E (2017b). Reservoir levels are available at a weekly frequency at ENTSO-E (2017c). The most disaggregated level with consistent data is the at the bidding zone level (5 geographic areas: NO1-NO5). As there is no publicly available data for inflow levels, we aggregate hourly generation to the weekly level and back out the implied inflows using the simplified law of motion $S_w = S_{w-1} + W_w - Y_w$, where W_w is the accumulated water inflow in week w.[12] We then uniformly allocate the weekly accumulated inflow onto each hour of the week and calculate the implied hourly reservoir level. The final data set is a balanced panel data set with N=5 electricity areas and H=8760 hours for 2019, leaving us with a total of 43,800 observations. Appendix C provides a detailed account of the data with descriptive statistics.

We estimate the model in (3) structurally, assuming that marginal generation costs follow an inverse, cumulative normal distribution with mean $\bar{c}$ and standard deviation σ_y, and that the marginal cost of approaching capacity bounds on the reservoirs is a smooth piecewise linear function. Specifically, for each of the hydroelectric power plants, we use the system of equations:

$$Y_h = \underline{Y} + (\overline{Y} - \underline{Y})\Phi\left(\frac{p_h - \bar{c} - \lambda_h}{\sigma_y}\right) \tag{4a}$$

$$\lambda_h = \beta\lambda_{h+1} + \frac{1}{2}\left(\kappa\left(\overline{S} - 2S_h\right) + \left(\kappa - \frac{2\gamma}{\overline{S}}\right)\left(\sqrt{S_h^2 + \delta^2} - \sqrt{(S_h - \overline{S})^2 + \delta^2}\right)\right), \tag{4b}$$

$$S_h = S_{h-1} + W_h - Y_h, \tag{4c}$$

$$S_H = S_0, \qquad S_0 \geq 0 \text{ given}, \tag{4d}$$

where Φ is the standard normal cumulative distribution and $\kappa, \gamma, \delta > 0$ are parameters in the smoothed piecewise linear function (see appendix C). We estimate structural parameters using maximum likelihood, formulated as a constrained maximization problem similar to a mathematical program with equilibrium constraints (MPEC, Su and Judd 2012). Appendix C outlines the formal estimation procedure and provides details on estimated parameters as well as supplementary data and figures.

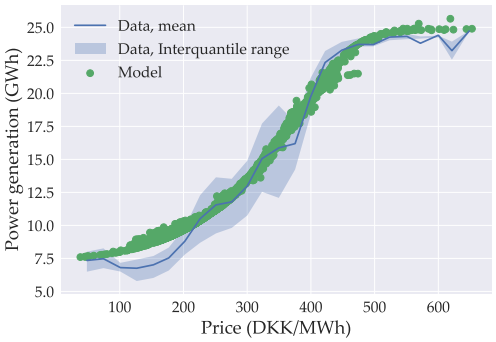
Figure 2.1 illustrates the fit of the estimated model. Panel (a) illustrates the upward sloping aggregate supply curve and indicates that hydroelectric plants ramp up production when prices

[12] This simplified law of motion that corresponds to $\eta = \gamma = 1$ in our theoretical framework is relatively standard for the specific type of hydroelectric power plant.

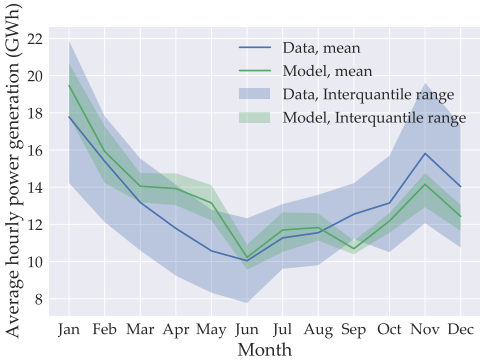
increase. Furthermore, the overall shape of the curve is captured well. Panel (b) illustrates the hourly average power generation for each month and shows that the model fits the general seasonality quite well, where production (on average) is highest during colder months and lower during the warm summer months.

Figure 2.1: Model fit of aggregate Norwegian hydroelectric generation, 2019

(a) Aggregate supply curve (Y_h)



(b) Seasonality in aggregate generation



Note: To minimize the noise in the plot, the "mean" label represents prices collected in intervals of 25 DKK/MWh, and evaluates the mean. The "interquantile range" indicates observations between 25% and 75% in the distribution.

Another validation of the model is that our estimated parameters are intuitively reasonable: We estimate the hourly discount factor to be 1, the average short run cost of generation ($\bar{c}$) is estimated to be around 3 DKK/MWh, and the total generation capacity of Norwegian plants is estimated to be 25.9 GW compared to the 27.7 GW as listed in ENTSO-E (2017d).

Finally, recall that the conventional linear programming framework is nested in our framework for low values of smoothing parameters. The estimation shows that at an aggregate level, the convex optimization problem with increasing marginal costs of generation is a better representation than the linear programming one ($\sigma_y \gg 0$). Indeed, in order to replicate the shape of the aggregate supply curve, the level of smoothing on generation costs, σ_y, is estimated to be roughly between 51-137 depending on what region of Norway we consider. Even for the lowest value of $\sigma_y \approx 50$ this indicates that the storage plant gradually starts emptying its reservoir when the current market price is between 0-100 DKK/MWh higher than the continuation value of energy in the subsequent hour.

3 Aggregation of hours within a year

While the storage model presented above is easy to solve, solving the larger model for the Danish electricity and district heat systems requires identifying equilibria in several markets from 2019-2040, which quickly becomes infeasible with a full hourly resolution. To reduce the dimensionality of this model, we aggregate hourly data into a number of states (or clusters) using a combination of conditional inference trees, a K-means algorithm, and a few *post* clustering adjustments. As the idea of aggregating hourly data is not novel (Wogrin et al. 2016; Gabrielli et al. 2018; Reguant

2019), the following briefly outlines our approach. Appendix D provides additional information and supplementary figures to illustrate the effects of the aggregation.

3.1 Clustering intra-yearly variational data

The model for the Danish electricity and district heat system draws on exogenous variables that represent hourly variation in (i) the productivity of technologies based on intermittent energy sources, (ii) short-run demand for electricity and district heat in various geographical regions, (iii) water inflow for hydroelectric power plants, and (iv) availability of transmission lines for trade in electricity.

Let $\overline{Y}_{t,i}$ denote the hourly average generation capacity for plant $i \in \mathcal{I}$ in year t. The hourly generation capacity is defined as $\overline{Y}_{t,h,i} = \nu_{\tau,h}\overline{Y}_{t,i}$, where $\nu_{\tau,h}$ indicates the hourly variation for a plant of intermittency type τ. Let $\nu_{g,h}$ denote the variation in hourly demand habits in the geographical region g, and $\nu_{(g,gg,h)}$ the variation in the capacity of transmission lines enabling exports in electricity from region g to gg.[13] Based on the exogenous variation in hourly demand and generation capacity of intermittent plants, we can measure the residual demand for electricity and district heat as

$$r_{t,g,h} = Y^d_{t,g}\nu_{g,h} - \sum_{i \in \mathcal{I}_g} \overline{Y}_{t,i}\nu_{\tau,h}, \tag{5}$$

where $Y^d_{t,g}$ is the yearly level of demand in region g, and $\mathcal{I}_g$ is the subset of plants in region g.[14]

The objective of the aggregation is to define variables $\tilde{\nu}_{x,k}$ for $x \in \{g, (g, gg), \tau\}$ defined over a smaller set of clusters K, rather than hours, such that key model characteristics are maintained. Determining what constitute key characteristics, however, depends on the analysis in question. An idealized but computationally greedy solution is to evaluate the fit of the clustered model in terms of how well the exogenous parameters fit, i.e. by comparing $\tilde{\nu}_{x,k}$ and $\nu_{x,h}$ directly. A second metric is to evaluate the model's ability to replicate the clustered version of the residual demand curve in (5). Evaluating the clustered model in terms of this objective allows us to use a lower value of K, but might miss important equilibrium dynamics of the model, such as the availability of transmission lines or the access to cheap stored energy. Instead, we validate the clustered model based on the more high level objective of intra-yearly prices. This is a less greedy solution than evaluating the fit of all individual $\tilde{\nu}_{x,k}$ and it takes important equilibrium dynamics into account.

We start by using conditional inference trees to identify seasonal breaks in the variational data (Hothorn et al. 2006). The seasonal breaks are used to ensure a realistic modelling of seasonal storage behavior (see appendix C.5). Next, we apply a K-means algorithm to cluster intra-seasonal data into representative states. The K-means algorithm assigns each hour to a cluster by minimizing the Euclidean distance between hourly data and cluster-averages. As we focus on making sure

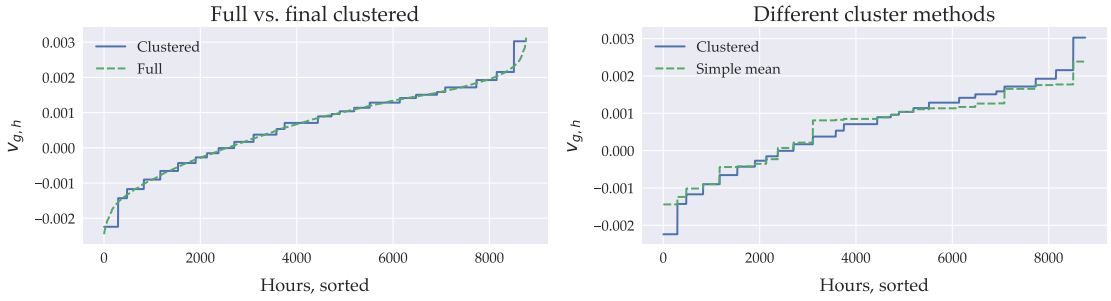
[13] In the empirical section 4 we include around 100 different variables: 53 types of variation in generation capacity, 10 geographical regions, and 36 transmission lines.

[14] This measure of 'residual demand' does not take into account that both intermittent plants and short run demand may respond to prices.

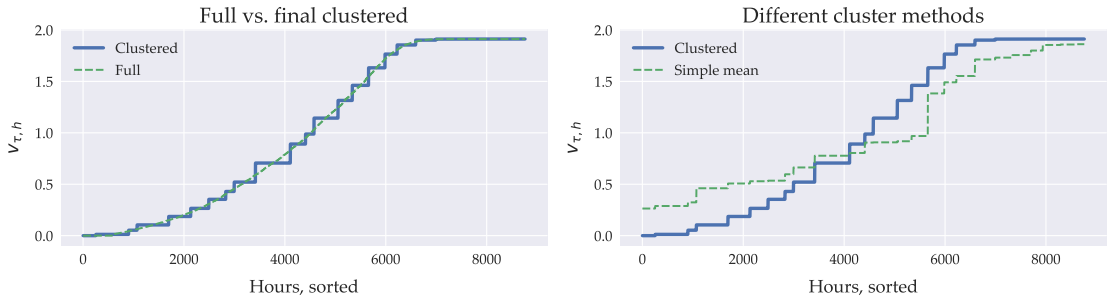
that the clustered model produces a realistic price variation, we use residual demands to identify the clusters. Next, we adjust the model data $\tilde{v}_{x,k}$ *post* clustering to target price variation: We sort clustered ($\tilde{v}_{x,k}$) and hourly variables ($v_{x,h}$) and fit the levels in the sorted $\tilde{v}_{x,k}$ to reflect the full range of values in $v_{x,h}$. As we do this for each variable separately, we potentially compromise realistic correlations between hourly variables in order to obtain a better fit of volatility in the clustered data. Finally, we adjust the clustered data in order to specifically target tail observations, and fit the variation in residual demands.[15]

Figure 3.1 shows how our adjusted clustering algorithm aggregates the hourly data: The 'Full' lines indicate the sorted hourly data prior to clustering, the 'Simple mean' lines indicate the effect of K-means clustering, and the 'Clustered' lines indicate our final version with *post* clustering adjustments. Panel (a) shows that the simple K-means approach averages out the tails of the distribution in residual demand; our adjustments fix this issue. Panel (b) shows that, even though the K-means algorithm does not target variation in wind directly, it captures a fair share of the variation. Again, our proposed adjustments ensure that the full distribution of wind variation is captured in the 'Clustered' version. Thus, even though we only target the residual demands in the K-means algorithm, the clustered data represents the full distribution for all variables.

Figure 3.1: The effect of clustering



(a) Residual demand



(b) Danish wind technology

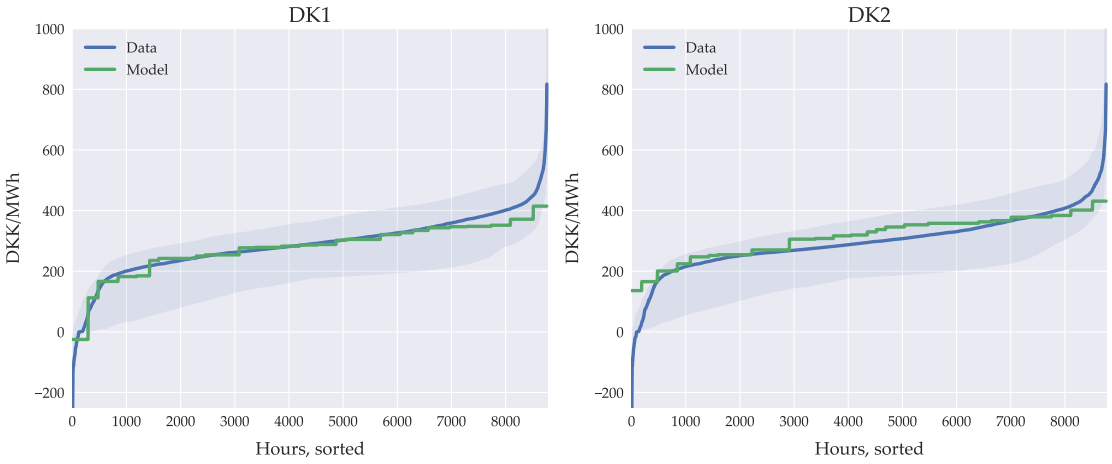
Note: The figure shows the effect of clustering with $K = 25$ and $S = 4$. The 'clustered' path uses the preferred clustering method that includes both fitting of model variation including residual demand, and adjustments to capture tail probabilities.

Figure 3.2 illustrates the price variation predicted by the model in the base year compared against observed prices for the two Danish electricity markets. The figure shows that while the

[15] See appendix D for more details on this.

model does not capture the thin tails of the price densities, prices are generally close to the ones observed in data. Importantly, we show in appendix D that using a simple K-means approach and increasing K to as high as 1,000 states does not generate a better fit in price variation than our preferred clustering method with $K = 25$.

Figure 3.2: Model fit of intra-yearly prices, 2019



Source: Energy Data Service (2021a).
Note: The shaded area indicates the interval of minimum/maximum prices observed between 2011-2020 in the data. The solid line indicates observations from 2019. DK1 and DK2 are abbreviations for the two electricity price regions Western and Eastern Denmark.

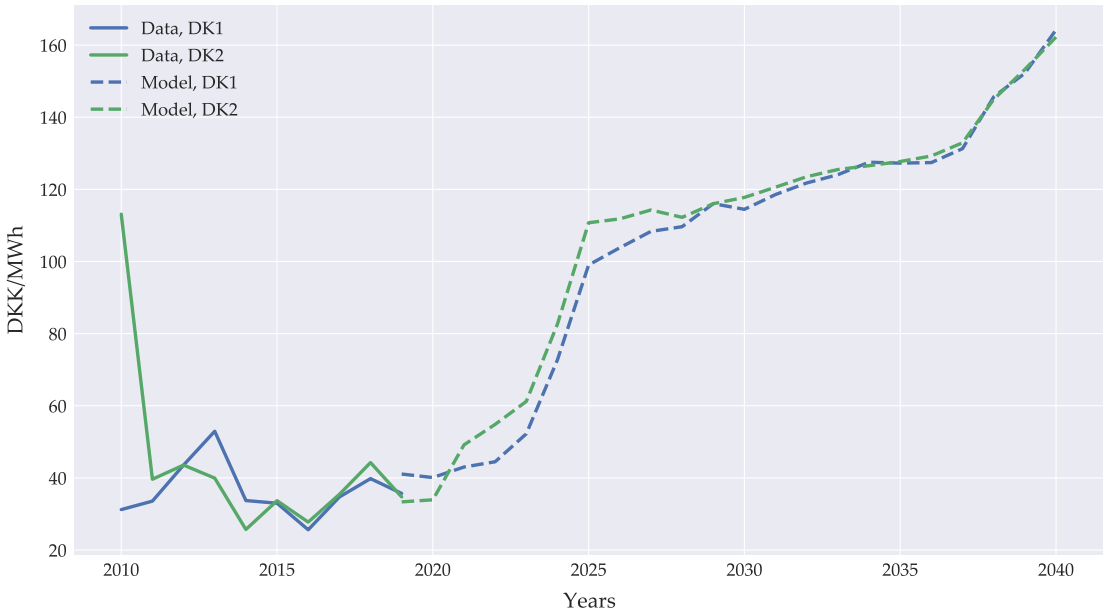
Furthermore, the model provides a very good fit of a second high level objective: Realistic technology-specific prices. Figure 3.3 shows the downlift for wind technologies defined as the difference between yearly average prices and the average price received for wind technologies. As argued by Hirth et al. (2015), the downlift for wind technologies is a key characteristic of intermittency. Because the marginal generation cost of wind is low, electricity prices are pushed towards zero when wind productivity is high. This implies that wind technologies receive, on average, a lower price of supplied electricity. Figure 3.3 confirms that the model is able to replicate the observed downlifts in the base year. Furthermore, the model predicts that the downlift is increasing over time, indicating a higher degree of intermittency in the electricity system; section 4 returns to this issue.

3.2 Clustered representation of storage plants

For the most part, aggregation of hours into representative states is straightforward. We simply scale capacities appropriately and the equilibrium concept remains the same: One price in each state balances demand and supply. The clustering, however, is inherently non-chronological, which poses a problem when representing storage plants' law of motion for stored energy. The following proposes a number of ways to handle this.

Full hour-to-cluster mapping

Figure 3.3: Downlift on wind, historic data and model



Source: Energy Data Service (2021a), Energy Data Service (2021b).
Note: The dashed lines are from the baseline scenario of the model of the Danish electricity and district heat systems, cf. section 4.

From the K-means algorithm, we can establish a mapping that specifies which hours are assigned to what cluster $\mathcal{S}_k$. In this way, we can reestablish chronology and simply solve the storage plant's problem as in equations (3). This is essentially a version of the multi-grid solution that only solves for equilibria on the grid of clusters, but the storage technologies' supply comes from optimization on the hourly grid such that

$$Y_k = \sum_{h \in \mathcal{S}_k} Y_h.$$

While this approach does not reduce the size of the storage problem, it still simplifies the problem considerably as it only solves equilibrium constraints for each cluster.[16] For our purposes, keeping the full hourly resolution for storage plants alone represents roughly 7 million constraints.[17] While this is feasible, it remains an inconveniently large system to solve alongside the rest of the model. Thus, we only use this method to check the validity of other solution methods.

State-average aggregation

An alternative is to use the identified hour-to-cluster mappings to define a Markov-state representation of the model. We can establish a transition probability matrix $\boldsymbol{P}$ where element $P_{i,j}$

[16] In regions where storage technologies make up a large share of aggregate supply this simplification may become important, as this fixes the price across all hours $h \in \mathcal{S}_k$; thus, if equilibrium prices react strongly to the supply from storage technologies, a full hourly model may imply significant price variation within clusters k that are otherwise similar. This can be remedied by including auto-correlation moments in the K-means clustering procedure.

[17] The model includes 13 plants that optimize over 8760 hours for 21 years. With a minimum of three variables included per hour this is 7,174,440 constraints.

specifies the probability of transitioning from cluster i to j. Similarly, we can define $\boldsymbol{P}^{-1}$ as backward looking probabilities, with $P^{-1}_{i,j}$ denoting the probability that the preceding cluster was type j given that the current cluster is i. Using cluster-averages, the storage problem can be reduced and solved on the grid of clusters as well. Specifically, we replace the system of equations in (3) with

$$y_k = \underline{Y} + (\overline{Y} - \underline{Y}) \, (g_y)^{-1} \left(p_k - \overline{c} - \zeta' \lambda_k + \kappa_y;\ \Sigma_y \right) \tag{6a}$$

$$\lambda_k = \eta\beta \sum_j P_{k,j} \lambda_j - g_s \left(\frac{S_k}{\overline{S}};\ \Sigma_s \right) \tag{6b}$$

$$S_k = \eta \sum_j P^{-1}_{k,j} S_j + \zeta(W_k - y_h) \tag{6c}$$

$$Y_k = n_k y_k, \tag{6d}$$

where y_k measures the cluster-average generation and n_k the number of hours that is mapped to cluster k.

Compared to the full hourly model this is a rather drastic simplification in two important ways: First, when the storage plant chooses the optimal generation in cluster k, the shadow value of storage is defined as the expected value of future states $\sum_j P_{k,j}\lambda_j$. Thus, the storage plant inherently faces uncertainty about future states that we do not include in the hourly model presented in section 2. Thus, compared to the full hourly model, this results in a less flexible representation of storage plants: As a plant cannot predict with certainty when it enters states with high/low prices, it will tend to smooth out production more than what is optimal in the deterministic model. Secondly, the storage constraint is only checked for cluster average values, $\sum_j P^{-1}_{k,j} S_j$. In this respect, the cluster-average model results in an overly flexible representation of storage plants. Consider, for instance, the case where a cluster with medium prices represents some hours that succeed a spell of low-price states, but also some hours that succeed a spell of high prices. In this case, the cluster-average model may not be constrained by storage capacity, while the full hourly model would predict that some hours are constrained by the lower bound on storage and some by the upper bound.

State-average aggregation with long-run uncertainty

The state-average approach simplifies the storage problem significantly, but leaves two primary issues: Uncertainty of future states and averaging out of potential violations of storage constraints. To resolve the issue of uncertainty, we extend the system of equations to be defined over cluster tuples (k, j). Thus, instead of only solving for a cluster-specific generation variable, y_k, we solve for the optimal generation for all combinations of $y_{k,j}$:

$$y_{k,j} = \underline{Y} + \left(\overline{Y} - \underline{Y}\right) \left(g_y\right)^{-1} \left(p_k - \overline{c} - \zeta' \lambda_{k,j} + \kappa_y;\ \Sigma_y \right) \tag{7a}$$

$$\lambda_{k,j} = \eta\beta \sum_i P_{j,i} \lambda_{j,i} - g_s \left(\frac{S_{k,j}}{\overline{S}};\ \Sigma_s \right) \tag{7b}$$

$$S_{k,j} = \eta \sum_j P_{k,j}^{-1} S_{j,k} + \zeta \left(W_k - y_{k,j} \right) \tag{7c}$$

$$Y_k = n_k \sum_j P_{k,j} y_{k,j}. \tag{7d}$$

In this formulation the storage plant decides the optimal generation in state (k, j) knowing that the succeeding cluster will be of type j. When forecasting beyond the future state j, however, the storage plant has to rely on expected values once again cf. equation (7b).

The representation increases the dimensionality of the system considerably from K to potentially K^2 cluster-tuple states. In our application with $K = 25$ we end up with 116 cluster-tuple states. We note that it is straightforward to solve the entire model on this extended grid of 116 cluster-tuple states. Doing so resolves another issue that also exists in the related literature on storage modeling, including the system states framework and the multi-grid approaches: If we only identify one equilibrium state per cluster, we do not take into account that prices are endogenous to the supply from storage technologies.[18] Naturally, defining the entire model, including equilibrium prices, over cluster-tuple states (k, j) resolves this inconsistency.

The system of equations in (7) is dramatically faster to solve than the full hourly model, but may come at the cost of accuracy. In particular, the approximation is sensitive to the cyclicality of data as storage levels ($S_{k,j}$) still rely on averages (see appendix E).

Validation of the storage approximation

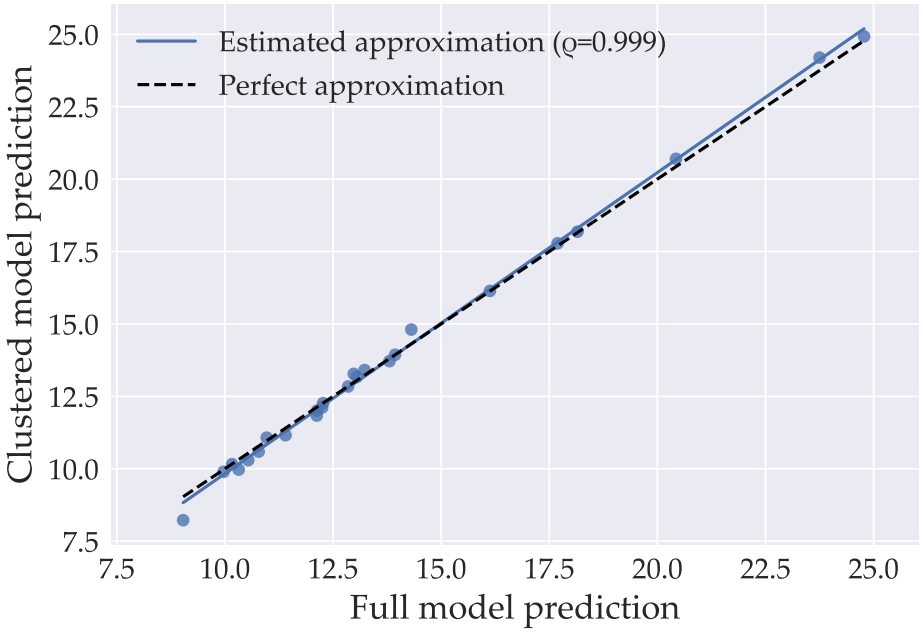
To evaluate the storage approximation, we perform two validity checks. First, we evaluate the model's ability to predict seasonal storage behavior using the estimated model for Norwegian hydroelectric power plants in section 2.2. Figure 3.4 compares the optimal generation in the hourly model to its counterpart with $K = 25$ clusters. The clusters are identified by considering hourly variation in water inflow and equilibrium prices with three seasonal breaks. The plot shows that the implied cluster-average generation in the hourly model almost coincides with the one from the clustered model.

Second, in the empirical analysis of electricity storage in Denmark (section 5), we compare results for the clustered and a full hourly storage model. To avoid running the entire model with an hourly resolution, we first solve the model with the clustered approximation and endogenous prices. Then, we take the clustered price variation as given and solve the hourly storage model. Figure 3.5 compares the clustered and the full hourly model for various sizes of batteries. We note that compared to the hydro model the approximation is less accurate, in particular when the batteries have shorter duration. Nonetheless, even for small values of duration, the storage approximation does not feature a significant bias. Thus, while the clustered approximation should always be used with some caution, we conclude that it provides reasonable results in our setting

[18] In the terminology of the cluster-tuple states this corresponds to assuming that the equilibrium price only depends on the current state, k, even though the supply from storage technologies depends on the next cluster-state as well. In other words, we implicitly assume that:

$$p_k = p_{k,j} = p_{k,l}, \ \forall j, l, \qquad \text{and} \qquad y_{k,j} \neq y_{k,l}.$$

Figure 3.4: Optimal generation in full and clustered model, Norwegian hydro-model



Note: ρ is the Pearson correlation coefficient between optimal generation in the full and clustered model.

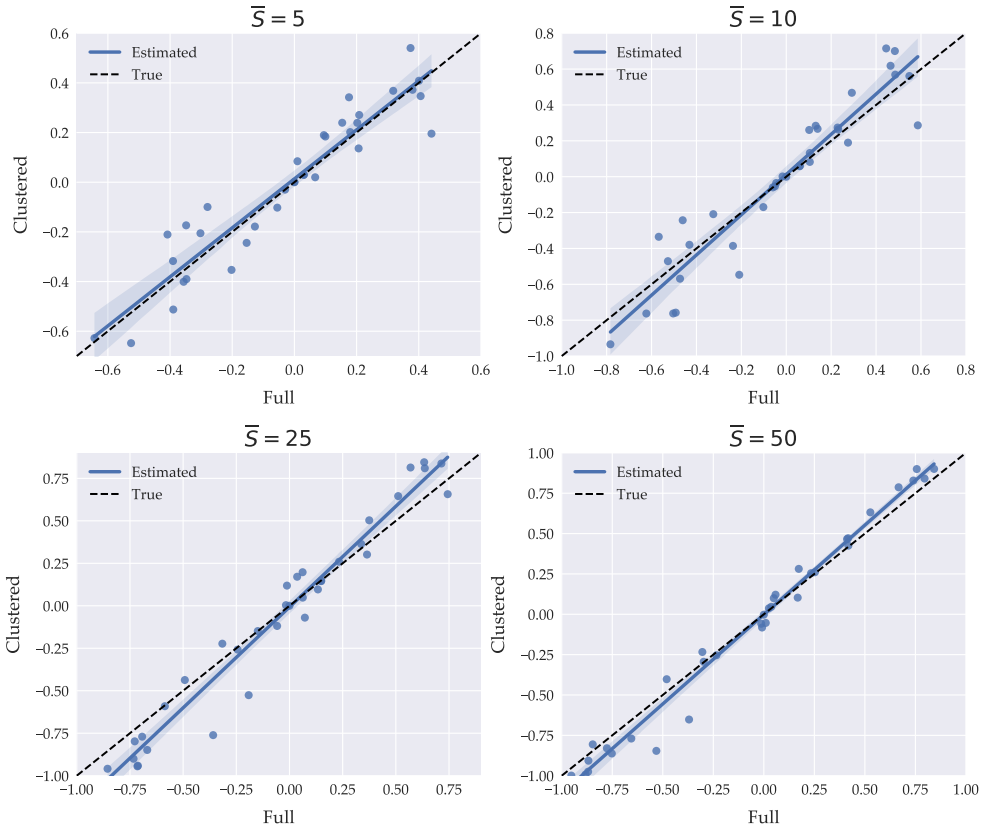
and makes the model computationally tractable.

4 The electricity system in and around Denmark

For the empirical part, we focus on Denmark and use the model for electricity and district heat systems described in Berg and Eskildsen (2021). Before we investigate the effects of introducing electricity storage in Denmark in section 5, we outline important characteristics of the current electricity system and relevant features of the baseline scenario from 2019-2040.

The electricity market in Denmark consists of three layers: A wholesale, distribution, and a retail market. The main component of this market is the wholesale market on the Nordic electricity exchange, Nord Pool. Through Nord Pool and the European Power Exchange (EPEX SPOT), Denmark is connected to a relatively well integrated transmission network that includes the Nordic countries Sweden, Norway, and Finland, and Central/Western European countries such as Germany, the Netherlands, and Great Britain (ENTSO-E 2021). Thus, while our focus is on Denmark, the model solves for equilibria on 8 selected electricity markets in this integrated market: Two Danish regions (West and East), Nordic countries (Sweden, Norway, Finland), and three Central/Western European regions (Germany, Austria, Luxembourg (DEATLU), the Netherlands, and Great Britain). Furthermore, while we generally focus on the potential for electricity storage, the model solves simultaneously for equilibria on Danish district heat markets. This is mainly due to the prominent role of combined heat and power plants in Denmark. We reiterate that the model focuses on the wholesale market and is simplified in a number of ways. We do not account for balancing markets nor do we generally consider the need for ancillary services. The implication

Figure 3.5: Optimal generation in full and clustered model, full model data



Note: The figure shows the cluster-average levels of generation by a Danish electricity storage plant in a model using the clustered approximation, and in one using the full hourly model, but with prices from the clustered approximation.

is that we only assess part of the potential value that stems from electricity storage, namely from price arbitrage. While the value from e.g. operating reserves may be significant, these markets can saturate quickly (Denholm, Jorgenson, et al. 2013). Thus, the profitability of large-scale battery investments are not likely to be affected by these considerations on the margin.

4.1 Data

The model is primarily informed by data collected and maintained by the Danish Energy Agency to use in their techno-economic model Ramses (Danish Energy Agency 2018a). The data contains highly detailed information on Danish electricity and district heat generating plants, as well as more aggregated data on plants operating in relevant neighboring countries. The data further contains projections of the future composition of plants in all the geographic areas. The projection for generation and transmission line capacities are based on scenarios published by ENTSO-E and ENTSOG.[19] Throughout the paper, unless otherwise specified, we take this path of generation

[19]Specifically, data is based on the scenario "National Trends" in the Ten-Year Network Development Plan 2020 and the Midterm Adequacy Forecast 2020, along with market statistics from ENTSO-E's Transparency Platform.

capacities as given. However, future research is dedicated to endogenizing investments in a way that captures the considerable regulations in this area.

The information on domestic plants is based on the national inventory database, "Energiproducenttællingen", with projections of future capacities based on the Danish Energy Agency's Energy and Climate Outlook 2021. We note that this projection is based on a frozen policy assumption, i.e. assuming that no new policy initiatives are introduced after ultimo 2020. In certain aspects, this baseline scenario may lead to overly pessimistic predictions for the value of electricity storage. For instance, while the Danish government has committed to a target of 70% reductions of greenhouse gas emissions by 2030, the frozen policy scenario only delivers a 55% reduction (Danish Energy Agency 2021). If emission reductions are achieved through electrification of other sectors (e.g. as suggested by Danish Council on Climate Change 2018), both the level of demand for electricity and the installed capacity of intermittent renewables is expected to increase significantly compared to our frozen policy baseline scenario. This would likely make electricity storage more valuable than the current baseline scenario suggests.

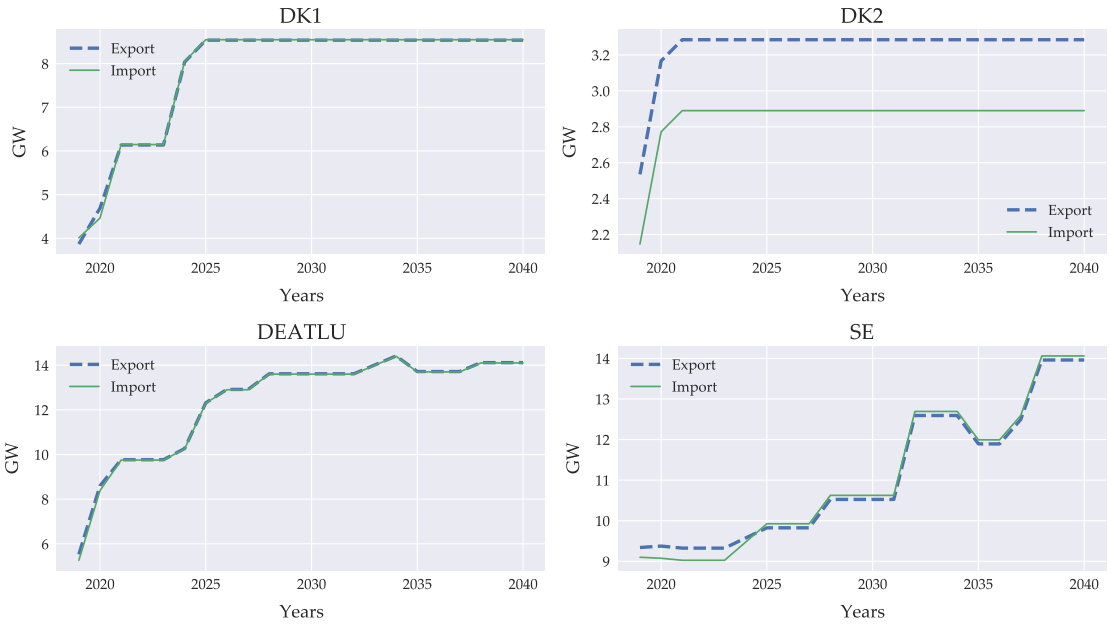
4.2 Baseline projections

Before we turn to simulating the potential of electricity storage in the Danish context, we highlight four important trends in the baseline scenario in our data related to investments in the transmission network, electricity storage technologies, renewable generation capacity, and dispatchable generation capacity.

As of 2021 the transmission network allows Denmark to trade with all of the countries included in the model, with the exception of Great Britain. However, significant expansions to the transmission network is underway (ENTSO-E 2021). Figure 4.1 shows the total access to exports and imports of electricity over time for a subset of regions. For instance, between 2019 and 2025 the export and import capacities are more than doubled for Western Denmark (DK1).

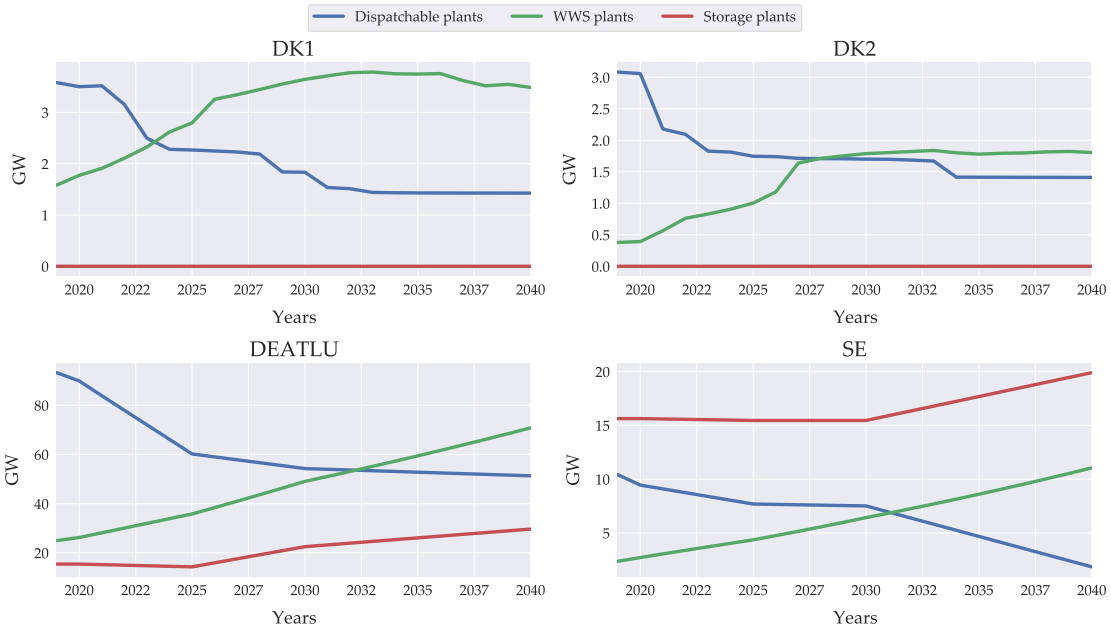
Furthermore, the generation capacities of various types of technologies are changing quite dramatically. Figure 4.2 illustrates the change in generation capacity for dispatchable, intermittent renewable (WWS), and electricity storage plants for selected regions. First, for all four regions, there is an increased share of WWS capacity, in particular for Denmark, Germany and The Netherlands. The implication of the higher penetration of WWS technologies is apparent in the residual demand curves (figure 4.3). For all regions the variation increases drastically over time to the extent that supply from WWS technologies by 2030 vastly exceeds demand in many hours. Second, all regions except Great Britain retire a large share of the generation capacity from dispatchable plants. And third, investments in electricity storage are quite heterogeneous across countries. In the Scandinavian countries (Norway, Sweden, Finland), the use of hydroelectric power with storage is critical already today. Through the 2030's, Sweden and Finland start investing in electricity storage technologies to supplement this capacity, whereas Norway steadily invests in electricity storage from 2020. To the south of Denmark, all countries invest quite heavily in electricity storage technologies, in particular in the region DEATLU that roughly doubles its available generation

Figure 4.1: Total Hourly Transmission Capacity



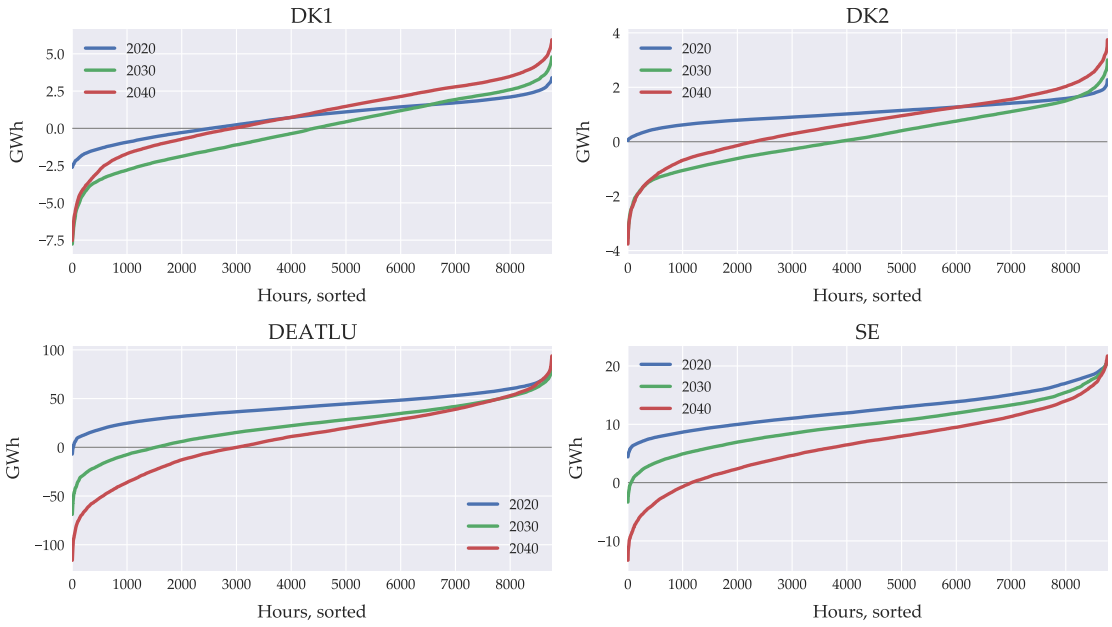
Note: The abbreviations denote the geographic areas Western Denmark (DK1), Eastern Denmark (DK2), Germany, Austria, and Luxembourg (DEATLU), and Sweden (SE).
Source: Own calculations based on data from the Danish Energy Agency.

Figure 4.2: Hourly average generation capacity



Note: The dispatchable supply is defined as the sum of all plants except wind, water, and sunlight (WWS) and storage technologies. Here, hydroelectric power plants with storage is included in storage category and run-off-river hydroelectric power is included in the WWS category.
The hourly average generation capacity is defined as the generation capacity times the yearly capacity factor. Thus, for a wind turbine with a generation capacity of 1 MW and a capacity factor of c, the hourly average generation is c MW.

Figure 4.3: Sorted residual demand on geography and years



Source: Own calculations based on data from the Danish Energy Agency.

capacity from 15 to 30 GW. In Denmark, however, there are currently no plans to introduce electricity storage on any significant scale.

The implication of these trends in investments is that intra-yearly price variation differs a lot from region to region. Figure 4.4 shows that for the case of Denmark, where WWS generation capacity increases, dispatchable generation capacity is retired, and there is no significant investment in electricity storage technologies, prices are significantly lower for most of the year, but with peak prices that increase dramatically over time. At the other end of the scale, Sweden invests significantly in both electricity storage and WWS technologies to experience lower prices all year round.

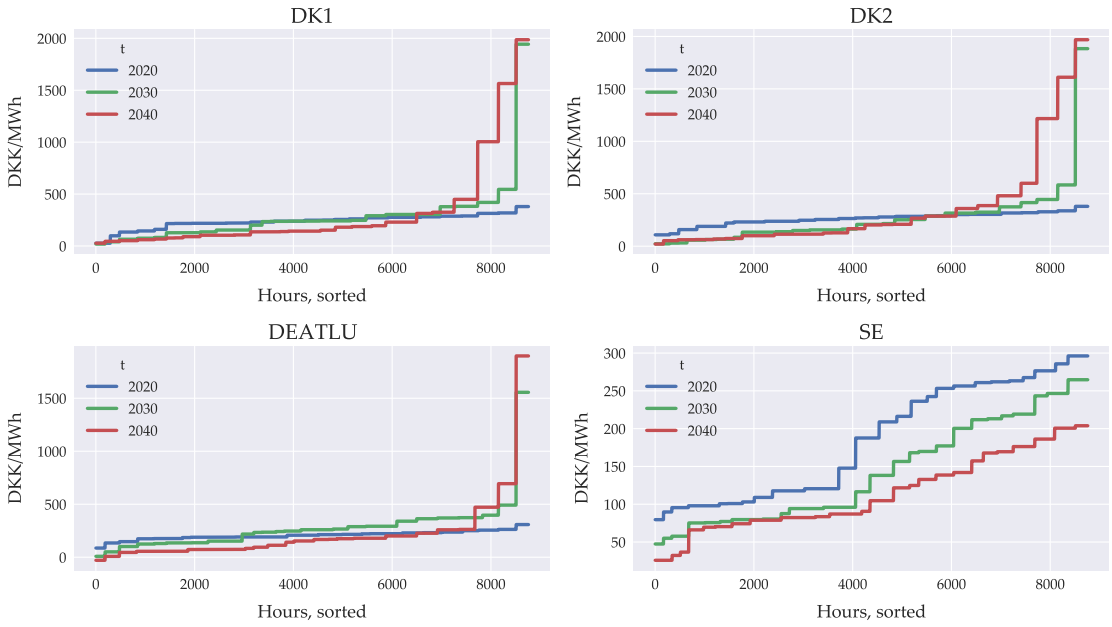
5 The effects of electricity storage in Denmark

To evaluate the effect of electricity storage in the system, we vary the level of generation capacity for new electricity storage plants from 0-4 GW in Western Denmark and 0-2.5 GW in Eastern Denmark, reflecting the relative size of demand.

5.1 The profitability of electricity arbitrage

We investigate three types of batteries in terms of storage-to-generation capacities: 2, 4, and 10 hours duration. While the larger batteries have higher investment costs, they are also less limited by the serial charge constraint. To assess the profitability of each type, we measure the sum of discounted short run profits (Π) assuming an annual discount rate of 8%, a technical lifetime of

Figure 4.4: Intra-yearly price variation, baseline scenario



20 years, and a variable cost of discharge of 2.5 \$/MWh.[20] The discounted profits are:

$$\Pi_t = \sum_{j=t}^{t+20} (1.08)^{t-j} \left[\sum_k \left(Y_k p_k - 2.5 Y_k \mathbb{I}_{Y>0} \right) \right], \qquad \mathbb{I}_{Y>0} \equiv \begin{cases} 1, & Y_k > 0, \\ 0, & Y_k \leq 0 \end{cases}$$

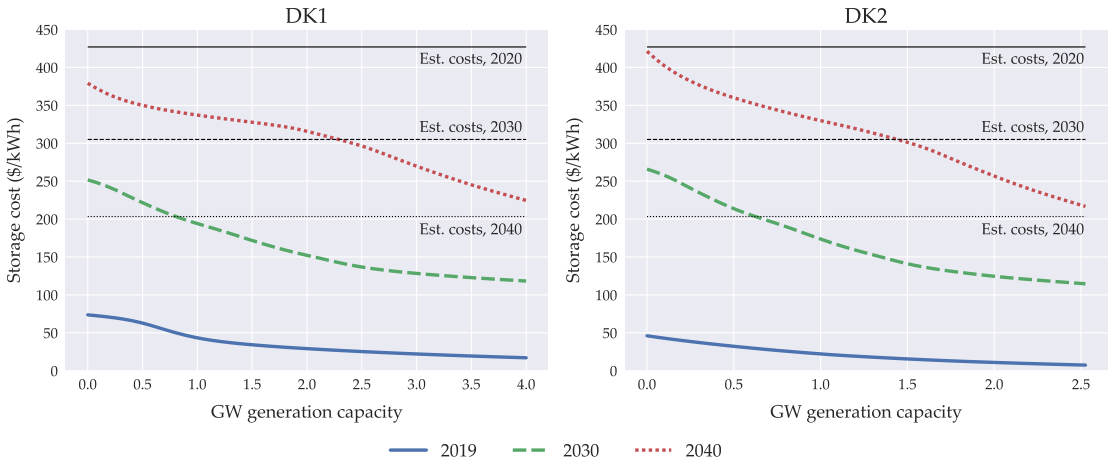
For the investment costs we use estimates from Baxter et al. (2020) for lithium-ion phosphate batteries for large plants (100 MW).[21] We find that the 2 hour duration battery is most profitable: While the 10h duration battery earns roughly 2-3 times the short run profits of a 2h duration one with a similar generation capacity, the investment costs are roughly 4 times higher. We refer to appendix G for results for 4h and 10h duration batteries.

Figure 5.1 illustrates the break even investment costs of storage for the given parameter values. In the baseline year investment costs far exceed the profits of bulk electricity storage in spite of WWS technologies providing more than 50% of the Danish supply in 2019. This result is mainly due to the combination of two factors: (i) The availability of large-scale storage especially in Sweden and Norway and (ii) the strategy of developing a transmission grid that allows for trade to balance markets. However, with the increased capacity share of WWS technologies in 2030-2040 in Denmark as well as neighboring countries, electricity storage starts to be profitable at investment costs around 250-400\$/kWh. In comparison, Linn and Shih (2019) find that storage costs has to drop below 160\$/kWh before being profitable in a simulated version of the ERCOT system (Texas) in 2030. We note, however, that with the projected investment costs, electricity storage is still only profitable after 2030 in our simulations. Once again, the investments in transmission grids from 2019-2030 plays an important role for this result. While Danish dispatchable generation capacity

[20] The lifetime and marginal costs correspond to the data from Danish Energy Agency (2018b).

[21] Baxter et al. (2020) only includes data for 2020 and 2030. Thus, for 2040, we simply project that costs drop by roughly 1/3 following the trend in Danish Energy Agency (2018b).

Figure 5.1: Break even investment costs of storage



Note: The discounted short run profits of the storage technology assumes an discount rate of 8% and a technical lifetime of 20 years. The vertical lines indicate cost estimates from Baxter et al. (2020).

is retired in this period, transmission lines to the Netherlands and Great Britain are commissioned with a combined import capacity of 2.1GW.

Finally, we note that discounted profits in figure 5.1 are downward sloping in generation capacity, indicating a diminishing return to capacity. This shows that as more electricity storage is installed, price volatility decreases which leads to lower potential for further arbitrage.

5.2 The effect of storage on key equilibrium outcomes

We evaluate the effect of storage on three key equilibrium outcomes: Prices, emissions, and profitability of intermittent renewables.

The effect on prices

At its core, electricity storage provides more stable equilibrium prices. Figure 5.2 confirms that, in particular for 2019-2030, large-scale electricity storage significantly reduces price volatility. However, the effect of storage on average electricity prices is quite different over time. Electricity storage affects mean prices through two channels: First, due to the loss of energy from storage, batteries are in fact net consumers of electricity. In our case, the roundtrip efficiency of the battery is 90% in 2019, implying that it consumes roughly 11% more than it supplies to the electricity grid. In other words, batteries consume more electricity when prices are low than it discharges when prices are high. All things equal, this drives up average prices.[22] Second, equilibrium prices may respond differently in low and high price states. For 2019 we find, in line with Sioshansi (2011), that the net effect of storage is an increase in prices. However, over time peak prices increase significantly to a price region where supply is highly price-inelastic (cf. the flat part of figure 5.3). Thus, beyond 2019, the asymmetric response of prices in high and low states drive down average prices (qualitatively in line with Linn and Shih 2019).

[22]The roundtrip efficiency is estimated to increase to roughly 95% by 2030 (Danish Energy Agency 2018b).

Figure 5.2: Effect of storage on equilibrium outcomes

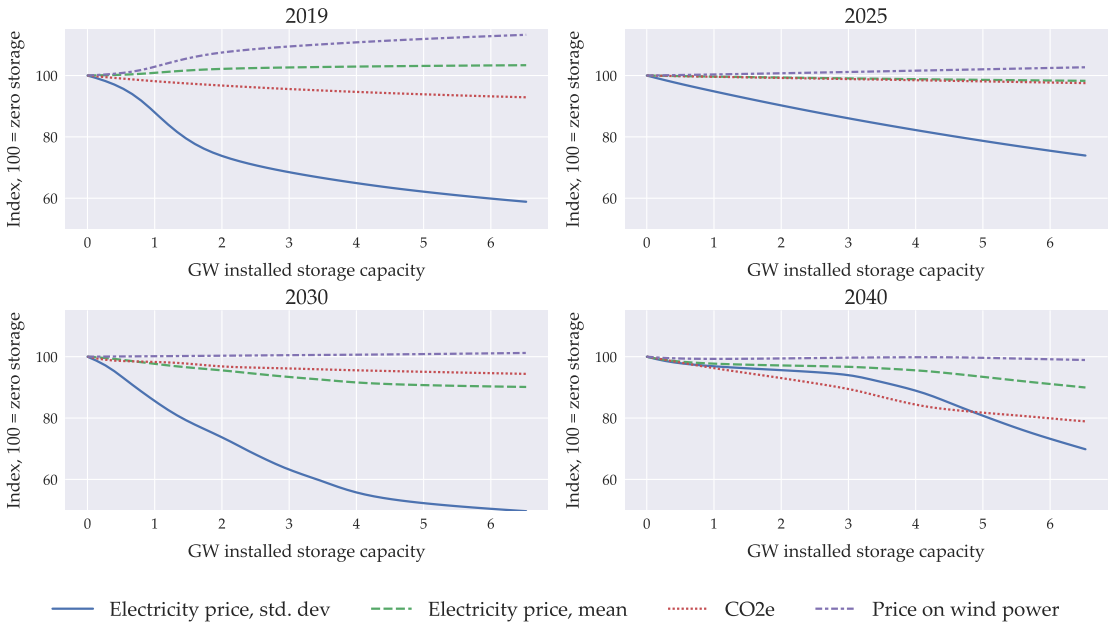
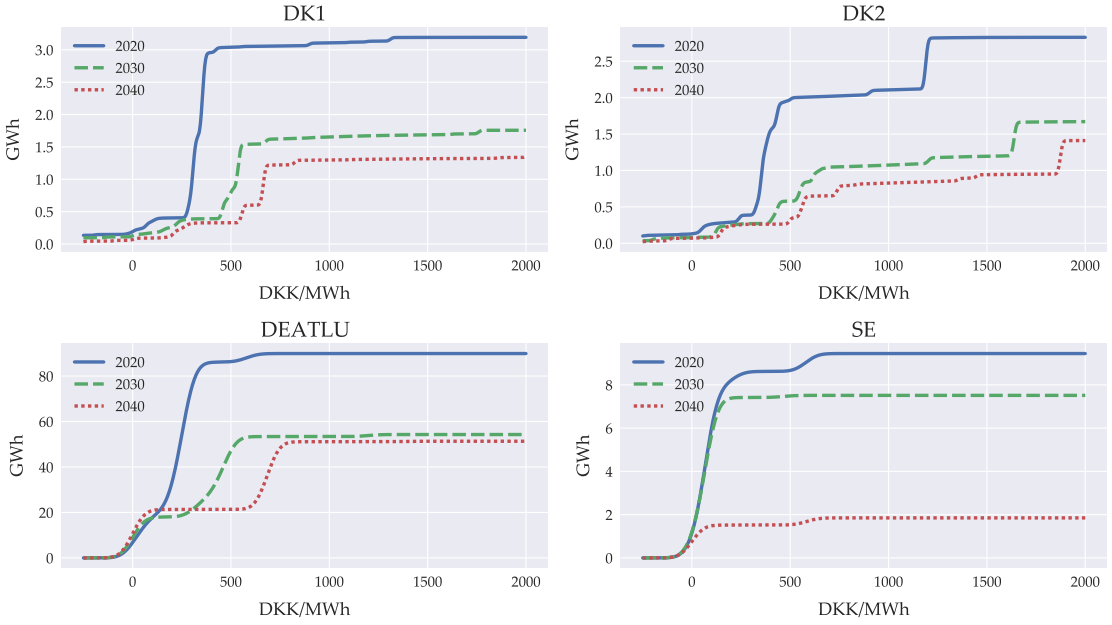


Figure 5.3: Hourly electricity supply from dispatchable plants



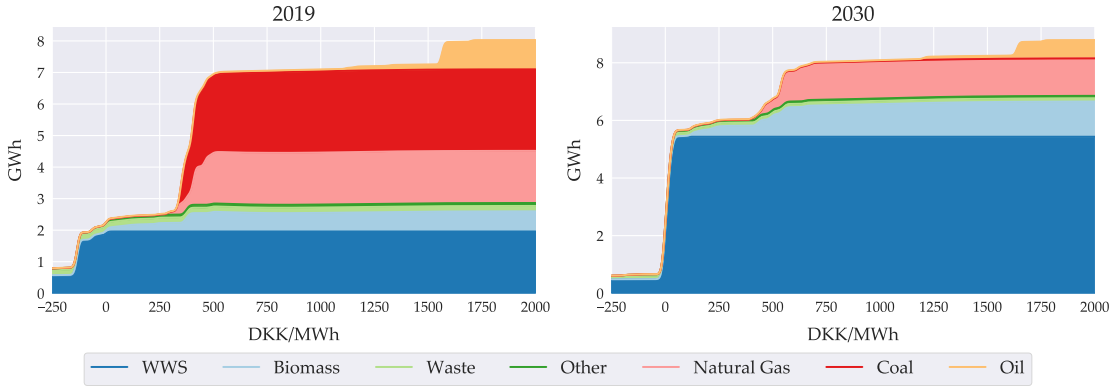
Note: The plot assumes a constant price of heat in Denmark of around 65 DKK/GJ.

The effect on emissions

Figure 5.2 further shows that throughout all years, the use of electricity storage lowers greenhouse gas emissions from the Danish electricity and district heat systems; a similar result holds for SO_2 and NO_x emissions (see appendix G). This is the opposite effect than what related literature finds for the case of Texas. The reason for this discrepancy can be found in the composition of plants in Denmark relative to Texas: Figure 5.4 displays the composition of primary energy goods used

to produce electricity for varying prices. In both 2019 and 2030, the marginal supply in low price states (−50-100 DKK/MWh) comes from clean energy such as WWS and biomass.[23] In 2019, model peak prices are only at roughly 400-500 DKK/MWh, in which case the marginal supply is a combination of coal, natural gas, and biomass. In 2030, model peak prices are much higher (1500-2000 DKK/MWh), in which case the marginal supply is largely based on oil. In both instances, the peak supply is more emission-intensive than in off-peak hours. Comparably, in the case of Texas, peak marginal supply (natural gas) is less emission-intensive than the off-peak marginal supply (coal). Thus, when electricity storage crowds out production in peak price states, it crowds out emission-intensive plants in Denmark and relatively clean plants in Texas. When electricity storage charges in low price states, the additional electricity generation is clean in Denmark and emission-intensive in Texas.

Figure 5.4: Hourly average supply of electricity by fuel types



Note: The plot assumes a constant price of heat in Denmark of around 65 DKK/GJ. the category "other" covers bio gas, bio oil, hydrogen, and uranium. In the Danish case, this is mostly bio gas and secondarily bio oil. The supply reflects the "hourly average" as it assumes an average capacity factor for relevant plants (e.g. WWS).

We note that our result of lower greenhouse gas emissions holds in spite of having fixed generation capacities. As intermittent renewables are generally considered complementary to electricity storage, endogenizing investments in generation capacities is likely to lead to larger reductions in emissions. Furthermore, we generally estimate emission leakage rates to be negative in our simulations, i.e. neighboring countries reduce emissions as well (see appendix G). An important caveat to this result is that the model does not include an endogenous price of CO_2 quotas in the EU Emissions Trading System.

The effect on the value of wind energy

In the baseline projection the downlift on Danish wind energy increases steadily from roughly 40 DKK/MWh in 2019 to around 160 DKK/MWh in 2040 as the generation capacity of intermittent renewables increase (cf. figure 3.3). This trend is a combination of wind production being negatively correlated with peak demand and having very low short run variable costs. The result is that

[23] We categorize biomass as a "clean"/CO_2-free energy source in our model, as this is the current regulatory status.

wind production is not at its highest when demand is high, and produces at very low prices. Due to these features electricity storage and wind power is considered to be complementary technologies (Bistline et al. 2020; Sioshansi 2011).

In our simulations, we generally confirm that electricity storage increases the *relative* profitability of wind power compared to other technologies. The so-called downlift on wind decreases by 20-60% with large scale electricity storage.[24] However, while increased electricity storage generally tends to increase the *absolute* value as well, this is not always the case. Figure 5.2 illustrates the price of wind energy as a function of installed storage capacity. In the years 2019, 2025, 2030 the price is increasing in electricity storage capacity. However, in 2040, the price of wind energy decreases with installed storage capacity. To explain this result, figure 5.5 plots the change in Danish equilibrium prices for all 25 intra-year states from installing all 6.5 GW generation capacity against wind production relative to demand for the same years (2019, 2025, 2030, 2040). In 2019, the lowest prices are primarily driven by domestic excess production of wind power. Thus, when batteries are charged in these low-price states, it drives up prices significantly (up to 210 DKK/MWh). Furthermore, in states with peak prices the discharged energy from the batteries causes small price drops. The reason for this asymmetric price response in 2019 is due to a relatively large back-up capacity of dispatchable plants. In 2040, however, this pattern reverses: First, in 2040, supply is inelastic in states with peak prices because large shares of dispatchable back-up capacity has been retired. This leads to large price drops in peak price states when electricity is discharged. Second, investments in transmission lines and neighbouring countries' investments in intermittent renewables imply that prices in the other tail of the distribution (low-price states) respond less to the introduction of electricity storage in Denmark.

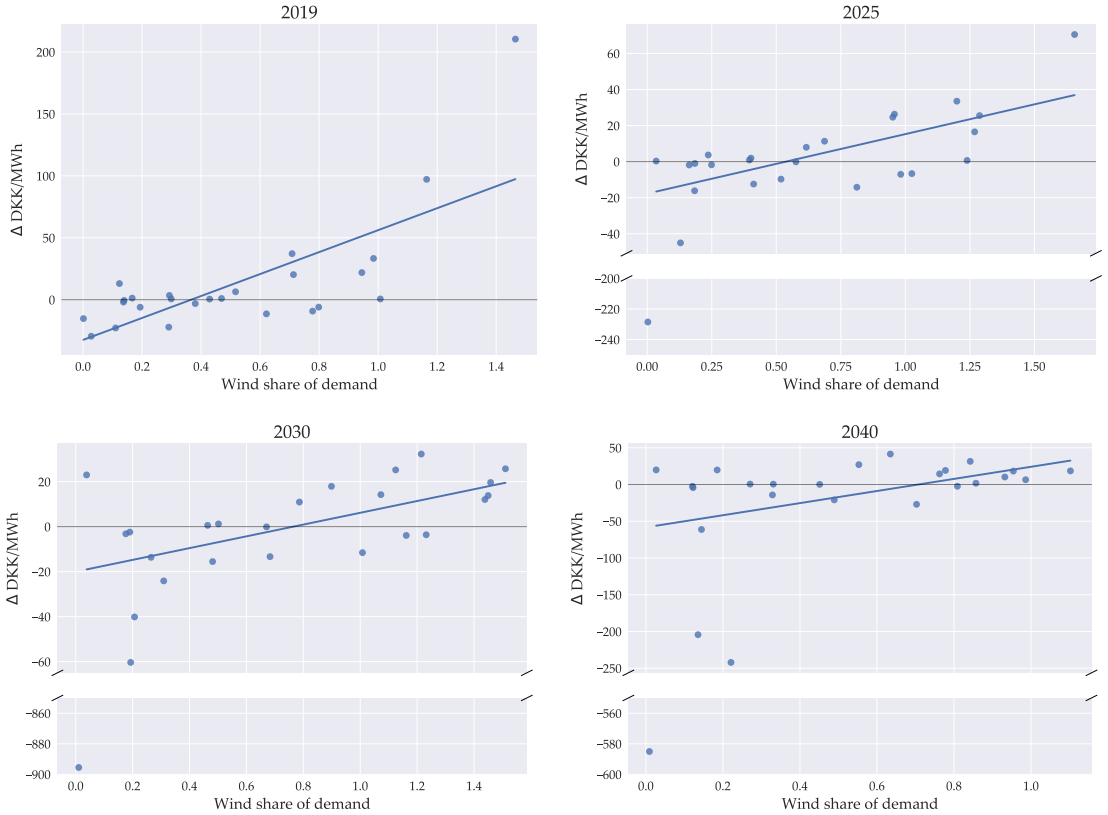
6 Conclusion

As economies increasingly rely on intermittent energy supply it becomes more difficult to stabilize supply and demand of electricity. In economic terms, the value of electricity storage derives from reallocating supply from states where electricity is abundant and valued at a low price to states where it is scarce and valued at a high price. In this sense, electricity storage can be an essential instrument to enable increased use of intermittent renewable energy. The literature on electricity storage, however, has not provided any conclusive answers in terms of the value and potential of the technology. Previous results range from concluding that electricity storage is cost-effective at sufficiently high shares of intermittent renewables in electricity generation, to concluding that it is not cost-effective and might even lead to increases in both emissions and average prices.

To assess the effects and value of electricity storage, we used a bottom-up model that accounts for competing strategies in mitigating intermittency such as electricity storage, trade in electricity, demand flexibility and back-up capacity of various types of dispatchable plants. We formulated a computationally tractable model of energy storage that can be specified as a system of nonlinear

[24] See figure G.3 in appendix G.

Figure 5.5: The effect of storage on prices and Danish wind production



Note: The change in price is the difference between full scale storage (6.5 GW generation capacity) and no storage. The lines indicate simple least square estimations of the correlation between the change in prices and the share of wind production.

equations and efficiently solved using gradient-based methods. In a structural estimation exercise we found that our formulation is able to replicate the generation and storage decisions of Norwegian hydroelectric plants. To further reduce the computational burden of solving the model, we aggregated all 8760 hours within a year to a smaller number of $K = 25$ representative states in a non-chronological manner. Instead of the traditional method that reestablishes chronology, we provided an approximation to the storage problem that keeps the size of the model relatively small. We showed that, in our applications, the approximation returns similar results as a model with full hourly resolution.

In the case of Denmark, we found that large scale deployment of electricity storage in the form of lithium-ion batteries generally leads to significant reductions in greenhouse gas emissions and average electricity prices. Emission reductions occur because electricity supply is relatively emission-intensive in peak hours; this supply is replaced by clean stored electricity. Average electricity prices fall in our model because the supply in peak price states is relatively price inelastic. Thus, when batteries provide marginally more electricity in peak states, the price decreases a lot compared to how much prices increase in low price states when batteries are charging. Even though batteries are net consumers of electricity (and therefore lowers net supply), the asymmetric price-response leads to lower average prices. These results stand in contrast to previous findings

for the Texas electricity system and highlight that the effects of electricity storage is specific to the market in question.

We also showed that, although the value of electricity storage is roughly 50% higher in Denmark than what has previously been estimated for Texas, it is still only likely to be cost-effective after 2030. This result emphasizes the importance of modeling the potential of other strategic substitutes for mitigating intermittency. In our framework electricity storage does not become profitable before 2030 because there are large investments in transmission lines to other European countries. Denmark is able to trade its way out of intermittency with neighbouring countries that either have large dispatchable back-up capacity or natural access to large, cheap, and flexible hydroelectric power. Only *post* 2030, when neighbouring countries have significantly retired back-up capacity and increased the capacity of intermittent energy, does electricity storage become profitable.

Finally, our empirical findings are subject to a number of important caveats: First, the results are based on a frozen policy assumption for Denmark that does not satisfy the political target for emission reductions for 2030. If this target is achieved through increased use of intermittent renewables in electricity generation and electrification of other key sectors, this is likely to increase the value of electricity storage. Second, electricity storage leads to significant reductions in price volatility and emissions; the former is currently not valued at all and the latter may be undervalued. Thus, taking these considerations into account, electricity storage may still be socially optimal.

References

Balyk, Olexandr et al. (2019). "TIMES-DK: Technology-rich multi-sectoral optimisation model of the Danish energy system". In: *Energy Strategy Reviews* 23, pp. 13–22. ISSN: 2211-467X.

Barton, John and David Infield (July 2004). "Energy Storage and Its Use With Intermittent Renewable Energy". In: *Energy Conversion, IEEE Transactions on* 19, pp. 441–448.

Baxter, Richard et al. (Dec. 2020). *2020 Grid Energy Storage Technology Cost and Performance Assessment*. Technical Report DOE/PA-0204. U.S. Department of Energy.

Berg, Rasmus K. and Janek Eskildsen (2021). "Electricity and District Heat systems in a Computable General Equilibrium Framework".

Bistline, John et al. (July 2020). "Energy storage in long-term system models: a review of considerations, best practices, and research needs". In: *Progress in Energy* 2.3, p. 032001.

Carson, Richard T. and Kevin Novan (2013). "The private and social economics of bulk electricity storage". In: *Journal of Environmental Economics and Management* 66.3, pp. 404–423. ISSN: 0095-0696.

Cebulla, F., T. Naegler, and M. Pohl (2017). "Electrical energy storage in highly renewable European energy systems: Capacity requirements, spatial distribution, and storage dispatch". In: *Journal of Energy Storage* 14, pp. 211–223. ISSN: 2352-152X.

Danish Council on Climate Change (2018). *Status for Danmarks klimamålsætninger og forpligtelser 2018*. Tech. rep. Copenhagen, Denmark: Danish Council on Climate Change.

Danish Energy Agency (2018a). *RamsesR docomentation*. https://ens.dk/sites/ens.dk/files/Analyser/ramsesr.pdf. Accessed: 03-09-2018.

Danish Energy Agency (2018b). *Technology Data for Energy Storage*. Tech. rep. Danish Energy Agency, p. 234.

Danish Energy Agency (2021). *Klimastatus -fremskrivning, 2021*. Tech. rep. Danish Energy Agency, p. 79.

de Sisternes, Fernando J., Jesse D. Jenkins, and Audun Botterud (2016). "The value of energy storage in decarbonizing the electricity sector". In: *Applied Energy* 175, pp. 368–379. ISSN: 0306-2619.

Denholm, P, J Jorgenson, et al. (2013). *The Value of Energy Storage for Grid Applications*. Tech. rep. National Renewable Energy Laboratory.

Denholm, Paul and Maureen Hand (2011). "Grid flexibility and storage required to achieve very high penetration of variable renewable electricity". In: *Energy Policy* 39.3, pp. 1817–1830. ISSN: 0301-4215.

Diaz, Gabriel, Andrés Inzunza, and Rodrigo Moreno (2019). "The importance of time resolution, operational flexibility and risk aversion in quantifying the value of energy storage in long-term energy planning studies". In: *Renewable and Sustainable Energy Reviews* 112, pp. 797–812. ISSN: 1364-0321.

Energy Data Service (2021a). *Elspot Prices*. https://www.energidataservice.dk/tso-electricity/elspotprices. [Online; accessed 19-10-2021].

Energy Data Service (2021b). *Production and Consumption - Settlement*. https://www.energidataservice.dk/tso-electricity/productionconsumptionsettlement. [Online; accessed 19-10-2021].

ENTSO-E (2017a). *Actual Generation per Production Type* (Version 30 October 2020) [Data set]. https://transparency.entsoe.eu/generation/r2/actualGenerationPerProductionType/show.

ENTSO-E (2017b). *Day-ahead Prices* (Version 30 October 2020) [Data set]. https://transparency.entsoe.eu/transmission-domain/r2/dayAheadPrices/show.

ENTSO-E (2017c). *Water Reservoirs and Hydro Storage Plants* (Version 30 October 2020) [Data set]. https://transparency.entsoe.eu/generation/r2/waterReservoirsAndHydroStoragePlants/show.

ENTSO-E (2017d). *Water Reservoirs and Hydro Storage Plants* (Version 30 October 2020) [Data set]. https://transparency.entsoe.eu/generation/r2/actualGenerationPerProductionType/show.

ENTSO-E (2021). *Completing The Map: Power System Needs in 2030 and 2040*. Tech. rep. Accessed: 17-10-2021. ENTSO-E, p. 70.

Eurostat (2021). *Electricity production capacities by main fuel groups and operator*. accessed 2 July 2021.

Gabrielli, Paolo et al. (2018). "Optimal design of multi-energy systems with seasonal storage". In: *Applied Energy* 219, pp. 408–424. ISSN: 0306-2619.

Helm, Carsten and Mathias Mier (2021). "Steering the energy transition in a world of intermittent electricity supply: Optimal subsidies and taxes for renewables and storage". In: *Journal of Environmental Economics and Management* 109, p. 102497. ISSN: 0095-0696.

Hirth, Lion, Falko Ueckerdt, and Ottmar Edenhofer (2015). "Integration costs revisited – An economic framework for wind and solar variability". In: *Renewable Energy* 74, pp. 925–939. ISSN: 0960-1481.

Hothorn, Torsten, Kurt Hornik, and Achim Zeileis (2006). "Unbiased recursive partitioning: A conditional inference framework". In: *Journal of Computational and Graphical Statistics* 15.3, pp. 651–674.

IEA (2021). *Net Zero by 2050*. Tech. rep. Lead authors and co-ordinators, Team, Bouckaert, Stéphanie and Pales, Araceli F. and Mcgalde, Christophe and Remme, Uwe and Wanner, Brent. IEA.

Jacobson, Mark Z. et al. (2017). "100% Clean and Renewable Wind, Water, and Sunlight All-Sector Energy Roadmaps for 139 Countries of the World". In: *Joule* 1.1, pp. 108–121. ISSN: 2542-4351.

Kotzur, Leander et al. (2018). "Time series aggregation for energy system design: Modeling seasonal storage". In: *Applied Energy* 213, pp. 123–135. ISSN: 0306-2619.

Krajačić, Goran, Neven Duić, and Maria da Graça Carvalho (2011). "How to achieve a 100% RES electricity supply for Portugal?" In: *Applied Energy* 88.2. The 5th Dubrovnik Conference on Sustainable Development of Energy, Water and Environment Systems, held in Dubrovnik September/October 2009, pp. 508–517. ISSN: 0306-2619.

Linn, Joshua and Jhih-Shyang Shih (2019). "Do lower electricity storage costs reduce greenhouse gas emissions?" In: *Journal of Environmental Economics and Management* 96, pp. 130–158. ISSN: 0095-0696.

Lv, Yibing et al. (2007). "A penalty function method based on Kuhn–Tucker condition for solving linear bilevel programming". In: *Applied Mathematics and Computation* 188.1, pp. 808–813. ISSN: 0096-3003.

Nahmmacher, Paul et al. (2016). "Carpe diem: A novel approach to select representative days for long-term power system modeling". In: *Energy* 112, pp. 430–442. ISSN: 0360-5442.

Poncelet, Kris et al. (2017). "Selecting Representative Days for Capturing the Implications of Integrating Intermittent Renewables in Generation Expansion Planning Problems". In: *IEEE Transactions on Power Systems* 32.3, pp. 1936–1948.

Reguant, Mar (2019). "The Efficiency and Sectoral Distributional Impacts of Large-Scale Renewable Energy Policies". In: *Journal of the Association of Environmental and Resource Economists* 6.S1, pp. 129–168.

Renaldi, Renaldi and Daniel Friedrich (2017). "Multiple time grids in operational optimisation of energy systems with short- and long-term thermal energy storage". In: *Energy* 133, pp. 784–795. ISSN: 0360-5442.

Sioshansi, Ramteen (2011). "Increasing the Value of Wind with Energy Storage". In: *The Energy Journal* 32.2, pp. 1–29. ISSN: 01956574, 19449089.

Su, Che-Lin and Kenneth L. Judd (2012). "Constrained Optimization Approaches to Estimation of Structural Models". In: *Econometrica* 80.5, pp. 2213–2230. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA7925.

Wogrin, Sonja, David Galbally, and Javier Reneses (2016). "Optimizing Storage Operations in Medium- and Long-Term Power System Models". In: *IEEE Transactions on Power Systems* 31.4, pp. 3129–3138.

Zerrahn, Alexander, Wolf-Peter Schill, and Claudia Kemfert (2018). "On the economics of electrical storage for variable renewable energy sources". In: *European Economic Review* 108, pp. 259–279. ISSN: 0014-2921.

Appendices

A Definitions

Definition 1 (Flexible Sigmoid Function)**.**
Define a flexible sigmoid function as a mapping $f : \mathbb{R} \to [0, 1]$ *with characteristic parameters* $\alpha \in \mathbb{R}$, $\Sigma \in \mathbb{R}^n_+$, *such that:*

i. f *is monotonically strictly increasing and continuously differentiable for* $\min(\Sigma) > 0$.

ii. f *approaches the discrete-choice function in the limit*

$$\lim_{\Sigma \to \mathbf{0}} f(x;\ \alpha, \Sigma) = \begin{cases} 0, & x < \alpha \\ 1, & x \geq \alpha \end{cases}.$$

iii. f*'s supremum (infimum) is* 1 (0)*:*

$$\lim_{x \to \infty} f(x;\ \alpha, \Sigma) = 1$$
$$\lim_{x \to -\infty} f(x;\ \alpha, \Sigma) = 0.$$

B A convex optimization problem

We assume that marginal costs depends on both Y_h and S_h in a separable way such that

$$\frac{\partial C}{\partial Y_h} = g_y\left(\frac{Y_h - \underline{Y}}{\overline{Y} - \underline{Y}};\ \bar{c}, \Sigma_y\right)$$
$$\frac{\partial C}{\partial S_h} = g_s\left(\frac{S_h}{\overline{S}};\ \Sigma_s\right)$$

Generally, we assume that g_y, g_s are increasing in Y_h, S_h respectively. This is consistent with the assumption of inverse flexible sigmoid functions that automatically ensures that $Y_h \in [\underline{Y}, \overline{Y}]$ and $S_h \in [0, \overline{S}]$. Thus, we can ignore the capacity and storage constraints in the following. In the following, we start by assuming that $\gamma_c = \gamma_d = 1$; Appendix B.1 shows that under appropriate assumptions, the results can be generalized to the case of other charge/discharge efficiencies.

With separability of the marginal cost function the problem of the storage plant is to maximize (for convenience repeated here):

$$V = \sum_{h}^{H} \beta^h \left[p_h Y_h - C(Y_h, S_h)\right],$$
$$\text{s.t. } S_h = \eta S_{h-1} + W_h - Y_h,$$
$$S_0 = S_H \geq 0, \text{ given.}$$

Using the law of motion we can write this as a problem of maximizing wrt. $S_1, ..., S_{H-1}$. This

yields the first order derivatives:

$$\begin{aligned}\frac{\partial V}{\partial S_h} =&\beta^{h-1}\left\{p_h\frac{\partial Y_h}{\partial S_h}-\left(\frac{\partial C}{\partial S_h}+\frac{\partial C}{\partial Y_h}\frac{\partial Y_h}{\partial S_h}\right)+\beta\left(p_{h+1}\frac{\partial Y_{h+1}}{\partial S_h}-\frac{\partial C}{\partial Y_{h+1}}\frac{\partial Y_{h+1}}{\partial S_h}\right)\right\}\\ =&\beta^{h-1}\left\{-\frac{\partial C}{\partial S_h}+\left(\frac{\partial C}{\partial Y_h}-p_h\right)+\beta\eta\left(p_{h+1}-\frac{\partial C}{\partial Y_{h+1}}\right)\right\},\end{aligned}$$

and second order derivative:

$$\frac{\partial^2 V}{\partial S_h^2}=-\beta^{h-1}\left\{\frac{\partial^2 C}{\partial S_h^2}+\frac{\partial^2 C}{\partial Y_h^2}+\beta\left(\eta\right)^2\frac{\partial^2 C}{\partial Y_{h+1}^2}\right\}<0,$$

where the sign comes from the fact that all second order derivatives are positive when the marginal costs are inverse sigmoid functions. The cross-derivatives are defined as

$$\frac{\partial^2 V}{\partial S_h S_{h-1}}=\beta^{h-1}\eta\frac{\partial^2 C}{\partial Y_h^2}>0.$$

It is straightforward to verify that the Hessian related to V is negative definite, i.e. the function V is concave. Thus, the first order conditions imply the global maximum for the problem. Next, rephrase the problem in terms of the Lagrangian instead:

$$\mathcal{L}=\sum_h^H\beta^h\left\{p_hY_h-C(Y_h,S_h)+\lambda_h\left(\eta S_{h-1}+W_h-Y_h-S_h\right)\right\}$$

The first order conditions are defined as

$$\begin{aligned}\lambda_h&=\left(p_h-\frac{\partial C}{\partial Y_h}\right)\\ \lambda_h&=\beta\eta\lambda_{h+1}-\frac{\partial C}{\partial S_h}\\ S_h&=\eta S_{h-1}+W_h-Y_h\end{aligned}$$

Inverting the partial derivative we get the closed form expression for the optimal Y_h:

$$Y_h=\underline{Y}+(\overline{Y}-\underline{Y})f_y\left(p_h-\overline{c}-\lambda_h+\kappa_y;\ \Sigma_y\right) \tag{8a}$$

$$\lambda_h=\beta\eta\lambda_{h+1}-g_s\left(\frac{S_h}{\overline{S}};\ \Sigma_s\right) \tag{8b}$$

$$S_h=\eta S_{h-1}+W_h-Y_h \tag{8c}$$

where f_y is a flexible sigmoid (inverse of g_y). λ_h measures the continuation value, and is a function of λ_{h+1} and the relative distance between S_h and the bounds on storage $(0,\overline{S})$.

We note that this system of equations is particularly easy to solve: Given prices and the initial level of stored energy S_0, the system can be ordered topologically and solved as follows: Guess at an value of the continuation value in the first hour, λ_1. Given this level, the dependency graph associated with the system is a Directed Acyclic Graph (DAG). In particular, the topological order

is given by $\{Y_1, S_1, \lambda_2, Y_2, S_2, ...\}$ and so on for all h. In other words, the numerical problem of solving the system of equations comes down to finding the value of λ_1 such that recursively solving forward on equations (8) such that $S_H = S_0$.

B.1 The case with a smoothed piece-wise linear function in the law of motion

Consider the storage optimization problem with different charge/discharge efficiencies. We replace the law of motion with

$$S_h = \eta S_{h-1} + \zeta\left(W_h - Y_h\right),$$

where we define ζ as the smooth piecewise linear function

$$\begin{aligned}
\zeta(x) &= \frac{\gamma_c + 1/\gamma_d}{2} x + \frac{\gamma_c - 1/\gamma_d}{2}\left(\sqrt{x^2 + \sigma_\zeta^2}\right) \\
\zeta'(x) &= \frac{\gamma_c + 1/\gamma_d}{2} + \frac{\gamma_c - 1/\gamma_d}{2}\frac{x}{\sqrt{x^2 + \sigma_\zeta^2}} > 0 \\
\zeta''(x) &= \frac{\gamma_c - 1/\gamma_d}{2}\frac{\sigma_\zeta^2}{\left(x^2 + \sigma_\zeta^2\right)^{3/2}}, \leq 0
\end{aligned}$$

where γ_d, γ_c are discharge, charge efficiencies, and $\sigma_\zeta \geq 0$ is a smoothing parameter, such that

$$\lim_{\sigma_\zeta \to 0} \zeta(x) = \begin{cases} (W_h - Y_h)/\gamma_d, & W_h - Y_h < 0 \\ \gamma_c (W_h - Y_h), & W_h - Y_h \geq 0 \end{cases}$$

Note that, as $\gamma_d, \gamma_c \leq 1$, the derivative is bounded by $[\gamma_c, (\gamma_c + 1/\gamma_d)/2]$. Similarly, the second derivative is bounded by $[\sigma_\zeta(\gamma_c - 1/\gamma_d)/2, 0)$.

In this case, the derivative of the value function is defined as

$$\frac{\partial V}{\partial S_h} = \beta^{h-1}\left\{\frac{1}{\zeta_h'}\left(\frac{\partial C}{\partial Y_h} - p_h\right) - \frac{\partial C}{\partial S_h} + \frac{\eta\beta}{\zeta_{h+1}'}\left(p_{h+1} - \frac{\partial C}{\partial Y_{h+1}}\right)\right\},$$

where ζ_h' is short hand for $\zeta'(W_h - Y_h)$. The second order derivative is defined as:

$$\frac{\partial^2 V}{\partial S_h^2} = \beta^{h-1}\left\{\frac{1}{(\zeta_h')^2}\left(\frac{\zeta_h''}{\zeta_h'}\left[p_h - \frac{\partial C}{\partial Y_h}\right] - \frac{\partial^2 C}{\partial Y_h^2}\right) - \frac{\partial^2 C}{\partial S_h^2} + \frac{\eta^2\beta}{(\zeta_{h+1}')^2}\left(\frac{\zeta_{h+1}''}{\zeta_{h+1}'}\left[p_{h+1} - \frac{\partial C}{\partial Y_{h+1}}\right] - \frac{\partial^2 C}{\partial Y_{h+1}^2}\right)\right\}$$

where we've used that

$$\begin{aligned}
\frac{\partial(\zeta_h')^{-1}}{\partial S_h} &= -\frac{\zeta_h''}{(\zeta_h')^3} \\
\frac{\partial(\zeta_{h+1}')^{-1}}{\partial S_h} &= \eta\frac{\zeta_{h+1}''}{(\zeta_{h+1}')^3}
\end{aligned}$$

The cross-derivatives are defined as:

$$\frac{\partial^2 V}{\partial S_h S_{h-1}} = \beta^{h-1} \frac{\eta}{(\zeta_h')^2} \left(\frac{\zeta_h''}{\zeta_h'} \left[\frac{\partial C}{\partial Y_h} - p_h \right] + \frac{\partial^2 C}{\partial Y_h^2} \right)$$

For the problem to be globally concave, we require specifically that:

Assumption 2 (Regularity assumption)**.**
The value $\gamma_c - 1/\gamma_d$ and/or σ_ζ are/is relatively small such that for all $Y \in [\underline{Y}, \overline{Y}]$:

$$\frac{\partial^2 C}{\partial Y^2} \geq -\frac{\zeta''(W-Y)}{\partial \zeta'(W-Y)} \frac{\partial C}{\partial Y} \tag{9}$$

Note that we can always find parameters such that this is the case, at least when approximating ζ using the smooth piecewise linear function as outlined above. Note for instance, that there can only potentially be an issue in discharging states as the left hand side is always positive, whereas the right hand side is only positive when $\partial C/\partial Y > 0$, which is the case for $Y > 0$ (for batteries/heat storage). For positive values ($\partial C/\partial Y > 0$), we know that, at least for sufficiently small values of σ, $(\partial^2 C/\partial Y^2)/(\partial C/\partial Y) > 1$. Thus, as the ratio ζ''/ζ' is similarly bounded, and proportional to σ_ζ, there always exists a smoothing parameter, such that the regularity assumption holds.

B.2 Penalty approach and approximate shadow values

The convex optimization framework presented in assumption 1 can similarly be interpreted as a penalty approach to replace the complementary and slackness conditions in the Karush-Kuhn-Tucker (KKT) conditions for optimality (see e.g. Lv et al. 2007).

Consider again the original linear programming problem with occasionally binding constraints in equation (1). The KKT conditions for optimality is then defined as

$$\lambda_h = \zeta' \left(p_h - \overline{c} + \underline{\eta}_h - \overline{\eta}_h \right) \tag{10a}$$

$$\lambda_h = \beta \eta \lambda_{h+1} + \underline{\mu}_h - \overline{\mu}_h \tag{10b}$$

$$S_h = \eta S_{h-1} + \zeta (W_h - Y_h) \tag{10c}$$

$$0 = \overline{\eta}_h \left(\overline{Y} - Y_h \right), \qquad \overline{\eta}_h, \overline{Y} - Y_h \geq 0 \tag{10d}$$

$$0 = \underline{\eta}_h \left(Y_h - \underline{Y} \right), \qquad \underline{\eta}_h, Y_h - \underline{Y} \geq 0 \tag{10e}$$

$$0 = \overline{\mu}_h \left(\overline{S} - S_h \right), \qquad \overline{\mu}_h, \overline{S} - S_h \geq 0 \tag{10f}$$

$$0 = \underline{\mu}_h S_h, \qquad \underline{\mu}_h, S_h \geq 0, \tag{10g}$$

where λ_h is the shadow value of storage, $\overline{\mu}_h, \underline{\mu}_h$ are shadow costs of storage capacities $(\overline{S}, 0)$, and $\overline{\eta}_h, \underline{\eta}_h$ are shadow costs of generation capacities $(\overline{Y}, \underline{Y})$. Compared to this setup, the penalty function approach defines an alternative objective, e.g.

$$\sum_{h=1}^{8760} \beta^h \left[p_h Y_h - \overline{c} Y_h - \frac{1}{t} \left(\left[\overline{\eta}_h \left(\overline{Y} - Y_h \right) \right]^2 + \left[\underline{\eta}_h \left(Y_h - \underline{Y} \right) \right]^2 + \left[\overline{\mu}_h \left(\overline{S} - S_h \right) \right]^2 + \left[\underline{\mu}_h S_h \right]^2 \right) \right], \qquad t > 0$$

where t is a scalar used to obtain convergence to the original KKT problem. The idea of the penalty function approach is to obtain a (non-feasible) solution to an initial problem with a large value of t, and then iteratively decrease the value of t towards zero.

This penalty function relies on the shadow values $\overline{\mu}_h, \underline{\mu}_h, \overline{\eta}_h, \underline{\eta}_h$. The approach we take in the main section (assumption 1), however, can be seen as a different type of penalty function. To see how, start by observing that in optimum the shadow costs in the complementarity constraints follow the pattern

$$\overline{\mu}_h^* \begin{cases} = 0, & S_h < \overline{S}, \\ > 0, & S_h = \overline{S} \end{cases} .$$

Similarly, when they enter the KKT conditions, the shadow values are on the form:

$$\underline{\mu}_h - \overline{\mu}_h \begin{cases} < 0, & S_h = 0 \\ = 0, & 0 < S_h < \overline{S} \\ > 0, & S_h = \overline{S} \end{cases} .$$

Let g_s denote a continuous and differentiable approximation to $\underline{\mu}_h - \overline{\mu}_h$; this could, for instance, be an inverse flexible sigmoid type function defined over $S_h/\overline{S}$. Let $G_s \equiv \int g_s dS_h$. Using G_s as the relevant penalty function, the first order conditions that follow from this problem is identical to the one from the convex optimization problem in (2).

C Estimation of storage model on Norwegian hydroelectric power plants

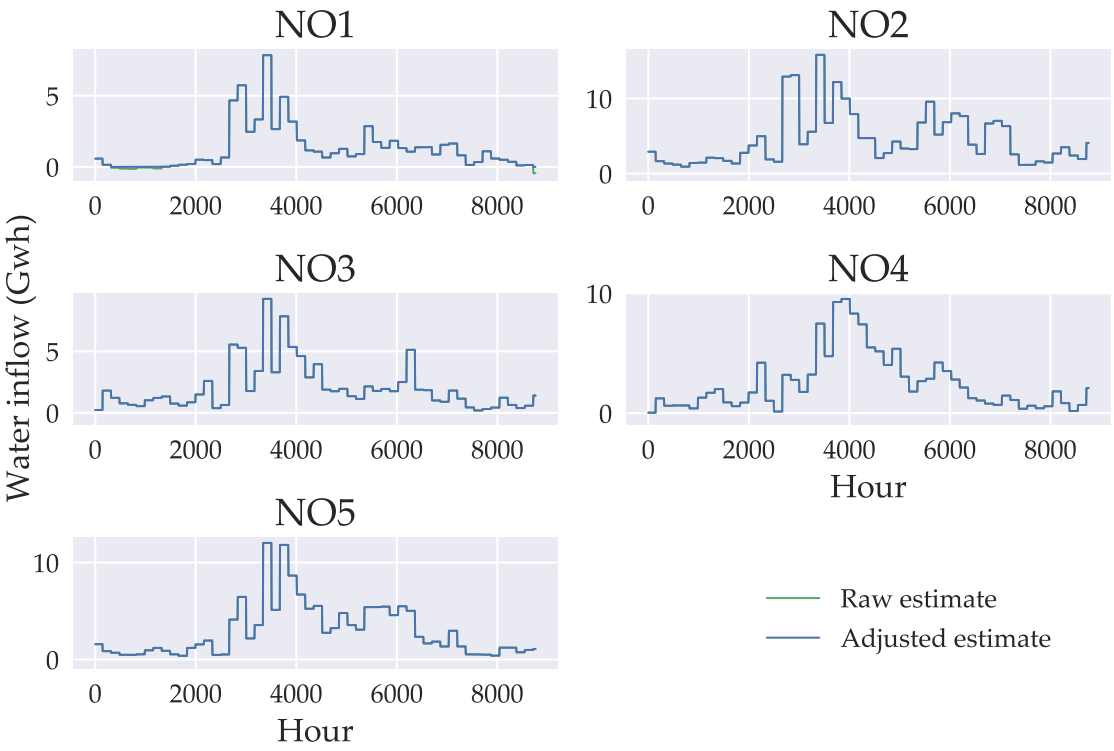
C.1 Data

The plants we are considering are Norwegian hydroelectric plants with reservoirs and adjacent rivers that provides an exogenous inflow of water into the reservoirs. Hourly hydroelectric production is retrieved from ENTSO-E (2017a) and spot prices from ENTSO-E (2017b) for 2019. There is no publicly available data for inflow levels, but reservoir levels are available at a weekly frequency at ENTSO-E (2017c). While production data is available at the plant level, reservoir levels are only available at the bidding zone level (i.e. NO1, NO2, NO3, NO4, NO5). With no obvious method for mapping plant reservoir levels at the bidding zone level down to the plant level, the most disaggregate level with consistent data is at the bidding zone level. Finally, there is no available data for water inflows. Instead, we aggregate hourly generation to the weekly level and back out the implied inflows using the simplified law of motion $S_w = S_{w-1} + \hat{W}_w - Y_w$, where $\hat{W}_w$ is our estimate of the accumulated water inflow in week w.[25] We note that this way estimating inflows implies that the imputed inflow is negative in week 3-8 and 52 for the NO1-area. We set these negative

[25] This simplified law of motion that corresponds to $\eta = \gamma = 1$ in our theoretical framework is relatively standard for the specific type of hydroelectric power plant.

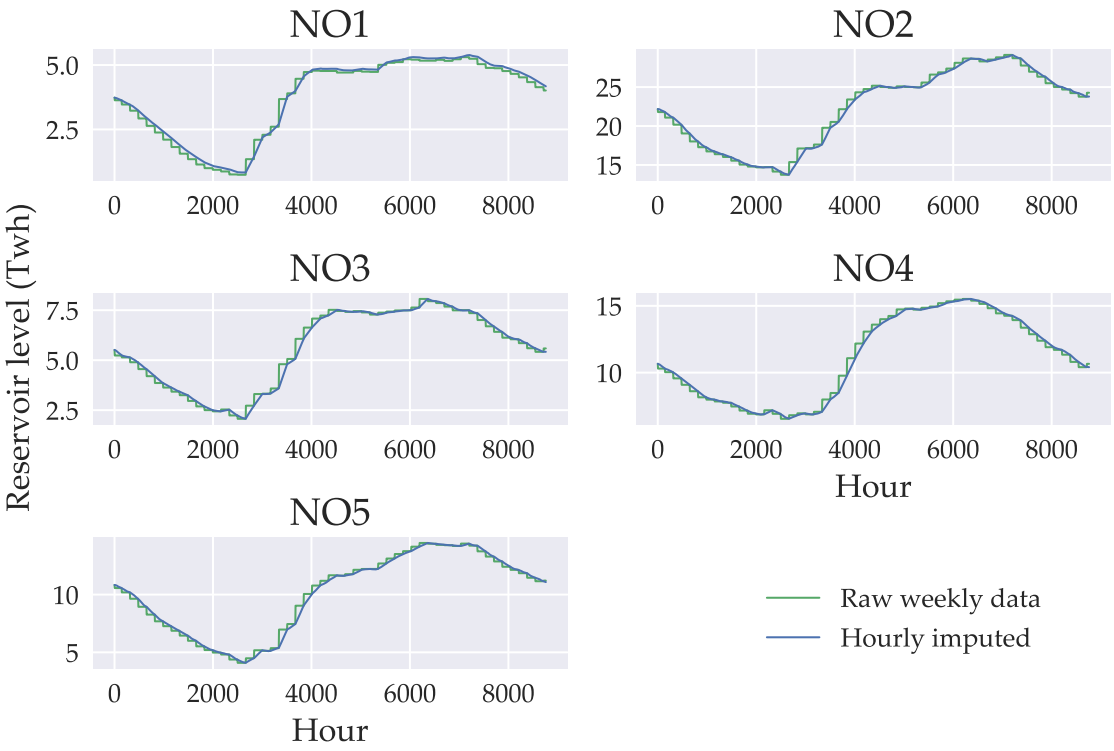
inflow values to zero. We then uniformly allocate the weekly accumulated inflow onto each hour of the week and calculate the implied hourly reservoir level using $S_h = S_0 + \sum_{s=1}^{h} \hat{W}_h - \sum_{s=1}^{h} Y_h$ for $h \in [1...8760]$. Figure C.1 illustrates the hourly inflow of water for five areas. For all areas except NO1 there is no adjustment to the estimated inflow, but for NO1, there is a small adjustment but it is virtually only visible in the last week of the year. Figure C.2 illustrate the weekly reservoir level collected from ENTSO-E (2017c) and the hourly imputed level. Finally, table 1 provides descriptive statistics for prices and production.

Figure C.1: Water inflows in 2019 at the bidding zone level



Note: The 'Raw estimate' is water inflow estimated using the law of motion $W_w = S_w - S_{w-1} + Y_w$. In the 'Adjusted estimate' we have set negative inflow values to zero.

Figure C.2: Reservoir levels in 2019 at the bidding zone level



Note: 'Raw weakly data' is the average reservoir level in a given week. When imputing the reservoir level at the hourly frequency we have assumed the weakly data is ultimo dated.

Table 1: Descriptive statistics for Norwegian hydroelectric plants, 2019

	Prices (DKK/MWh)	Power generation (GWh)
mean	290.687	2.616
std	60.278	1.485
min	10.303	0.406
median	288.336	2.296
max	817.151	9.772
N	5	
H	8760	
$N \times H$	43,800	

C.2 Maximum Likelihood

To structurally estimate the model we assume that observed electricity generation (Y_h) is contaminated with mean-zero i.i.d normal measurement error

$$\eta_{i,h} \equiv Y_{i,h} - Y^*_{i,h}\left(S_{h,i}, p_{i,h}; \theta\right) \sim \mathcal{N}\left(0, \sigma^2_\eta\right), \tag{11}$$

where i indexes the geographical area in norway (NO1-NO5), $Y^*_{i,h}(\cdot)$ is the policy function given by the system of equations in (4), and θ is a vector of estimation parameters. Using the hourly data from 5 regions, we have a balanced panel data set with $N = 5$ electricity areas and $H = 8760$ hours for 2019, leaving us with a total of 43,800 observations. Assumption (11) implies that the

likelihood of observing the data is

$$\mathcal{L}(\theta) = \prod_i \prod_h \phi(\eta_{i,h}(\theta)), \qquad \phi(\eta_{i,h}(\theta)) = \frac{1}{\sqrt{2\pi\sigma_{eta}^2}} \exp\left(-\frac{\eta_{i,h}(\theta)^2}{2\sigma_\eta^2}\right). \tag{12}$$

The estimation if formulated as a constrained maximization problem similar to the mathematical program with equilibrium constraints (MPEC) method proposed by Su and Judd (2012). Formally,

$$\begin{aligned} \max_{\theta \in \Theta} \ \log\left(\mathcal{L}(\theta)\right) \propto \sum_i \sum_h \log\left(\eta_{i,h}(\theta)\right) \\ \text{subject to system (4)}, \end{aligned} \tag{13}$$

where Θ is the parameter space discussed below.

C.3 Identification

Figure C.3 illustrates the shape of the marginal cost functions and how they depend on relevant parameters. The marginal generation costs are centered around the level $\bar{c}$, and diverges when it approaches the constraints $\underline{Y}, \overline{Y}$. A larger value of σ_y indicates that the costs increase more gradually when generation increases. The marginal storage costs is a smooth piecewise linear function that has two regimes: Outside the constraints $[0, \overline{S}]$ the marginal costs increase at a rate of κ and within the bounds the marginal costs increase at a rate of $2\gamma/\overline{S}$. The parameter δ determines how smooth the marginal costs move from one regime to the other. We note that whereas the marginal generation costs automatically ensures that $Y_h \in [\underline{Y}, \overline{Y}]$ this is not the case for the chosen parametrization of storage costs. Instead, the implication is that the reservoir constraints may be marginally violated if prices fluctuate by more than $\gamma + \kappa$ DKK/MWh during a year. It is indeed possible to apply the same cost structure as for generation costs, and thus ensure that $S_h \in [0, \overline{S}]$, however, the current parametrization is chosen to increase numerical stability.[26]

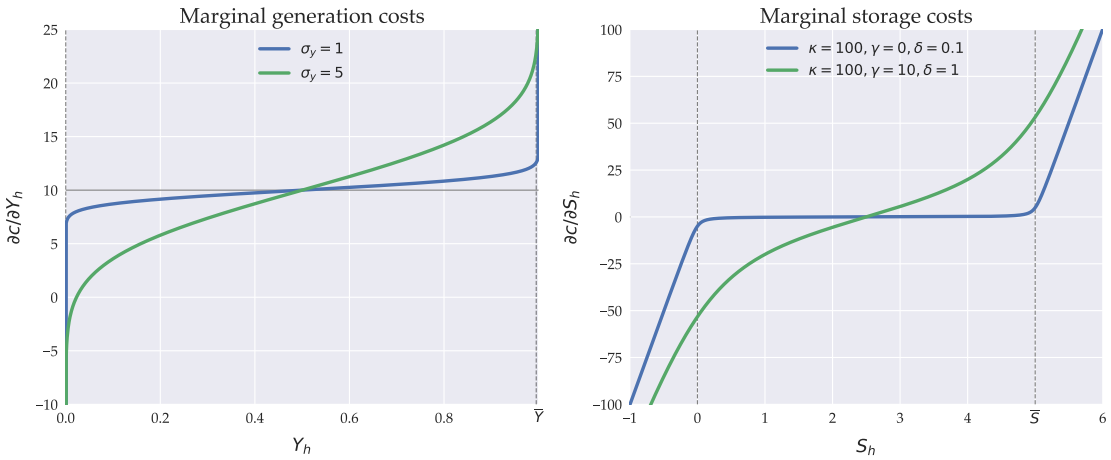
Unfortunately, we do not have enough variation in storage levels to estimate the parameters κ, γ and δ in the last term in (4b). The intuition is that storage constraints are never binding and hence, the last term in (4b) is always approximately zero. Instead, we choose the values of κ, γ and δ to balance the trade off between computational tractability and approximation errors.[27] Specifically, we set $\kappa = 100$, $\gamma = 0$ and $\delta = 0.0001$. This leaves us with 17 parameters to estimate, such that $\theta = \{\sigma_{y,i}\, \overline{Y}_i, \underline{Y}_i, \beta, \bar{c}\}$ for $i = \{\text{NO1}, \text{NO2}, \text{NO3}, \text{NO4}, \text{NO5}\}$. The parameter space Θ is restricted to

$$\Theta = \left\{\theta \in \mathbb{R}_+^{17} | \sigma_{y,i}, \underline{Y}_i, \bar{c} \geq 0, \beta \in [0, 1]\right\}. \tag{14}$$

[26] Note that in the representation in (4), any guess on p_h, λ_h returns a well-defined value of Y_h through (4a). However, with the way the marginal storage costs enter through (4b), a guess on a level of λ_h may result in an attempt to evaluate the marginal costs with a value of S_h outside the constraints $[0, \overline{S}]$ which causes a numerical problem if the costs are not well-defined in this region.

[27] The cost of storage should be approximately zero within the storage constraints.

Figure C.3: Marginal costs



Note: Capacities are set at $\underline{Y} = 0$, $\overline{Y} = 1$, $\overline{S} = 5$, and the cost level in the generation function is set at $\bar{c} = 10$.

C.4 Results

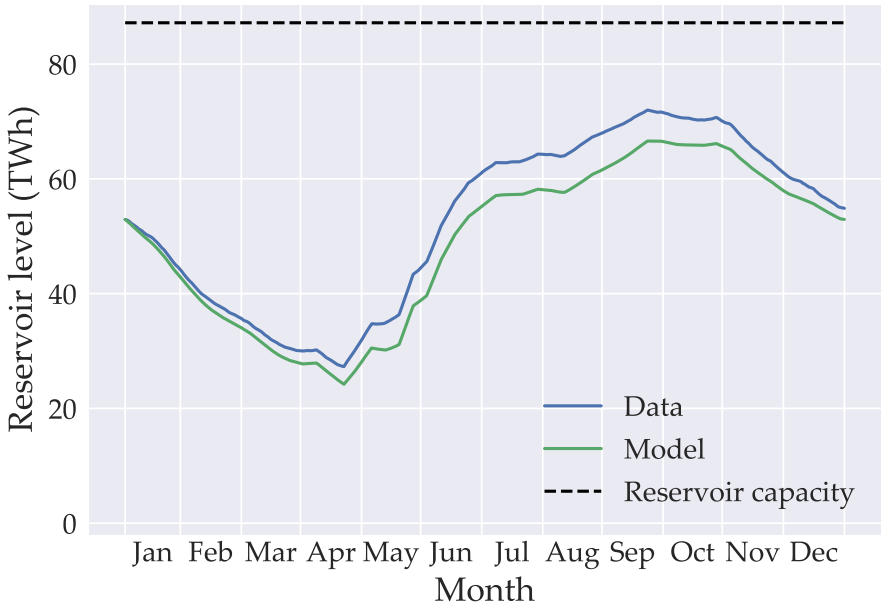
The full set of estimated parameters is given in table 2.

Table 2: Estimated parameters

		NO1	NO2	NO3	NO4	NO5
$\overline{Y}_i$	Maxmum generation capacity (MW)	1,891.1	10,492.5	3,387.4	3,854.5	6,238.7
$\underline{Y}_i$	Minimum generation capacity (MW)	1,242.6	1,418.9	1,591.3	782.3	2,350.7
$\sigma_{y,i}$	Shape parameter in marginal cost of generation	51.7	136.7	93.3	127.4	74.5
$\bar{c}$	Level parameter in marginal cost of generation (DKK/MWh)			3.0		
β	Hourly discount factor			1.0		

The main section 2.2 looks at how hourly generation predicted by the model matches data. A related check is to look at the implied reservoir levels of the model relative to data; figure C.4 does this. The figure shows the tendency of hydroelectric power plants to deplete reservoirs during the beginning of the year until mid spring, refilling it during summer until late September, and then depleting it once again from late autumn to the end of the year.

Figure C.4: Model fit of aggregate Norwegian reservoir level, 2019



C.5 The importance of seasons in the clustered model

Panel (a) of figure C.5 illustrates the fit of the reservoir level of the clustered model with $K = 25$ states. The clustered model shows the same tendency as the full model, where reservoir levels are low in the end of April, rising towards the beginning of October and falling towards the end of the year. What turns out to be crucial in capturing the seasonal structure in the storage level is the splitting of the year into different seasons. As illustrated in panel (a) of figure C.5, the clustered model with $K = 25$ states contains four seasons. These four seasons are estimated using conditional inference trees based on Hothorn et al. (2006) by assuming three seasonal splits in the exogenous hourly data. The estimation implies that January-April is the first season characterized by low water inflow and high prices. The second season is May-July with high water inflow and low prices. The third season is roughly August-September with moderate water inflow and relatively low prices. Finally, the last season is characterized with moderate prices and low water inflow. Our clustered representation of storage technologies ensures that the storage level in the clustered model crosses the storage level of the full model whenever the season changes. To illustrate this point even further, panel (b) of figure C.5 shows the reservoir level in clustered model with $K = 30$ states but no seasons. Here the K-means clustering algorithm is allocating hours in the beginning of the year together with hours in the end of year. Because the seasonality in the reservoir level is acyclical, the average storage level in the clustered model becomes the mean between a high level and a low reservoir level. Adding an additional five clusters does not mitigate this.

Figure C.5: Aggregate Norwegian reservoir level in full and clustered models

(a) Model til K=25 states and 3 seasonal splits (b) Model with K=30 states and no seasonal splits

Note: The seasonal structure of the exogenous data (here hourly prices and water inflow) is estimated using conditional inference trees based on Hothorn et al. (2006). In panel (a) we assume there are three seasonal breaks, giving rise to four seasons. In panel (b) we assume there is no seasonality.

D Clustering of data

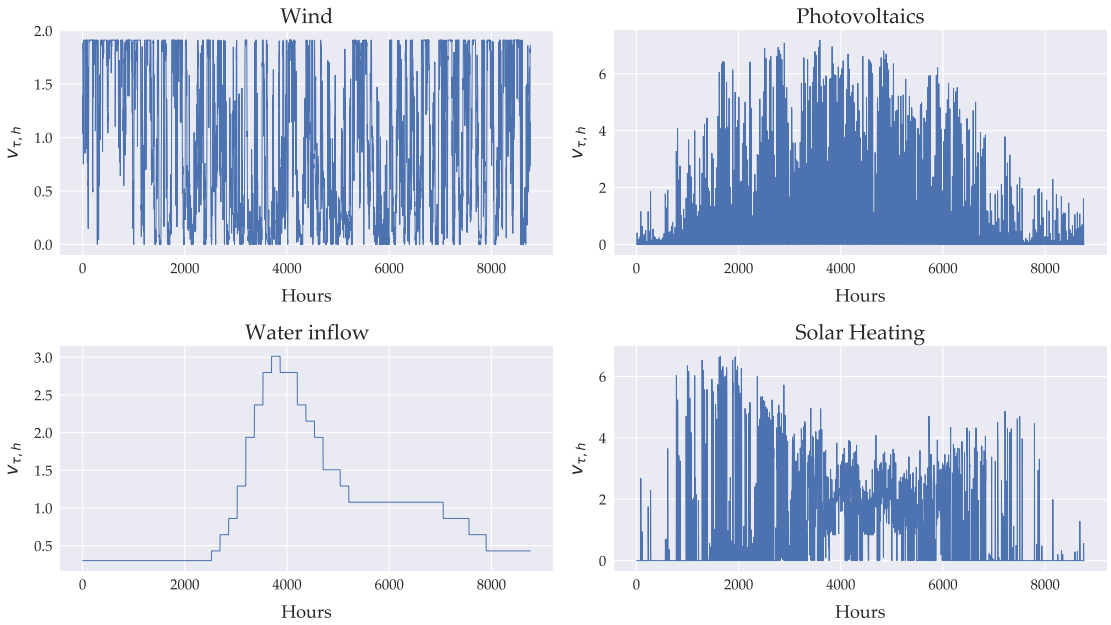
To scale down the size of the model we go through a number of steps: Conditional inference trees to detect chronological seasonal breaks, a simple K-means approach within each season to aggregate hours that are similar into representative stats, and *post* clustering fitting. We go through the steps in detail and provide supplementary evidence for how it works. We note, however, that as we include more than 100 variational patterns, we do not include figures for all types of variation, but only selected different types.

D.1 Ctree

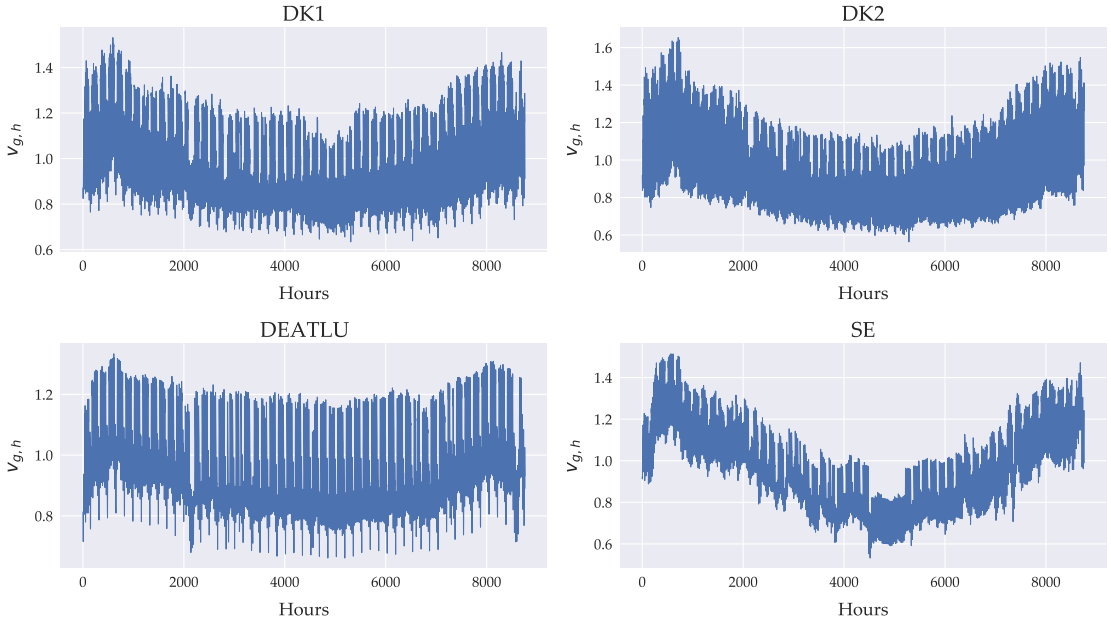
We consider four types of hourly variation: Generation capacity, demand, availability of transmission lines, and residual demand. Figure D.1 plots the variation for a selection of each of these types. We note that the strongest seasonal cyclicality is present in the water inflow, followed by the sunlight driven technologies (solar heating and photovoltaics, figure D.1a). The hourly demand, in particular in Sweden, is also clearly seasonal with smaller demand during the summer in general (D.1b). Availability of transmission lines (D.1c), however, does not exhibit any strong seasonality. In our preferred specification, where the number of intra-yearly states K is kept relatively small and $S = 4$, we use water inflow and residual demand variation in Ctree to detect seasonal breaks. For larger values of K, S we prefer to include data on variation in water inflow, demand, and sunlight technologies' generation capacity. We reiterate that we specifically focus on identifying the seasonality in water inflow, as this is important for modelling the behavior of hydro-electric plants with storage. If it was not for the storage plants, we would simply prefer to go straight to the K-means algorithm and skip the Ctree approach altogether.

Furthermore, we do not restrict seasons to start in hour 1 of the year. In other words, it is possible for a season to run e.g. 1000 hours from $h = 8000, ..., 8760, 1, 2, ...239$.

Figure D.1: Examples of hourly variation, unsorted

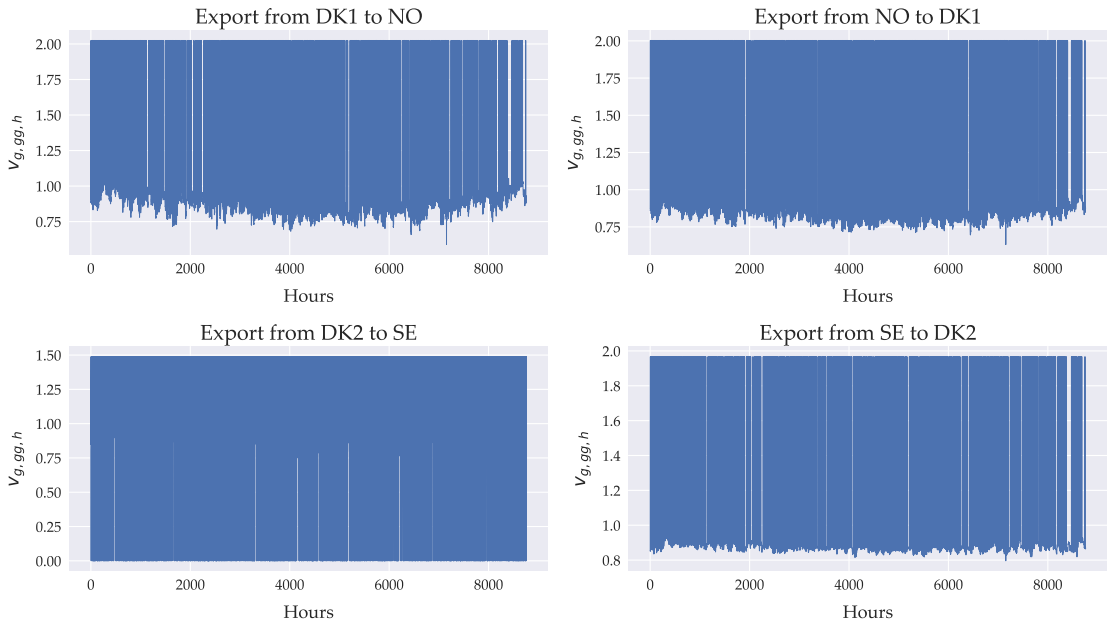


(a) Variation in types of generation capacity



(b) Variation in demand

The figure continues on the next page...



(c) Variation in transmission line availability

D.2 K-means

Within each season we apply the K-means algorithm to identify similar hours and aggregate them as representative states. For instance, the variation in hourly demand exhibits strong cyclicality within each day (24 hours) and weekly (weekend vs. workdays). Similarly, sunlight technologies are obviously more productive during the day.

Choosing how to perform the K-means clustering, however, depends on the circumstances, and what we seek to achieve. In general, aggregating hours into clusters is done in order to reduce the dimensionality of the model. In itself, the K-means clustering is used to identify similarities in data. Applying the K-means algorithm, we essentially attempt to maintain all correlational information on all relevant data simultaneously. One can say that the K-means algorithm targets a large number of low level objectives. However, for our purposes at least, we are not directly interested in obtaining a realistic correlation e.g. between when photovoltaic and wind technologies are productive. Instead, we are focused on the following high level objectives: First, we measure the validity of the clustering by the fit of intra-yearly variation in equilibrium prices. This set of statistics measures the state of the entire *system*, instead of the individual parts: The equilibrium price is high if demand is large relative to the intermittent supply, the potential for cheap imports is exhausted, and storage technologies are unable to balance this residual demand with cheap excess electricity from low price states. Second, we measure the validity of the clustering by the fit of technology-specific prices. The technology-specific price is essential for the value of a technology, and is thus instrumental in determining the optimal level of investments. Furthermore, it indicates whether or not the clustered model features a realistic correlational structure: If the average price received by wind technologies is lower than the average electricity price, the clustered wind variation correlates with low residual demand and/or hours with access to a lot of cheap import.

Focusing on the high level objectives allows us to obtain a very good model fit for very low number of clusters. However, this procedure comes at a cost: Specifically, when focusing on high level objectives, we cannot ensure that the clustering remains valid if the model moves far beyond the baseline scenario in which we have validated it. Specifically, our clustering method can be vulnerable to large changes in the composition of intermittent renewables. For instance, consider the case of Denmark, where in the baseline year, wind production is by far the dominant intermittent renewable source of energy, e.g. compared to photovoltaics. When we validate our clustered model by looking at how well the model produces intra-yearly variation in prices and yearly average technology-specific prices, we do not put a lot of weight on the variation in generation capacity of photovoltaics. Assume for instance that the clustered model has overestimated the correlation between generation capacity for photovoltaics and both wind capacity and hourly demand. In this case, the technology specific price for photovoltaics might be accurate in the baseline scenario, but become increasingly wrong if wind capacity becomes less important over time, and capacity of photovoltaics increase. In other words: The validity of the clustering of the model, relies on the model not moving too far away from the baseline scenario in terms of the composition of intermittent renewables.

Given these considerations, the data that we use in clustering the model depends crucially on the number of available clusters. For instance, if we only have K=25 clusters, the K-means algorithm is used to identify variation in residual demands and water inflow; the former, as this is essential for the formation of equilibrium prices (our high level objective) and the latter to obtain a realistic model for hydro-electric storage plants. In instances where K increases to e.g. 100 or even 1000, we use the K-means algorithm to identify the variation in individual patterns $v_{x,h}$, as this allows us to target the low level objectives. This makes validity of the clustering more robust outside the baseline scenario of the model.

The number of hours assigned to each season is simply based on the share of hours allocated to each season. The K-means algorithm minimizes the euclidean distance of the clustered ($v_{x,k}$) and hourly values of variables $v_{x,h}$ by partitioning the H_s hours in the given season into K_s clusters following the mapping $\mathcal{S}_k$:

$$\min \sum_{k} \sum_{h \in \mathcal{S}_k} \left(v_{x,h} - v_{x,k}\right)^2, \qquad v_{x,k} \equiv \frac{1}{n_k} \sum_{h \in \mathcal{S}_k} v_{x,h},$$

where n_k is the number of hours assigned to the cluster k.

Finally, we perform the *post* cluster fitting of the clustered data. First, we fit each individual clustered variable to the full range of hourly data. We do this by mapping the clustered values back to hourly values using the $\mathcal{S}_k$ mappings, and sort this data. Then, we sort the corresponding hourly data, and match the level in the sorted clustered data with the levels in the full hourly data; thus, essentially, the K-means algorithm is only used to identify the ordering of the clusters: The levels in all clusters are fitted to the hourly levels afterwards. Second, we make sure that p% of

the tail observations are adjusted to reflect the minimum/maximum of the full hourly distribution. To balance the clustered variable, the middle $1 - 2p\%$ of the clustered observations are used to balance the variable to make sure that $\sum_k v_{x,k}$ remains unchanged. Finally, we adjust the variation in demand in order to fit the full hourly data variation in residual demand functions. This means that all variables except for hourly demand, $v_{g,k}$, is fitted to represent the full variation in its hourly counterpart. The clustered demand component, in turn, ensures that residual demands are fitted to their hourly counterparts as well.

The following sections go through the effects of clustering and goes into more detail about how the *post* clustering adjustments change the results as well.

D.3 Clustering with K=25

The K-means algorithm is focused on identifying variation in residual demand and water inflow. Thus, in countries as Denmark where wind energy is an important source of electricity, the simple K-means clustering naturally obtains a relatively good fit for wind producers. Figure D.2 illustrates this. Similarly, as variation in water inflow is used in the conditional inference tree to detect seasonal breaks and in the K-means clustering algorithm, this fit is naturally also pretty good with the simple K-means algorithm. Figure D.3 illustrates this.

Figure D.2: Clustering of variation in generation capacity for a type of wind producers

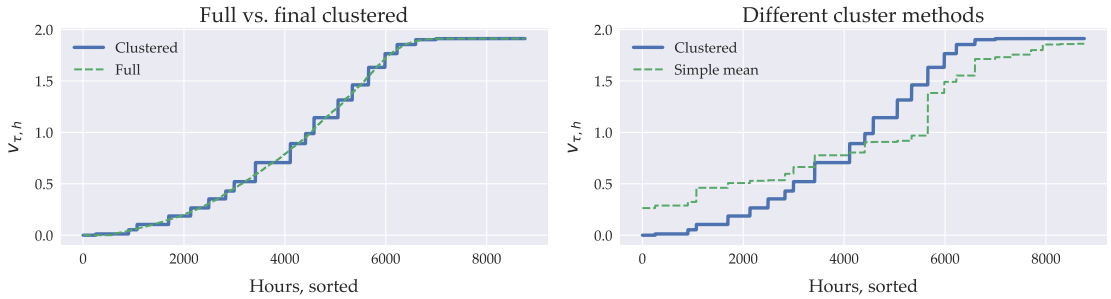


Figure D.3: Clustering of variation in a type of water inflow

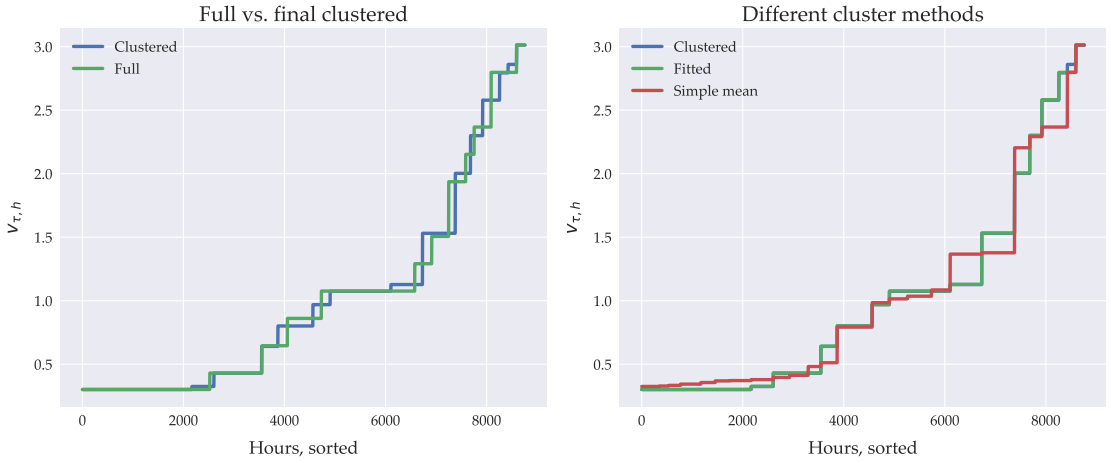
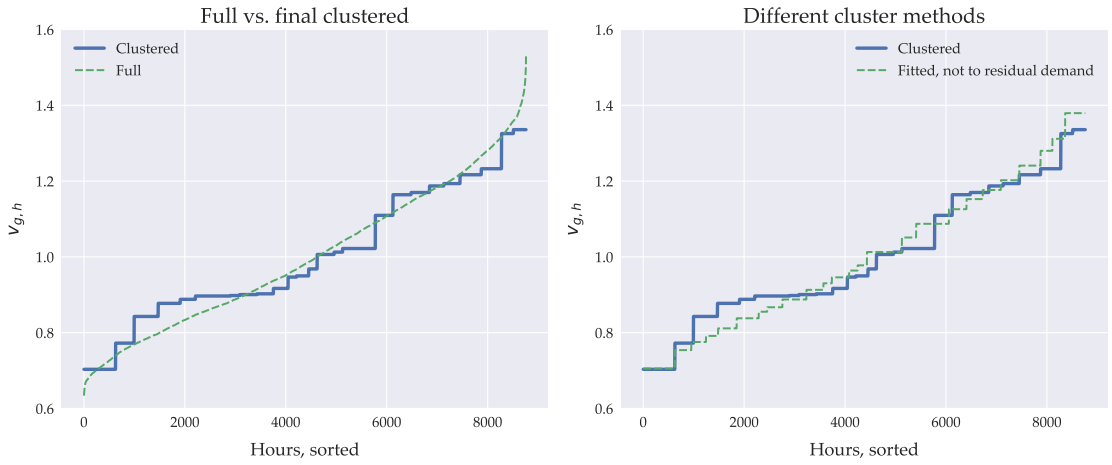


Figure D.4 illustrates the effect of targeting the variation in residual demand directly. As the left panel shows, the our final 'clustered' fit is not as accurate as for wind, water, and residual demand proved to be. The reason for this is simply that whereas all the other variables (e.g. generation capacity for wind in figure D.2) are manipulated to fit the full hourly data, the variation in demand is used to target the residual demand. The reason for this is that residual demand is an endogenous variable, and thus cannot simply be fixed at appropriate levels.

Figure D.5 shows an example of how our method performs with an hourly variation that is neither taken into account in the conditional inference tree nor the K-means clustering algorithm. The variation in availability of transmission lines are not a priority in the aggregation, as the variation to a relatively large degree is random. In this case, the simple K-means clustering

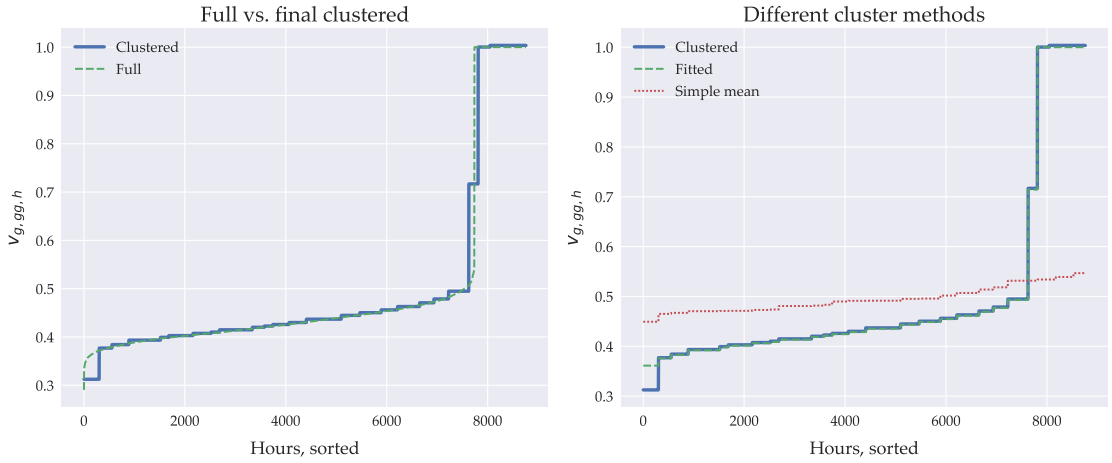
Figure D.4: Clustering of variation in hourly demand



Note: The figure shows the effect of clustering with $K = 25$ and $S = 4$. The path for 'Clustered' uses the preferred clustering method that includes both fitting of model variation including residual demand, and adjustments to capture tail probabilities.

algorithm barely includes any variation at all; this is drastically altered by the fitting procedure.

Figure D.5: Clustering of variation in trade capacities



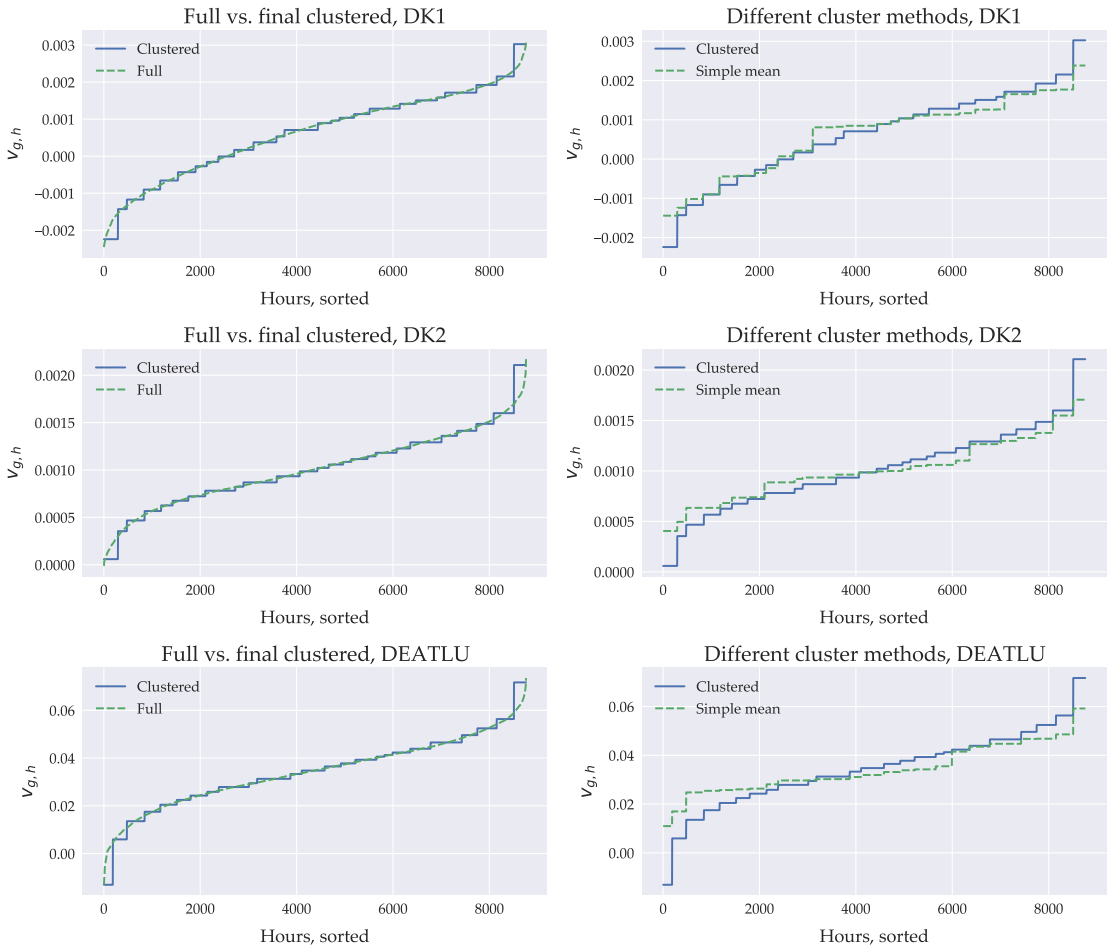
Note: The figure shows the effect of clustering with $K = 25$ and $S = 4$. The path for 'Clustered' uses the preferred clustering method that includes both fitting of model variation including residual demand, and adjustments to capture tail probabilities.

As previously mentioned, both algorithms (Ctree and K-means) are specified to identify variation in residual demand patterns. Furthermore, for the K-means algorithm, we specify that variation in domestic residual demand is important relative to foreign.[28] The result can be seen in figure D.6: The simple K-means clustering does pretty well in this case, in particular for the domestic areas (DK1,DK2). However, the post-cluster fitting ensures that the whole distribution

[28]Specifically, the K-means algorithm minimizes the Euclidean distance between hourly and clustered data; in this objective function, we weigh the variation in domestic residual demand with a factor 2, foreign residual demand 1, and variation in water inflow 0.5.

is included in the clustered version.

Figure D.6: Clustering of variation in residual demand



The *post* clustering adjustments may, as mentioned earlier, result in a worse fit in terms of correlations: The adjustments to fit the full hourly volatility is made for each variable separately, and does not take correlations into account. The effect of these adjustments are plotted in the following two figures for the case of wind: Figure D.7 plots the estimated correlation between variation in productivity of a Danish wind technology against variation in residual demand. In this case, the correlation is in fact slightly improved by the post-clustering adjustments. There is, however, no guarantee that this will generally be the case in general. However, evaluating the technology-specific prices, gives an indication that the correlation between a specific generation capacity variable and equilibrium outcomes are realistic; this is the second high-level objective that we evaluate in the main section 3. The figure D.7 shows the estimated correlation for wind productivity and variation in hourly demand. In this case, the final clustered version outperforms the 'simple mean' clustering method again. In this case, however, there is still some discrepancy between our final clustered version and the 'full' model with hourly data.

Figure D.7: Effect of clustering on correlations: Residual demand and wind.

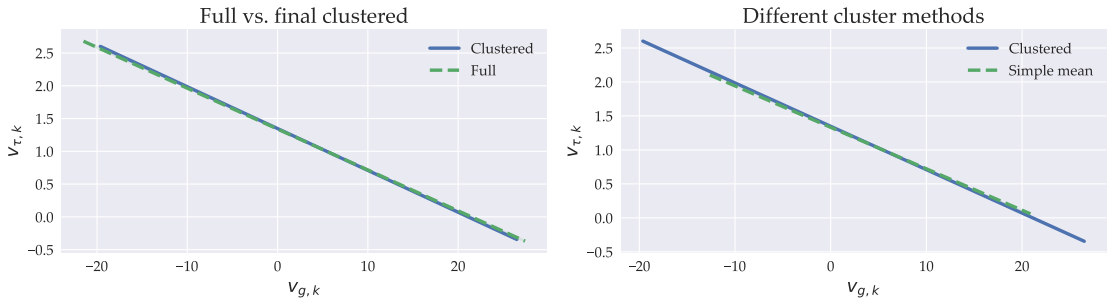
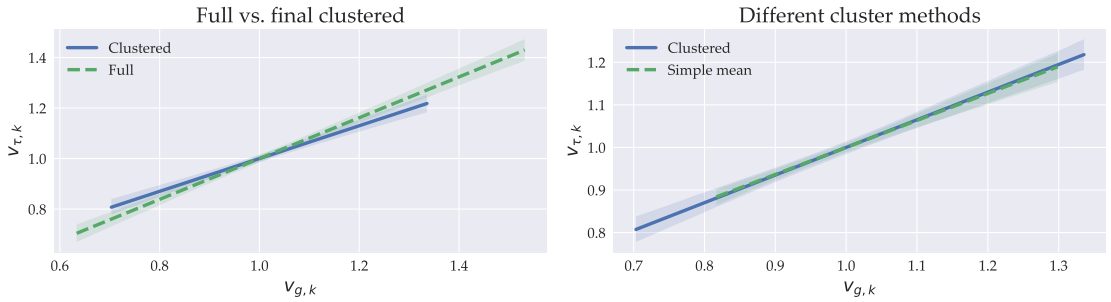


Figure D.8: Effect of clustering on correlations: Demand and wind.



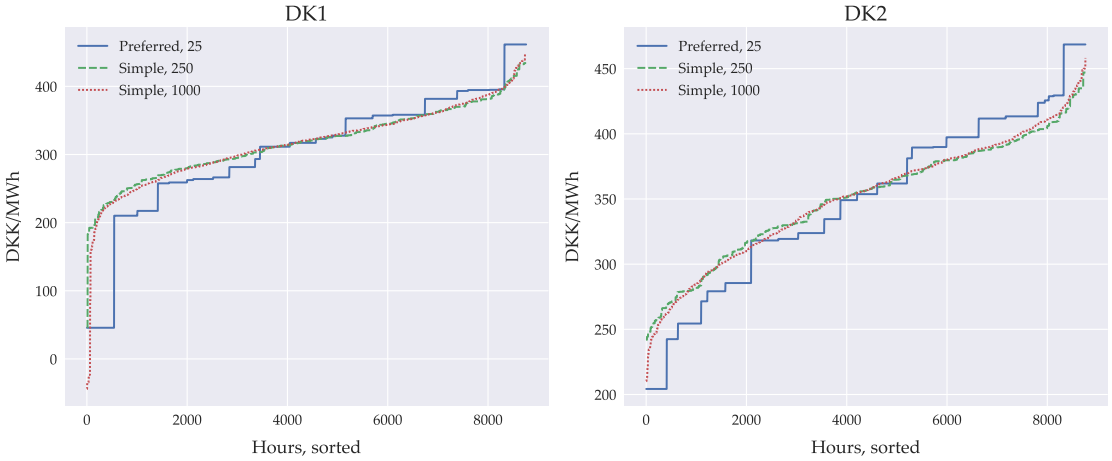
D.4 The effect of changing K

To investigate the effect of increasing the number of clusters, we compare our preferred clustering method with $K = 25$ to a more simple clustering method with varying K. In our preferred method, described further in the sections above, we only parse aggregate information to the K-means algorithm, and perform a number of *post* clustering adjustments to fit the variation in data. While this ensures that the intra-yearly variation is captured, the *post* clustering adjustments may have the adverse effect of skewing the correlations between different type of intra-yearly patterns. Thus, when we consider the effect of increasing K, the relevant alternative to our method is one where we use a large K and parse a lot of information to the K-means algorithm, in order to avoid performing the *post* clustering adjustments. To this end, the following investigates the effect of increasing K drastically, but employing a simple K-means algorithm that uses all available information on intra-yearly variation.

Figure D.9 illustrates the effect of increasing K dramatically from 25 to 250 and 1000. The figure confirms that, in terms of capturing price variation, the preferred clustering method performs as well as the simple method with $K = 1000$.

Next, figure D.10 illustrates the model's ability to capture the downlift on wind, measured as the difference between average prices and the specific prices received on wind. The figure shows that the preferred clustering method with $K = 25$ yields a similar result as the simple method with large K's, but still closer to the observed levels in data in our baseline year 2019.

Figure D.9: Intra-yearly variation in prices with varying K



Note: The "preferred" clustering method uses aggregate information to cluster data and *post* cluster adjustments. The "simple" clustering method uses a lot more information and relies on a conventional K-means approach. Both approaches initially uses conditional inference trees to identify seasonal breaks in data.

Figure D.10: Downlift on wind with varying K

Note: The "preferred" clustering method uses aggregate information to cluster data and *post* cluster adjustments. The "simple" clustering method uses a lot more information and relies on a conventional K-means approach. Both approaches initially uses conditional inference trees to identify seasonal breaks in data.

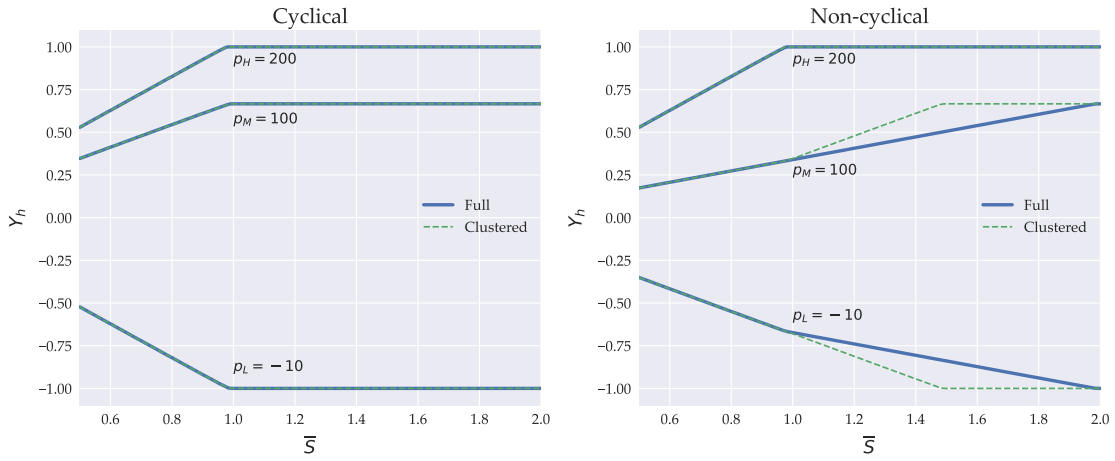
E Simulations with the state-average approximation

To illustrate the effect of the state-average approach, we set up a simple toy model with a number of simple scenarios. In all of the scenarios, we consider a small number of hours (7), where prices fluctuate between load ($p_L = 10$), medium ($p_M = 100$), and high ($p_H = 200$) levels. To isolate the effect the methodological effect of approximating optimization problem of storage plants, we assume no price variation within the three clusters.

We consider the simple case of a battery-like technology that does not have any inflow, and a dispatch efficiency of roughly 1. The generation capacity is normalized to $[-1, 1]$ such that it, at most, can charge/discharge one unit. The marginal cost is set at 5. The function g_y is chosen as a standard normal cumulative distribution with $\sigma_y = 10$. The marginal cost of the state constraint is a smooth piecewise linear function as suggested in appendix The marginal cost of the state constraint is a smooth piecewise linear function as suggested in appendix B.1 with slope $1e4$ when crossing the state constraints $[0, \overline{S}]$.

In the first case, we simulate the effects of both short run and long run storage, but in a way that is cyclical: For the first 6 hours the model alternates between low and medium price states. In the longer run (hour 7), the price is, however, at its highest (state H). This creates an incentive to partially save for the short run (state M instead of L), but simultaneously gradually store for the long run. The left panel in figure E.1 illustrates that for this cyclical case, the clustered model is spot on - both in cases where the state constraint is almost always binding, when it is only occasionally binding, and when it is never binding. Importantly, the clustered method will generally be spot on in both the "always binding" and "always never binding" cases. However, it turns out that the "occasionally binding" scenario is a result of the neat cyclicality.

Figure E.1: Supply as a function of the storage capacity



To show the point that the "occasionally binding" constraint depends on the cyclicality of the clustered model, we consider the case with mostly the same setup, except for the order of the first six hours: Instead of alternating between low and medium price states, we assume a more non-

cyclical order (specifically M, L, L, M, M, L, H). The right panel in figure E.1 illustrates the result; In the intermediate ranges of $\overline{S}$, the clustered model deviates slightly from the the simulation of the full hourly model.

To drive the points home further, we may consider the supply functions for all three states as a function of a specific price. Figure E.2 illustrates the case with a relatively large storage capacity $\overline{S} = 2$, where we vary the price in low states from the original -10 to 300. When the storage constraints are rarely binding, the approximation is accurate for both cyclical/non-cyclical data.

Figure E.2: Supply as a function of the price p_L, with $\overline{S} = 2$

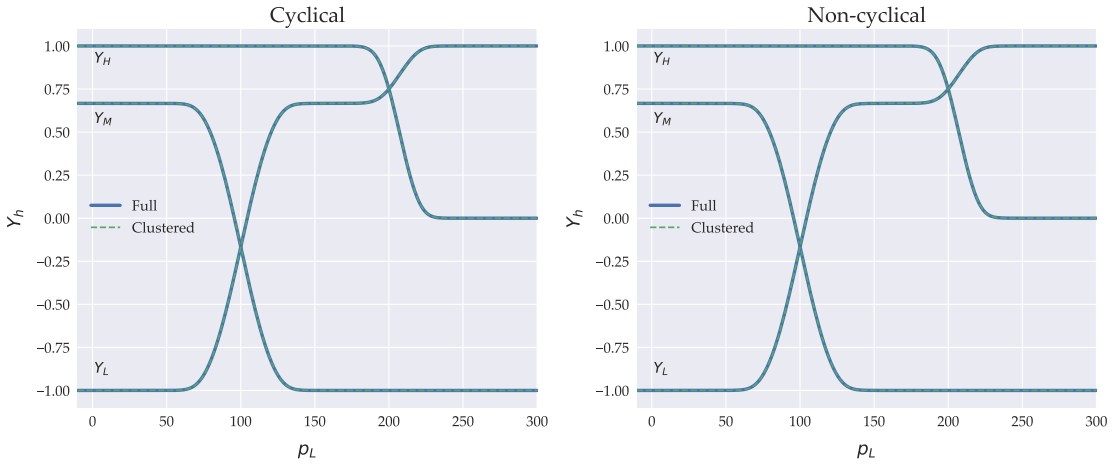
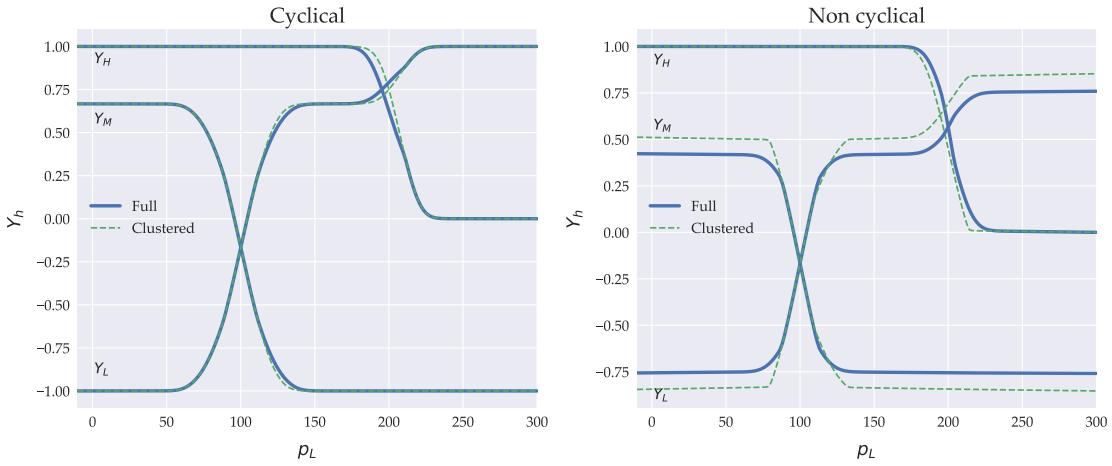


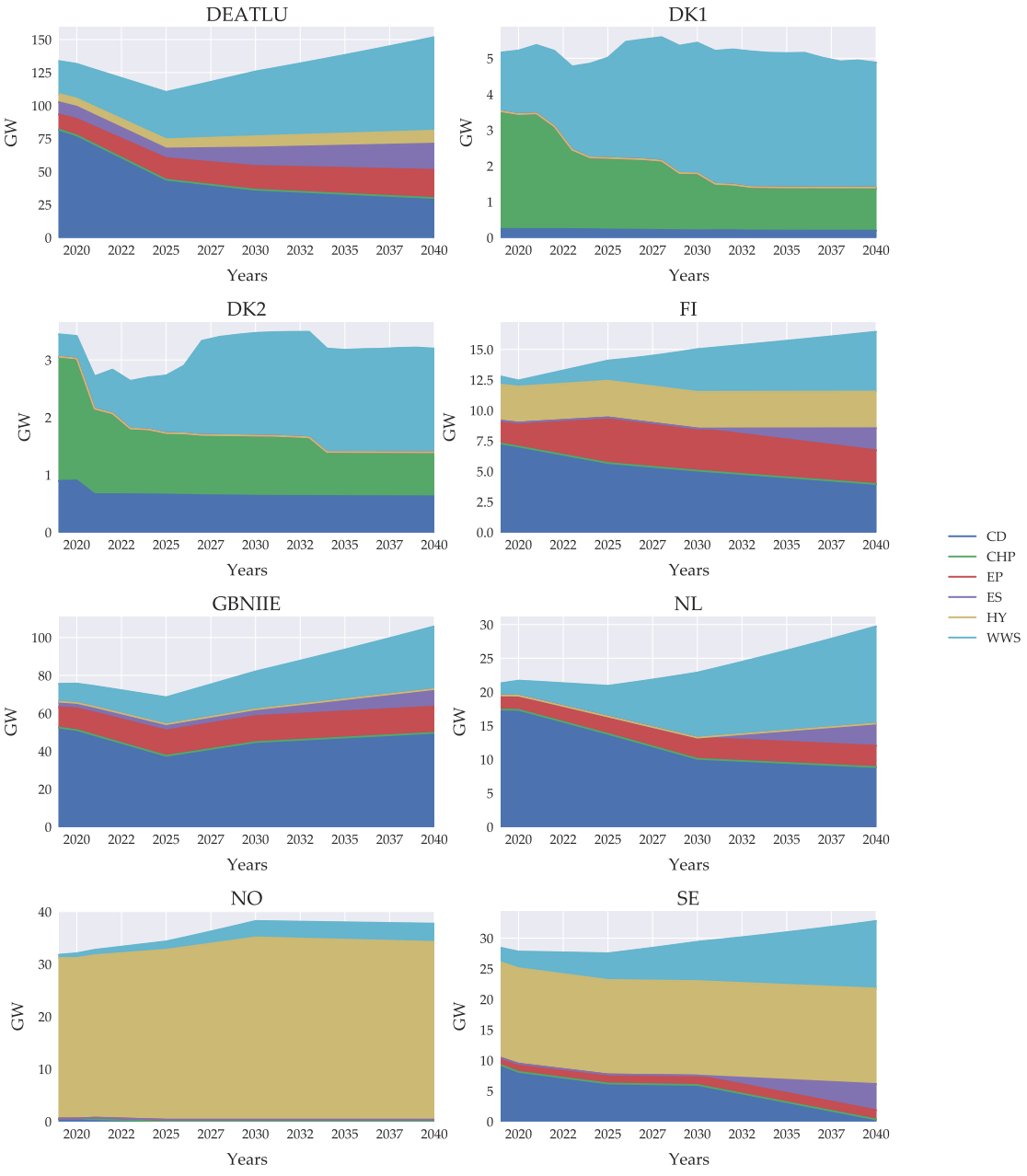
Figure E.3 emphasizes this point by repeating the average supply in each state, but in the "occasionally binding" ranges with $\overline{S} = 1.25$: It is only in the combination of "occasionally binding" and "non-cyclical" regime where the clustered model deviates slightly. This result is independent of the specific price variation.

Figure E.3: Supply as a function of the price p_L, with $\overline{S} = 1.25$



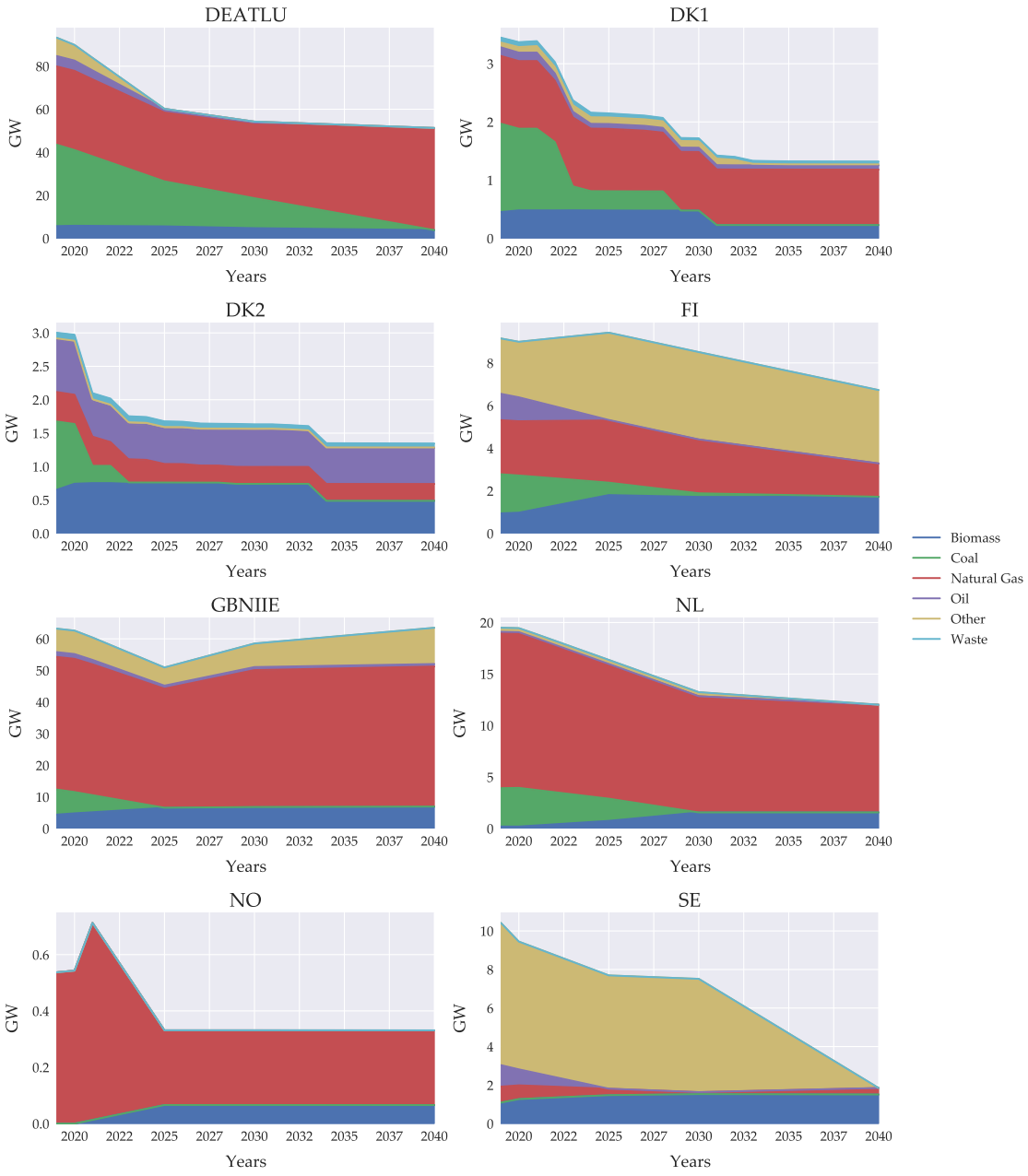
F Supplementary figures for baseline scenario

Figure F.1: Hourly average generation capacity for each country, by plant type



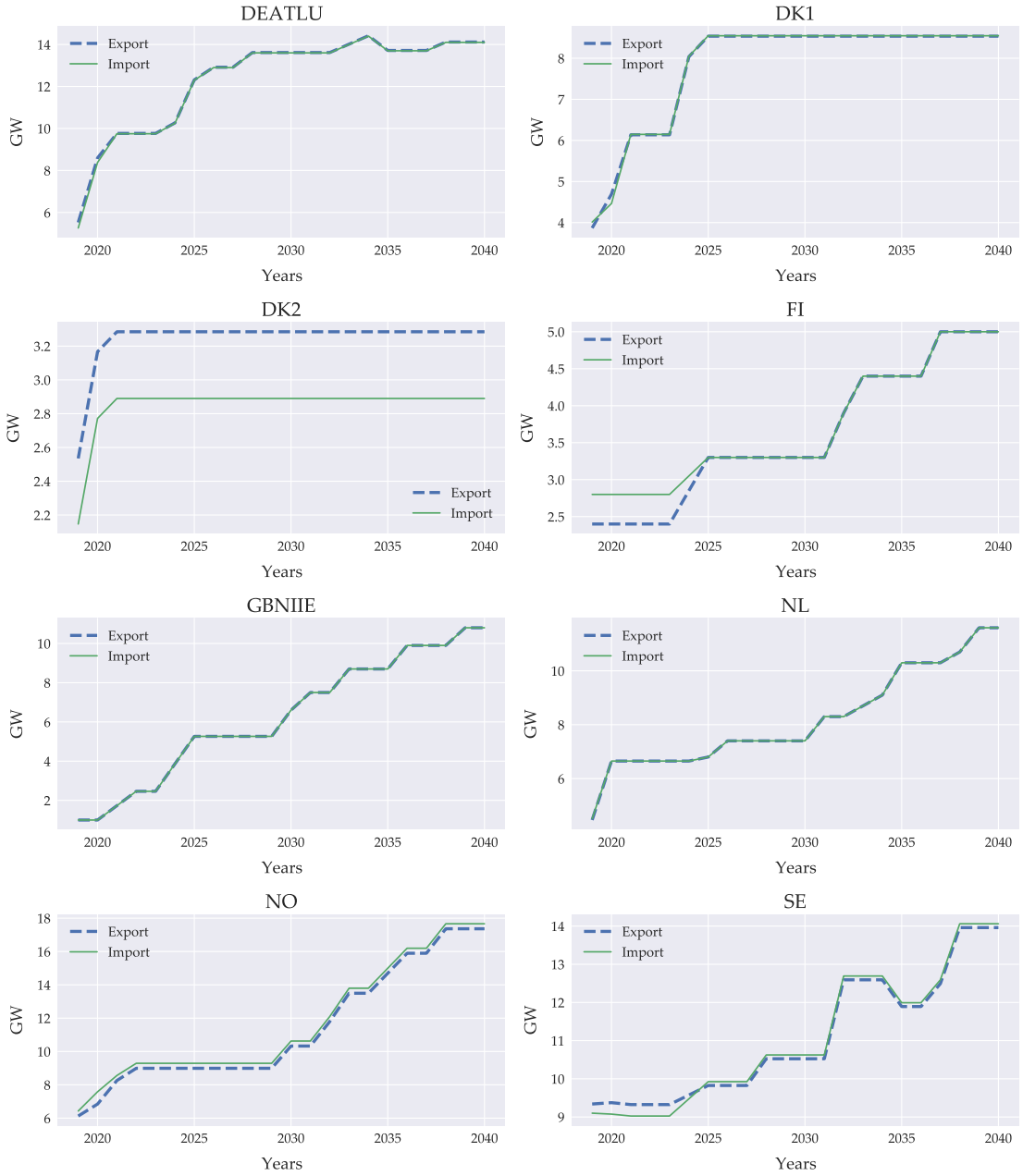
Note: Abbreviations denote condensation plants (CD), Combined Heat and Power plants (CHP), Exogeonous Production (EP), Electricity Storage (ES), Hydro plants with reservoir (HY), Wind, Water, and Sunlight (WWS).

Figure F.2: Hourly average generation capacity for each country, by fuel type



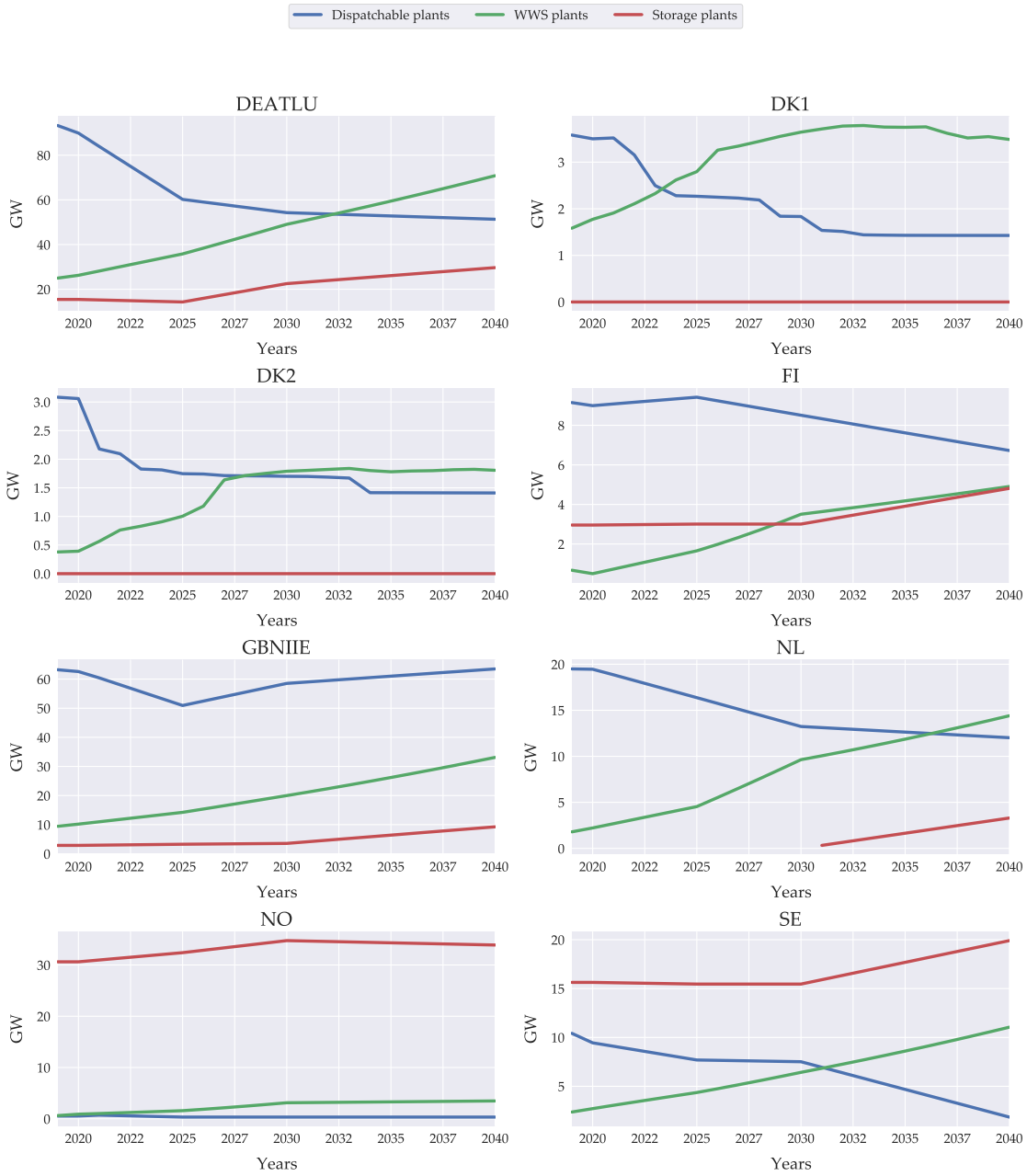
Note: The category "coal" covers coal and lignite, "oil" covers both fuel and gasoil, "Natural Gas" also covers synthetic natural gas, biomass covers various types of wood (pellets, chips, waste) as well as straw and peat. Finally, the category "other" covers biogas, biooil, hydrogen, and uranimum.

Figure F.3: Total Hourly Transmission Capacity by Country



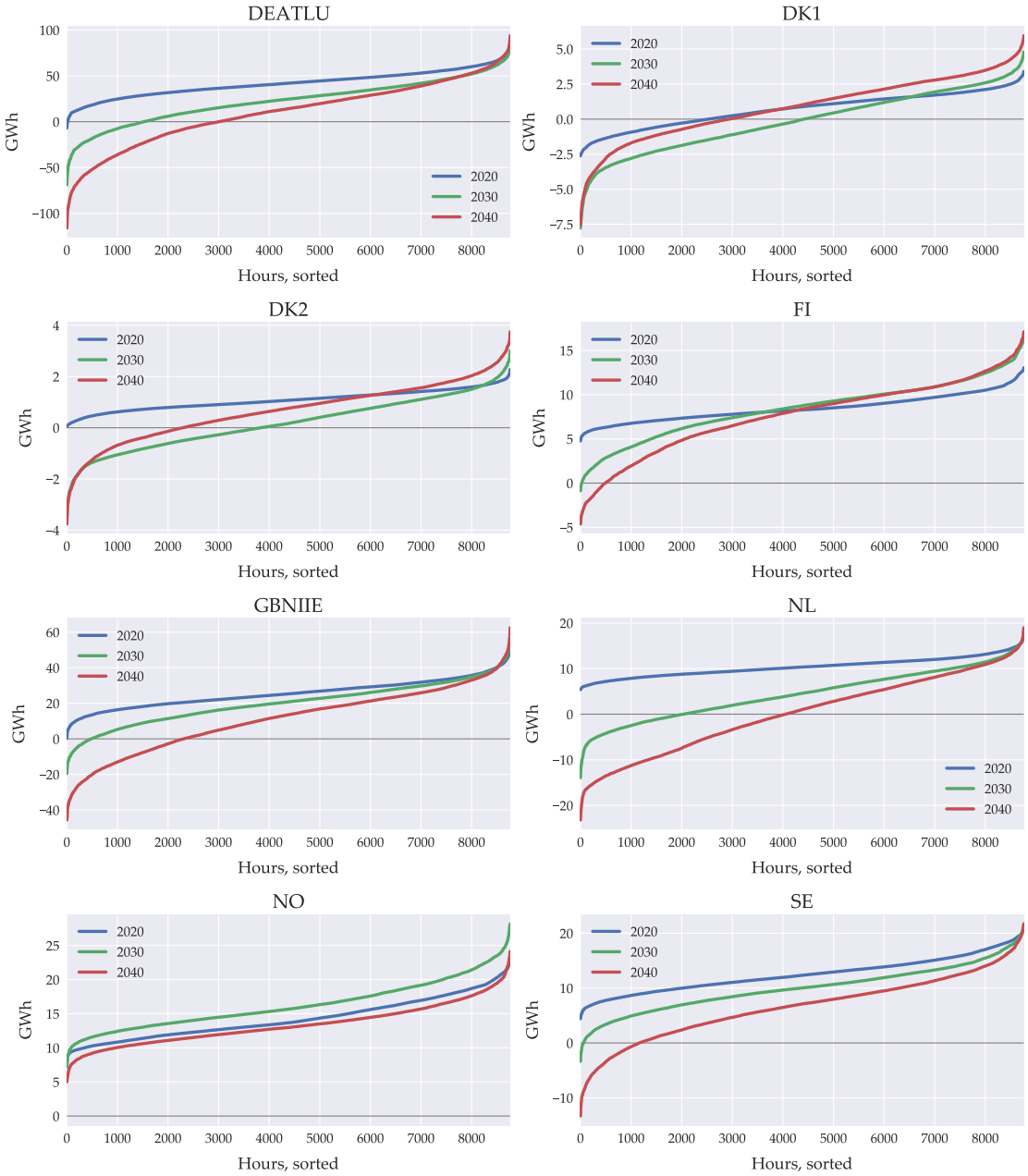
Note: The abbreviations denote the geographic areas Western Denmark (DK1), Eastern Denmark (DK2), Germany, Austria, and Luxembourg (DEATLU), Finland (FI), Great Britain, Northern Ireland, and Ireland (GBNIIE), Netherlands (NL), Norway (NO), and Sweden (SE).
Source: Own calculations based on data from the Danish Energy Agency.

Figure F.4: Aggregate hourly average generation capacity, types of plant



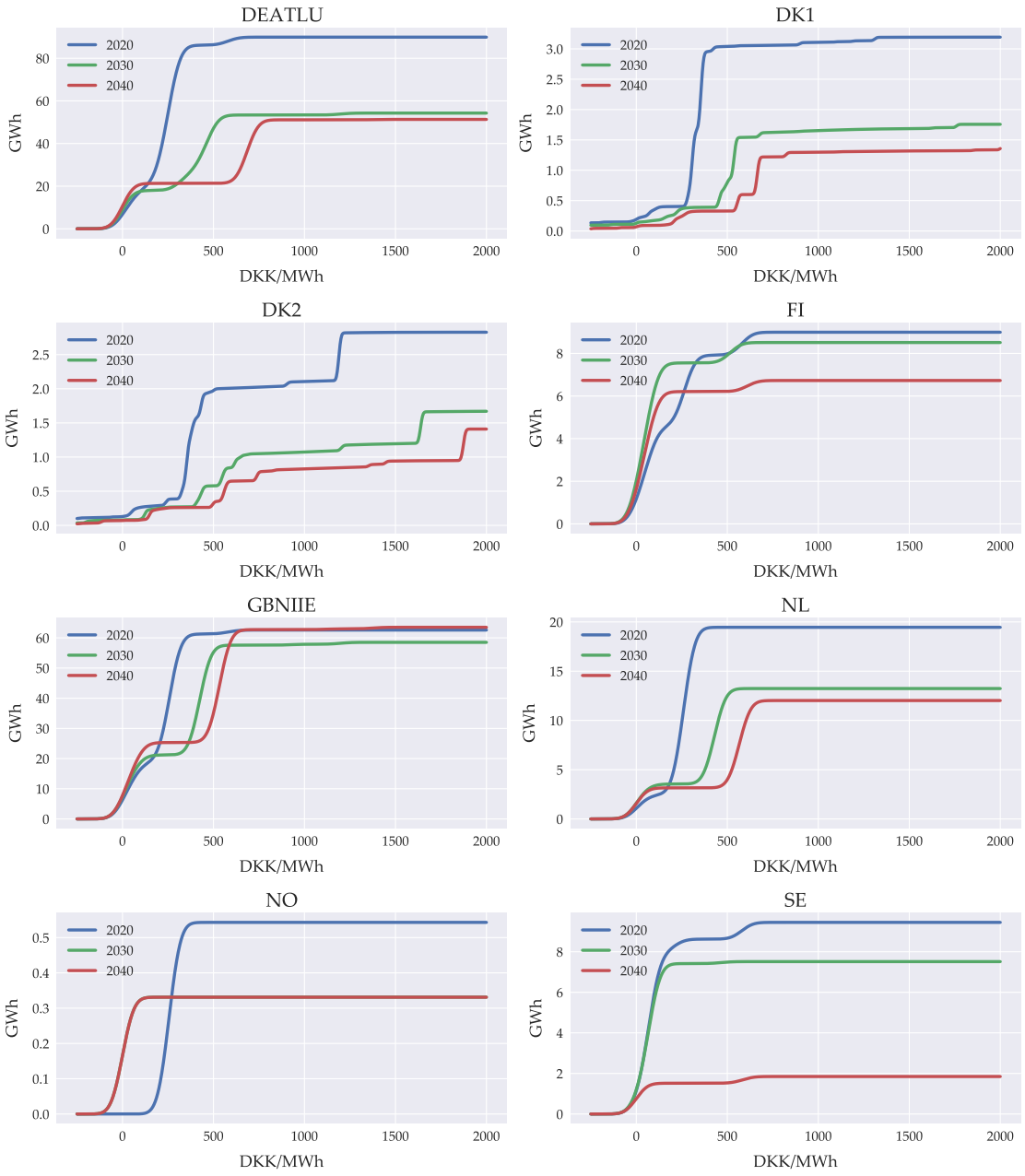
Note: The abbreviations denote the geographic areas Western Denmark (DK1), Eastern Denmark (DK2), Germany, Austria, and Luxembourg (DEATLU), Finland (FI), Great Britain, Northern Ireland, and Ireland (GBNIIE), Netherlands (NL), Norway (NO), and Sweden (SE).
Source: Own calculations based on data from the Danish Energy Agency.

Figure F.5: Sorted residual demand on geography and years



Note: The abbreviations denote the geographic areas Western Denmark (DK1), Eastern Denmark (DK2), Germany, Austria, and Luxembourg (DEATLU), Finland (FI), Great Britain, Northern Ireland, and Ireland (GBNIIE), Netherlands (NL), Norway (NO), and Sweden (SE).
Source: Own calculations based on data from the Danish Energy Agency.

Figure F.6: Aggregate supply from dispatchable plants, non-storage

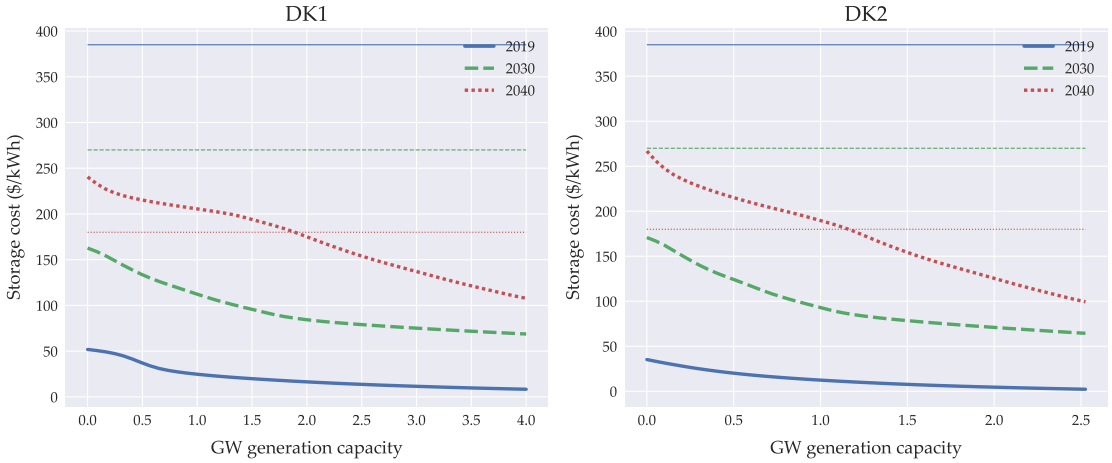


Note: The abbreviations denote the geographic areas Western Denmark (DK1), Eastern Denmark (DK2), Germany, Austria, and Luxembourg (DEATLU), Finland (FI), Great Britain, Northern Ireland, and Ireland (GBNIIE), Netherlands (NL), Norway (NO), and Sweden (SE).

G Supplementary figures for simulations

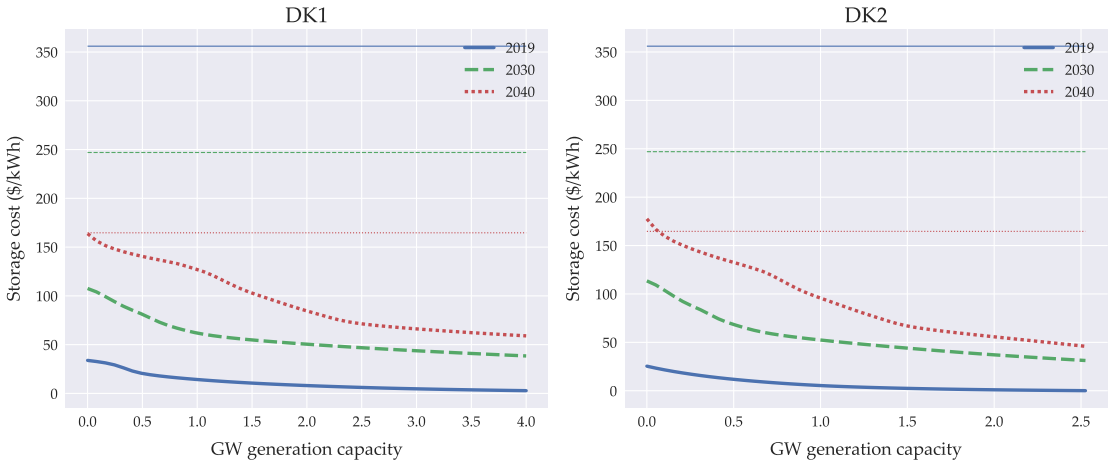
Break even investment costs:

Figure G.1: Break even investment costs of storage, 4h duration battery



Note: The discounted short run profits of the storage technology assumes an discount rate of 8% and a technical lifetime of 20 years. The vertical lines are estimated costs from Baxter et al. (2020).

Figure G.2: Break even investment costs of storage, 10h duration battery



Note: The discounted short run profits of the storage technology assumes an discount rate of 8% and a technical lifetime of 20 years. The vertical lines are estimated costs from Baxter et al. (2020).

Change in key variables:

Figure G.3: Change in key variables relative to case without storage, 2h duration

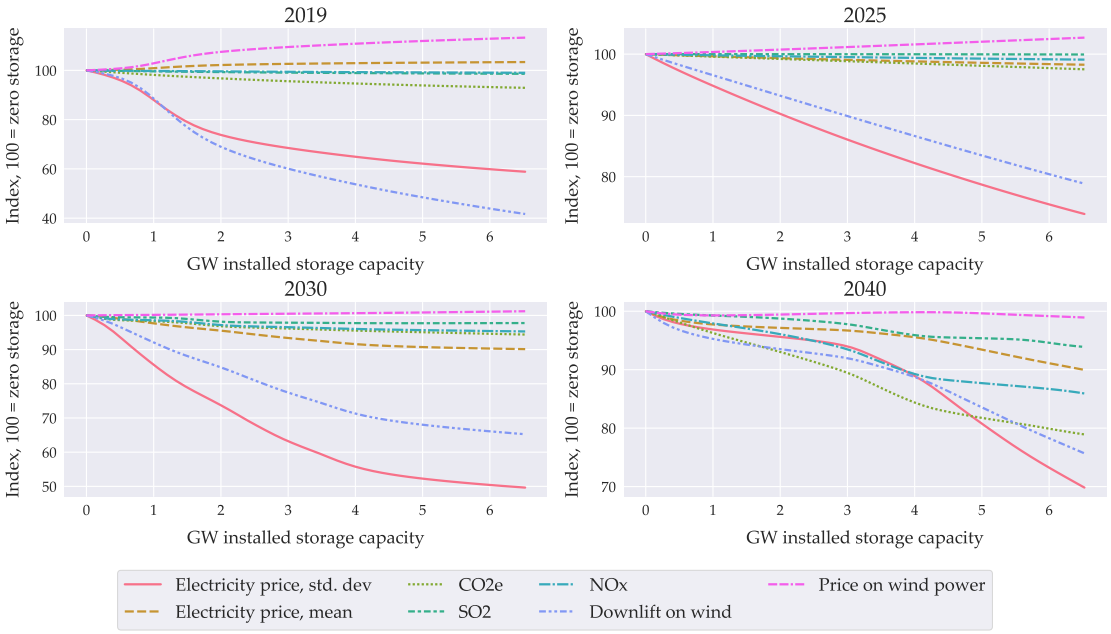


Figure G.4: Change in key variables relative to case without storage, 4h duration

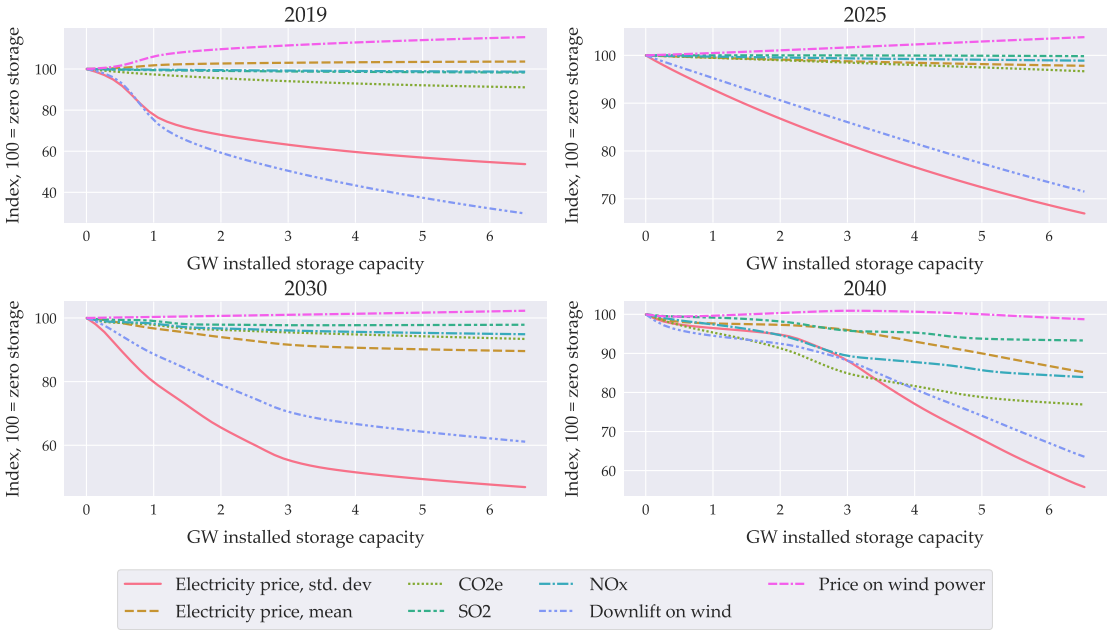
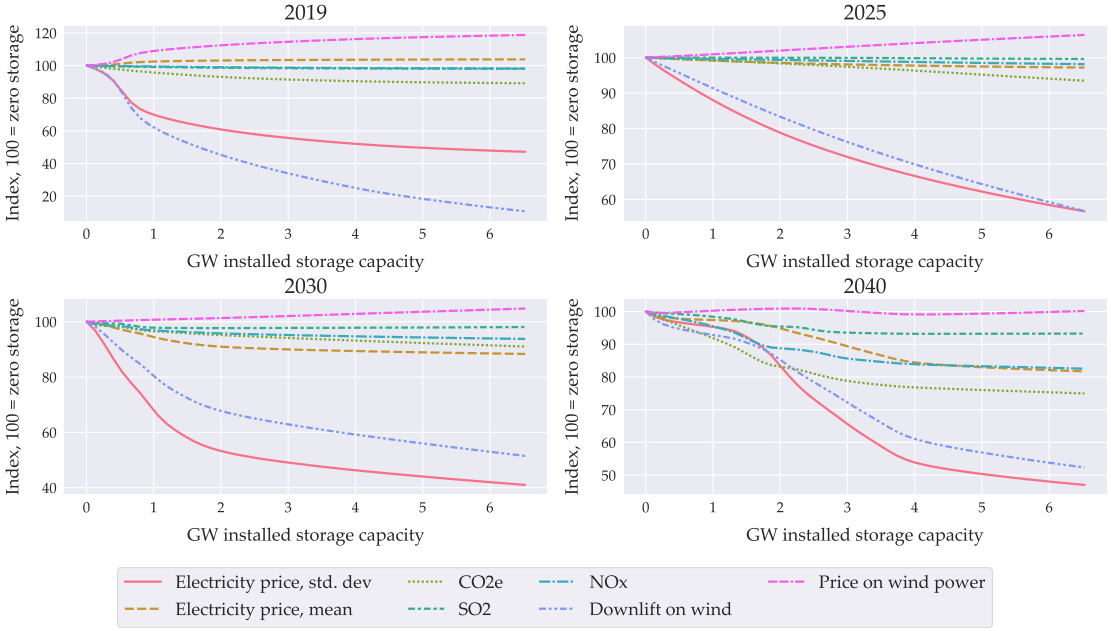


Figure G.5: Change in key variables relative to case without storage, 10h duration



Change in emissions

Figure G.6: Baseline emission trajectory

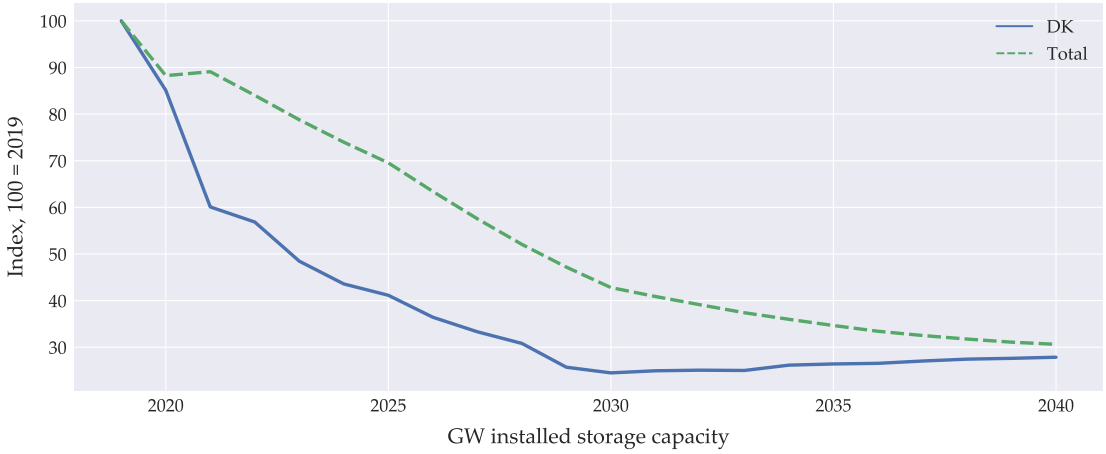
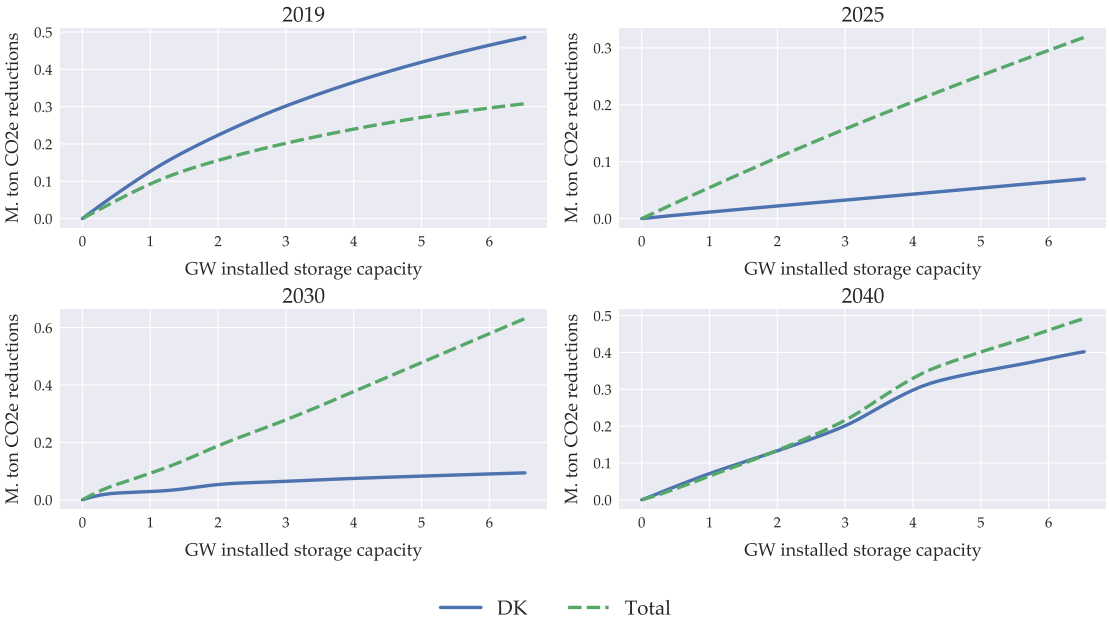
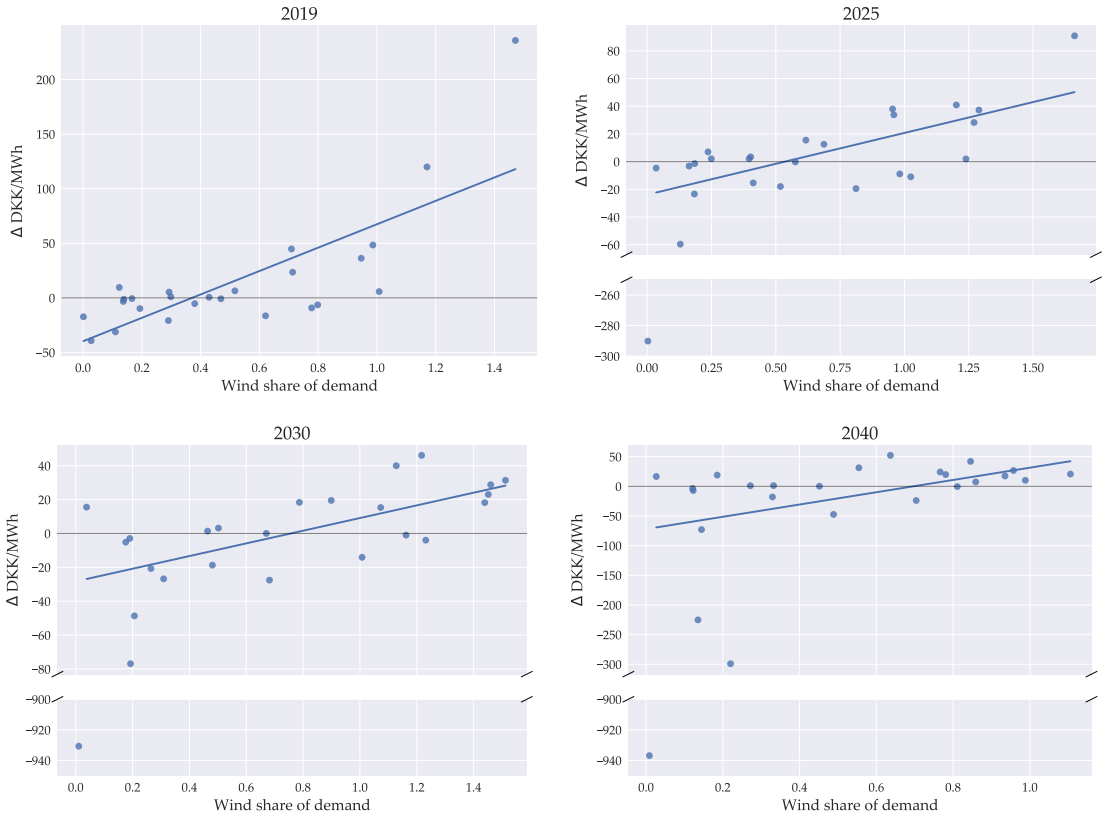


Figure G.7: Emission reductions



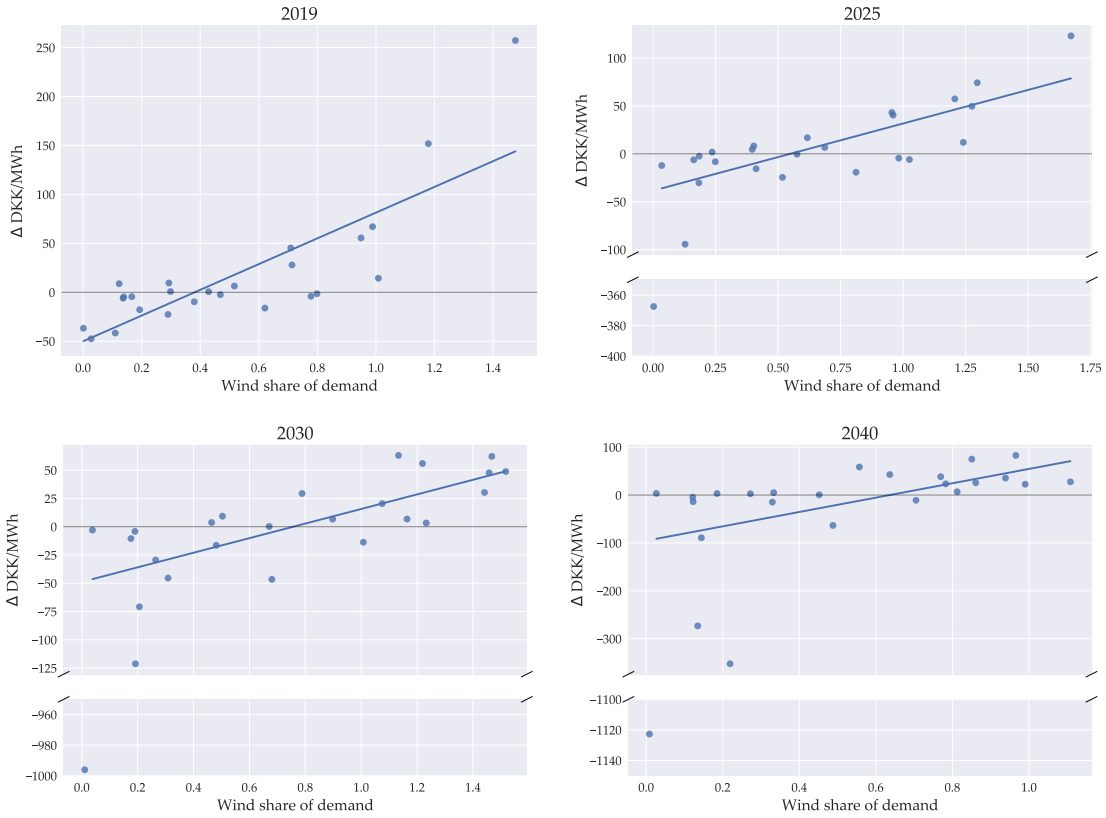
Change in hourly prices compared to wind production

Figure G.8: The effect of storage on prices and Danish wind production, 4h duration



Note: The change in price is the difference between full scale storage (6.5 GW generation capacity) and no storage. The lines indicate simple least square estimations of the correlation between the change in prices and the share of wind production.

Figure G.9: The effect of storage on prices and Danish wind production, 10h duration



Note: The change in price is the difference between full scale storage (6.5 GW generation capacity) and no storage. The lines indicate simple least square estimations of the correlation between the change in prices and the share of wind production.

Part 3

Solving the Big MAC Challenge: Endogenizing Technological Marginal Abatement Costs

Solving The Big MAC Challenge:*

Endogenizing Technological Marginal Abatement Costs

Rasmus K. Berg† Jonathan Leisner‡ August T. Nielsen§

December 29, 2021

Abstract

The paper develops a method for integrating technology data on emissions abatement in economic models. The methodology provides better estimates of the marginal costs of abating emissions - a key variable for designing effective environmental and climate policy - as it addresses four issues with existing methods: First, it distinguishes between end-of-pipe and input-displacing technologies and shows that these are substitutes; adopting end-of-pipe technologies reduces emissions related to inputs, thus making dirty technologies cheaper relative to competing clean input-displacing ones. Second, it ties the use of technologies to the relevant use of resources, thus making the costs of technologies a function of general equilibrium prices. For instance, if a large portion of abatement technologies are related to electrification, potential feedback effects from electricity prices are taken into account. Third, it features sluggish technology adoption both across firms and over time. Fourth, we allow technologies to be overlapping, such that adoption of one technology reduces the potential of another. We illustrate the core features through simulations in a toy model.

Finally, we develop general calibration methods that allow us to fit the model to input-output data as well as technology data characteristics, such as the degree of adoption of technologies in a baseline year.

Keywords: Marginal abatement costs; CGE model; End-of-pipe abatement; Input-displacing abatement; Hybrid modelling.

JEL classification codes: C6; Q4; Q5.

*For useful comments and discussions, we thank Ulrik Beck, Peter Stephensen, and Jens Kirk from the modelling group DREAM and GreenREFORM. Furthermore, we thank our colleagues at University of Copenhagen Frikk Nesje, Peter Kjær Kruse Andersen, Peter Birch Sørensen, and Janek Eskildsen for valuable feedback and discussions.

†University of Copenhagen. Oster Farimagsgade 5, 1353 Copenhagen K, Denmark. E-mail: rasmus.kehlet.berg@econ.ku.dk.

‡University of Copenhagen. Oster Farimagsgade 5, 1353 Copenhagen K, Denmark. E-mail: jl@econ.ku.dk.

§University of Copenhagen. Oster Farimagsgade 5, 1353 Copenhagen K, Denmark. E-mail: aetn@econ.ku.dk

1 Introduction

In the presence of an externality like the one presented by global warming, conventional environmental economic theory suggests that two key parameters be established to ascertain the proper course of action: The damage and abatement costs of the externality. This information is sufficient to establish a tax or tradable permit system that can restore efficiency. Beyond this stylized setup, the marginal abatement costs remain essential to determine an efficient regulation: In the case of greenhouse gas emissions, various governmental bodies have set quantitative targets such as reaching net-zero emissions within some time horizon. With quantitative targets, estimating abatement costs is important to determine the necessary level of e.g. a carbon tax. In cases where imposing a carbon tax is politically infeasible, estimates of abatement costs are important for guiding the choice of where and how to impose alternative regulation.

This paper offers a new method for estimating abatement costs utilizing a model formulation where abatement technologies are represented as optimizing agents allowing utilization of a wide range of technological data hitherto not possible in applied economic models. The method solves a number of challenges present in conventional approaches. Most notably, it explicitly distinguishes between end-of-pipe and input-displacing technologies, showing that these are substitutes: Adopting end-of-pipe technologies reduces the effective emission content of inputs, making dirty input-displacing technologies cheaper relative to competing clean ones.

In the environmental economics literature, the estimation of abatement costs has largely focused on identification of three channels of abatement (Bovenberg et al. 2008): First, given a fixed mix of inputs, output reduction lowers emissions at the rate of the emission intensity. Second, given the scale of output, input substitution towards cleaner inputs lowers the emission intensity and thus emissions. Third, certain technologies target emissions of pollutants directly, achieving emission reductions without otherwise altering the input mix nor the scale of production. These technologies are end-of-pipe in character, akin to retrofitting a filter to an existing production process. The channel denoted input substitution captures a number of different types of substitution towards cleaner inputs. For instance, it captures if firms increase their energy efficiency by substituting capital for energy inputs, or a composition effect of relatively clean firms increasing their scale of output relative to more dirty ones. However, it also covers the case where a firm adopts a relatively clean technology to displace a relatively dirty one, e.g. a heat pump that produces heating services using electricity in place of a boiler using fossil fuels. Following Amundsen et al. (2018a), we refer to this type of input substitution as adoption of input-displacing technologies.

One approach to measure the costs and potential of abatement is to use computable general equilibrium (CGE) models that naturally incorporate general equilibrium effects, such as rebound and sectoral reallocation effects, that can be important to the assessment of environmental policies and their related costs (Grepperud and Rasmussen 2004; Gillingham et al. 2013). Furthermore, CGE models account for the output reduction channel of abatement and to some extent for the input substitution channel. In such models the key parameters that govern the potential for

abatement are the elasticities of substitution where identification is typically based on historical data.[1] This approach, however, does not utilize data on technological abatement options that are collected in so-called technology catalogues or organized directly as marginal abatement cost curves (for an example of the latter, see Enkvist et al. (2007)).

In recent years, there have been several attempts to integrate bottom-up information on technological abatement options into CGE models in order to estimate abatement costs (Rive 2010; Kiuila and Rutherford 2013a; Kiuila and Rutherford 2013b; Faehn and Isaksen 2016; Weitzel et al. 2019). A common feature of the proposed methods is to synthesize the information on an abatement technology into two parameters: The cost and reduction potential of adopting it. The technologies are then ordered by their unit cost of abating emissions and usually collected in a step-wise marginal abatement cost (MAC) curve. In this approach, the various technologies are stacked in ascending order with the least costly first, with the cost and reduction potential being exogenously determined by technical data.[2] An abatement technology is then simply adopted in the model once the price of emissions exceeds the cost of the technology.

While capturing the technological abatement options through an exogenous MAC curve is computationally convenient, this reduced-form approach has a number of drawbacks that we solve in this paper. We propose a method for modelling technological abatement options that essentially builds on a return to basics for economists: Each technology is represented by a profit maximizing firm that produces its productive services under perfect competition. As with any other firm in the economy, technology firms draw on inputs from other sectors and face installation costs of capital in their production. This representation of technologies as optimizing entities solves four problems associated with the conventional MAC curve approach.

First, the literature on MAC curves has either handled input-displacing technologies implicitly by CES elasticities or explicitly by including them in MAC curves alongside end-of-pipe technologies, thus treating the two categories of technologies identically. A novel feature of our framework is that end-of-pipe technologies and input-displacing technologies are modeled distinctively and explicitly: Input-displacing technologies are modeled as profit maximizing firms that produce a specific service that is used as inputs by other firms. In the case of a heat pump technology, the technology firm produces heating services. Together with other technologies that provide heating services, these technologies are combined in a nested production function to produce an aggregate of heating services which is then used as an intermediate input in production of other goods and services. End-of-pipe technologies on the other hand produce reductions in emissions and are thus placed outside the nested production tree. In this sense the input-displacing and end-of-pipe technologies act as substitutes: A firm that adopts a cost effective end-of-pipe technology essentially achieves a price reduction on dirty inputs as it circumvents the cost of emissions imposed by regulation.[3] Thus, adopting an end-of-pipe technology makes the input-displacing technologies that

[1] See for instance documentation for the large scale CGE models GTAP and MAKRO Hertel et al. (2003) and Kronborg et al. (2019).

[2] An important exception to this is Weitzel et al. (2019); we revisit this exception later in the introduction.

[3] We generally assume that some type of regulation makes emissions costly; otherwise, there would simply not

rely on cleaner inputs relatively more expensive. Treating all technologies as end-of-pipe, e.g. as in Faehn and Isaksen (2016), will thus tend to underestimate the costs of abatement as the end-of-pipe technology lowers the price of dirty inputs and the profitability of the competing input-displacing technology. Also, including input-displacing technologies as if they were end-of-pipe can lead to the wrong predictions of the use of inputs: Treating a heat pump as an end-of-pipe technology will result in a decrease in the prices on dirty inputs (e.g. oil) and thus generally predict an increase in the use of oil, whereas in reality the heat pump displaces oil and coal and rather requires an increase in the use of electricity. In the general equilibrium, where the price of e.g. electricity is endogenous, this distinction further affects the estimated costs of abatement.

Second, the related literature usually constructs the MAC curve outside the model and keeps the costs of technologies exogenous throughout simulations. In our representation of technology firms, production of abatement services requires the use of inputs and thus relies on the use of capital, fuels, labor etc. Capturing the link between abatement costs and general equilibrium prices is important in cases where prices are likely to change. Consider for instance the technological abatement options that come in the form electrification: In this case, adopting abatement technologies requires a large increase in the use of electricity, which in turn may drive up equilibrium prices; this underestimates the costs of adopting the abatement technology.

Third, we depart from the previous literature to capture a more realistic adoption rate of technologies. All firms in the model solve a dynamic profit maximization problem, taking into account the expected path of future prices. By including convex installation costs of investments, the optimal rate of adoption of input-displacing and end-of-pipe abatement technologies is gradual over time. Thus, if regulation of emissions tightens quickly, firms face large installation costs; ignoring this channel will thus tend to underestimate abatement costs. Furthermore, whereas the traditional MAC curve approach assumes a constant marginal cost across firms of any given abatement technology, we allow costs to be heterogeneous according to some distribution. This implies that the threshold for when firms consider an abatement technology to be profitable is also heterogeneous, giving rise to gradual adoption for the representative firm even if installation costs of capital are negligible. This gradual within-period adoption is an application of the smoothing methodology of Berg and Eskildsen (2021) and requires us to specify the distribution that these firm-specific costs follow.

Finally, we offer a method for handling technologies with overlapping potentials. When two or more technologies have overlapping potentials, they essentially compete to abate the same emissions; if a firm uses one technology, it renders the other competing technology redundant. In the traditional approach, overlapping potential is handled outside the model by adjusting the reduction potentials of individual technologies. For instance, if two technologies can provide the exact same emission reductions, i.e. they have 100% overlap of potentials, the technology that is most expensive, given a set of baseline prices, is simply removed from the model (reduction potential is zero). In our setup where the cost of adopting a technology depends on equilibrium

be any adoption of end-of-pipe abatement technologies.

prices, the ordering of which technology is cheapest is no longer fixed. Thus, instead of the *ad hoc* adjustment to reduction potentials, we explicitly let the two technologies compete in the same market for emission reductions. If two carbon capture and storage technologies are available, but only one can be used at each plant at any given time, they overlap, and adopting one must imply a lower degree of adoption of the other. Our modeling approach ensures that this is the case.

Several papers attempt to integrate technological end-of-pipe abatement options into CGE models (Weitzel et al. 2019; Kiuila and Rutherford 2013b; Rive 2010). Whereas most papers introduce an exogenous technological MAC curve, Weitzel et al. (2019) endogenize costs by including technology-specific input shares. Using a quadratic cost approximation, they integrate the technological MAC curve with endogenous costs using a mixed complementarity format. This implementation introduces a large number of complementarity-slack conditions.[4] We differ from this approach in two important ways: First, we include several other aspects such as distinguishing between input-displacing and end-of-pipe technologies, adjustment costs of adopting new technology, and overlap in technologies' potentials. Second, we provide a computationally efficient formulation of the end-of-pipe abatement module as a system of closed-form equations.

In terms of modeling input-displacing technologies, there is, to the best of our knowledge, no general methodology for integrating technological data in general equilibrium models. In this regard, the closest related literature is the body of papers that rely on techno-economic data to describe specific sectors in the general equilibrium (see e.g. Böhringer 1998; Schumacher and Sands 2007; Böhringer and Rutherford 2009). For instance, Schumacher and Sands (2007) focus on the German iron and steel production industry based on technological data on different production 'routes' that use different energy inputs. We show that under certain simplifying assumptions, we replicate the modeling approach in Schumacher and Sands (2007).[5]

Our proposed method for integrating bottom-up technology data in CGE models is practically appealing as well: For existing CGE models that build on nested production trees, we include input-displacing technologies by systematically extending the existing tree structure. To include end-of-pipe technologies, all firms who emit must choose how much end-of-pipe abatement to undertake; we show that optimal adoption only requires that prices of dirty inputs are updated according to a simple formula. This makes implementation in existing models straightforward. Furthermore, we provide methods for calibrating these extended models to input-output tables. The methodology presented can be applied to a range of areas and technology data including conventional expert-based marginal abatement cost (MAC) curves as in Kiuila and Rutherford (2013b).

The detailed information on technologies comes from technology catalogues that for each technology include data on e.g. the associated average unit cost in some base year, required inputs, and emission reduction potential for end-of-pipe technologies. Examples of this type of data include

[4] This has important implications for the solution time: Weitzel et al. (2019) report that the solution time increases by 14% and 168% when solving the model in years 2025 and 2030 respectively.

[5] Appendix E elaborates on this.

a large portion of energy demand and supply technology data as published by the Energy Technology Systems Analysis Program.[6] The parameters introduced by our setup are identified by the data in technology catalogues such that no formal estimation is necessary. Essentially, technology catalogues provide a different identification strategy for parameters central to CGE models, which allows us to rely on the best available information on specific future technologies rather than on historical data on technologies in general.

The rest of the paper is organized as follows: Section 2 defines and explains input-displacing and end-of-pipe technologies. Section 3 presents the general model framework. Section 4 outlines the method for calibrating a CGE model with our methodological extensions to technology data and input-output data. This section further discusses shortcomings of our approach and some avenues for future research. Section 5 provides simulations in a toy model to illustrate core features of the framework. Finally, section 6 concludes.

2 Technologies

Firms have a wide range of available abatement technologies that can be categorized as either end-of-pipe or input-displacing. An end-of-pipe technology, as the name suggests, is a technology that does not alter the input structure nor the outputs of the production process, but instead affects the emission intensity of production. An important example of such a technology is carbon capture and storage (CCS), whereby firms capture the carbon, e.g. from combustion, before it is released into the atmosphere as carbon dioxide. The direct effect of installing the technology is thus entirely measured as lower emissions at some cost. Naturally, however, adopting the technology may still induce indirect changes in the optimal use of inputs and the scale of production. This occurs when a firm adopts an end-of-pipe technology, thereby lowering the relative cost of emission intensive inputs by effectively making them less emission intensive.

An input-displacing technology, in contrast, is a technology which directly affects the input structure of the production process in a particular way. Consider for instance a steel producer who relies on high temperature industrial facilities to melt steel. These high temperatures, sometimes above 1500 degrees Celsius, can be reached by a variety of possible technologies that use fossil fuels such as coal, gas, coke, or electricity in different amounts. A producer of steel can only realistically substitute from one energy source to another in this context by installing a new technology. Such input-displacing technologies do not change the output of production, but simply provide a productive service (e.g. melting) that can be thought of as an intermediate good in the production of the particular output (e.g. steel).

Our modeling framework reflects the distinct features of the two categories of technologies. The following presents each of them in turn.

[6] The class of data that we refer to here includes all data retrieved from www.iea-etsap.org/ where the convention is to characterize each technology by capacity, capital cost, fixed operating and maintenance costs, variable operating and maintenance costs, as well as some measure of energy-efficiency.

2.1 Stylized technology: End-of-pipe

Consider two stylized CCS technologies applicable in the production of cement summarized in table 1. Installing and running them requires capital as well as the energy inputs electricity and oil respectively.

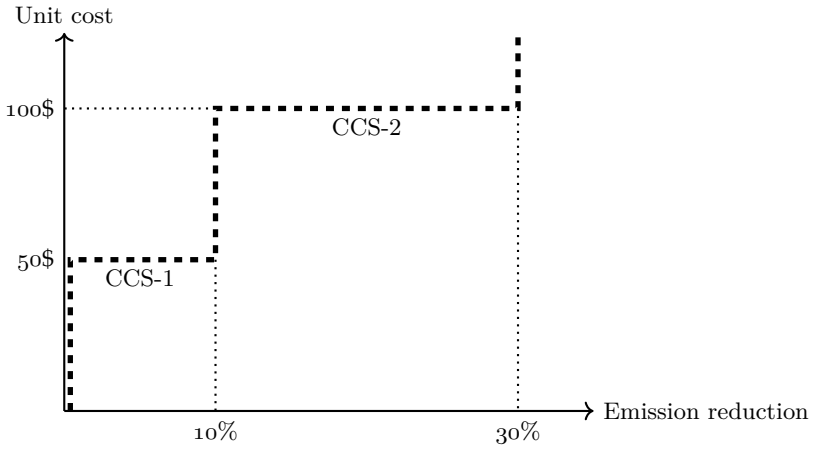
Table 1: Two stylized end-of-pipe technologies

Technology	Inputs	Unit cost	Emission reduction potential
CCS-1	Capital, oil	50 $	10 %
CCS-2	Capital, electricity	100 $	20 %

Notes: The unit cost is measured for a given set of input prices in some base year. The emission reduction potential measures the fraction of the industry's emissions that will be abated if the corresponding technology is fully adopted.

Traditionally, the integration of end-of-pipe technologies in CGE models relies on technological marginal abatement cost (MAC) curves constructed from data of the form in table 1 (see e.g. Rive 2010; Faehn and Isaksen 2016). A technological MAC curve is a step-wise function in the (emission reduction, unit cost)-space. For this very simple example with just two technologies, the MAC curve would look as depicted in figure 2.1. Intuitively, the curve shows that if the cost of emissions rises above 50 $, technology CCS-1 is phased in.

Figure 2.1: Stylized marginal abatement cost curve

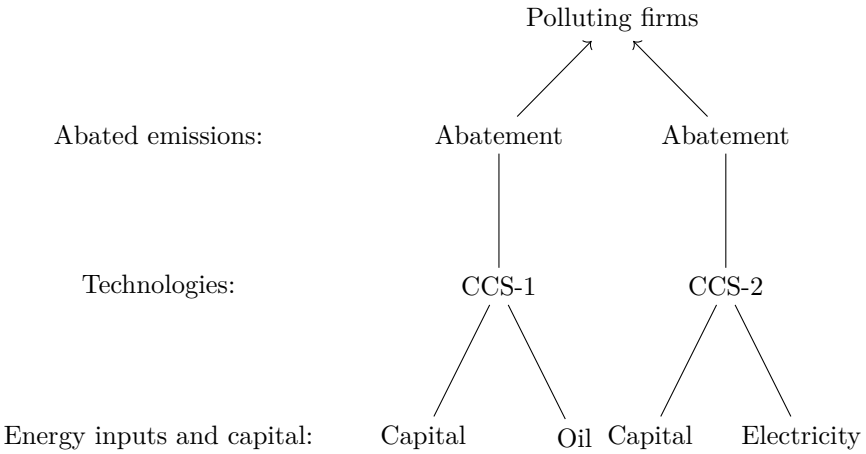


Incorporating MAC curves like this in models can be done either through approximating it using a smooth functional form or by incorporating the step-wise curve directly (Böhringer and Rutherford 2009). An example of the former is Faehn and Isaksen (2016) who incorporate a smoothed version of a technical marginal abatement cost curve by adjusting two otherwise exogenous parameters in the model: The total factor productivity of the representative firm and the emission intensity of its output. The price of emissions determines the optimal point on the MAC curve and hence a reduction of the emission intensity, which implies a cost, modeled as a reduction in total factor productivity. In this conventional approach, the first technology in figure 2.1 will be phased in smoothly when the price on carbon approaches 50 $ regardless of the state of the

economy (e.g. prices on oil, electricity and the degree of adoption of other technologies).

Rather than regarding the MAC curve as a fixed pre-computed object, our framework takes a step back and represents each step as a firm that uses inputs to produce end-of-pipe abatement services. This effectively makes the MAC curve of our model an endogenous object determined in equilibrium. Polluting firms abate emissions by purchasing abatement services from end-of-pipe abatement firms. For the two technologies of table 1, this gives rise to the nesting structure shown in figure 2.2.

Figure 2.2: Stylized nesting structure: End-of-pipe



The nesting structure makes it explicit that end-of-pipe technologies combine inputs to produce abated emissions. The costs of technologies, and the related abatement services, are tied to equilibrium prices through the bottom-nest. In this nesting structure, technologies rely on capital implying that the cost depends on the rental rate of capital. Relatedly, imposing standard installation costs in the use of capital for abatement technologies is a straightforward way to model sluggish adoption of these technologies.

The representation of the two technologies in figures 2.1 and 2.2 implicitly assumes that the two technologies are independent, i.e. that the reduction potential of technology CCS-1 does not depend on how much CCS-2 is phased in. In the conventional setup where the MAC curve is completely exogenous this is not a problem: To account for potential overlap of the technologies' reduction potentials, one can simply adjust the width of the plateaus before implementing the MAC curve in the model. When the costs of the technologies are fixed, the CCS-1 technology will always be adopted before CCS-2. Thus, any overlap is handled by appropriately scaling down the reduction potential of CCS-2. When the costs of technologies are endogenized, the order by which technologies are adopted is no longer fixed: Since technologies rely on different inputs, each technology's costs react differently to changes in input prices. With endogenous ordering, overlap in reduction potentials must be handled in other ways than simply scaling selected reduction potentials down. Section 3 explains how we model this overlap.

2.2 Stylized technology: Input-displacing

Input-displacing technologies alter the input structure of a production process. Take as an example an electrical heat pump which provides the energy service 'heating' in some sector, say, the paper and pulp industry, summarized in table 2.

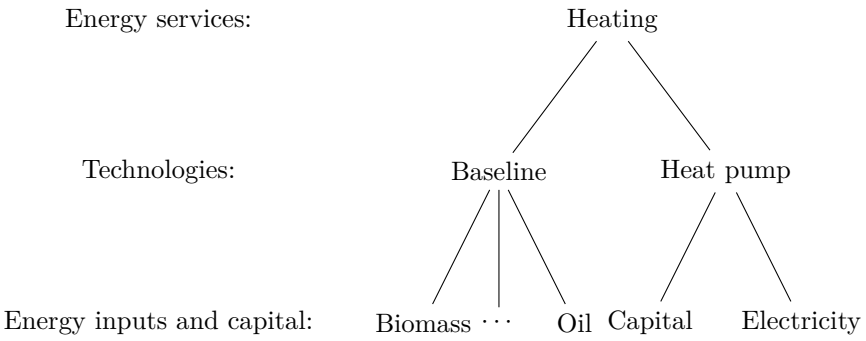
Table 2: A stylized input-displacing technology

Technology	Inputs	Unit cost	Energy demand coverage potential
Heat pump	Capital, electricity	25 $	10 %

Notes: The unit cost is measured for a given set of input prices in some base year. The energy demand coverage potential measures the fraction of the relevant energy service, e.g. heating, that the technology covers if fully adopted.

As with end-of-pipe technologies, we represent this technology as a representative firm which maximizes profits by combining its inputs and selling an energy service; in this case heating. We acknowledge that the technology data available may vary a lot. In particular, some technology catalogues are not exhaustive, meaning that the technologies from the catalogue do not cover all ways of producing an energy service (e.g. heating). To account for the case of non-exhaustive technology catalogues, we introduce an otherwise unlabeled 'baseline' technology that competes with the heat pump to supply heating. This essentially ensures that all catalogues, in combination with the baseline technology, become exhaustive. The resulting nesting tree is depicted in figure 2.3.

Figure 2.3: Stylized nesting structure: Input-displacing



In this setup, the heat pump competes against the baseline technology in a nest with a high degree of substitutability since they supply the same service. Traditionally, when modeling input substitution possibilities, the model does not present technologies explicitly as representative firms. Instead, one only has a function that aggregates different fuels into the energy service (e.g. heating) directly. This would correspond to simply leaving out the heat pump in the example above. All substitution between energy goods is then governed by the parameters in the aggregator function for the baseline technology, e.g. the elasticity of substitution in the case of CES. By explicitly modeling technologies, we replace part of the black-box substitution implied by the elasticity of substitution and replace it by competition between technologies that supply the same energy

service.

2.3 Identification of technologies

We identify parameters on technology firms from 'technology catalogues', which are detailed expert-based assessments of current and potential future technologies. Because our model represents technologies explicitly as their own node in a nested tree structure, the construction of the tree is dictated by the relevant technology catalogue.

A model built on specific technologies from a catalogue has two challenges. First, because technology catalogues are rarely exhaustive, it does not necessarily identify the status quo technologies from which an industry produces its goods. Second, the model should be able to reproduce national accounts data on the inputs and outputs produced by the industry. By modeling baseline technologies, we solve both challenges: They represent the status quo technologies that the technologies from the catalogue compete against. As baseline firms capture residual activity not covered by catalogues, they are residually identified from industry-level input-output data from national accounts.

The granularity with which these baseline technologies can be specified depends on the degree of detail in the input-output data used in the application. We assume that data on final goods production, total emissions, energy service use, input use for all energy services, as well as the industry's total input use is known from national accounts. Appendix D outlines calibration cases where less detailed data is available.

3 Model

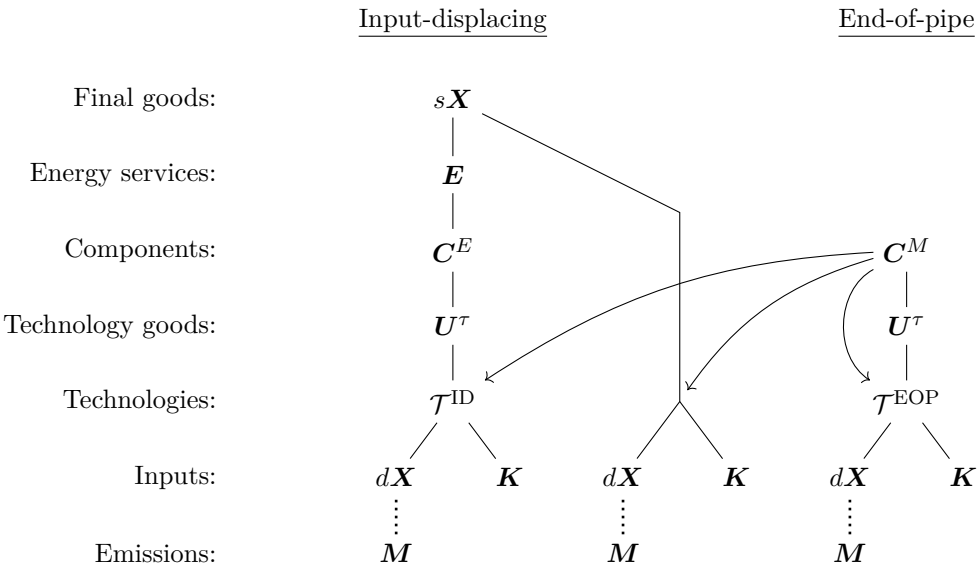
Consider a sector $s \in S$ that produces a vector (boldface notation) of non-capital final goods $s\boldsymbol{X}$, using a combination of non-capital goods $d\boldsymbol{X}$ and capital goods $\boldsymbol{K}$. The vector of non-capital goods $\boldsymbol{X}$ contains both final goods and inputs, and we use the notation $d\boldsymbol{X}$ to denote demand for these goods and $s\boldsymbol{X}$ to denote the supply, when such distinction is necessary. As the same types of inputs are used in various firms, we use $d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)$ to specify inputs used by technology firm τ. For simplicity, we drop sectoral and time subscripts whenever such distinction is not necessary.[7]

The nesting tree of this model is more general than the stylized ones in section 2, but the overall structure is the same: There are two nested trees, a tree with input-displacing technology firms at the bottom that produce energy services, and a tree to the right with end-of-pipe technology firms at the bottom that produce abatement services. The input-displacing tree further includes a subtree that is unrelated to energy services (middle part of figure) to illustrate that the model may include an arbitrarily nested production function beyond the energy service part covered by technology catalogues. The nesting structure includes four new elements compared to the stylized cases in figures 2.2-2.3: Emissions $\boldsymbol{M}$, technology goods $\boldsymbol{U}^{\tau}$, components $\boldsymbol{C}$, and final goods $s\boldsymbol{X}$.

[7] Appendix A summarizes all notation.

Final goods producers combine intermediate goods $\boldsymbol{E}$, which we refer to as energy services, as well as non-capital inputs $d\boldsymbol{X}$ and capital goods $\boldsymbol{K}$ in the production of final goods $s\boldsymbol{X}$. Profit-maximizing firms produce energy services E from *components*, $\boldsymbol{C}^E$. Component firms produce the C from *technology goods*, $\boldsymbol{U}^C$. Technology goods are outputs of profit-maximizing technology firms $\tau \in \mathcal{T}^{\text{ID}}$ who also use capital and non-capital inputs. The complete nesting structure is illustrated in figure 3.1. Although we interpret each nest as being populated by a profit-maximizing firm, an equivalent interpretation is that the entire tree represents a single representative firm where each level in the tree indicates cost minimization in that part of the firm.

Figure 3.1: The organization of subsectors



Note: The arrows indicate flows of end-of-pipe abatement from the end-of-pipe component firms to other firms in the industry that use energy inputs and thus generate emissions. All firms who generate emissions are allowed to purchase end-of-pipe abatement services if such possibilities are technically available for their sector. Baseline technologies are not shown.

A technology firm in the end-of-pipe tree $\tau \in \mathcal{T}^{\text{EOP}}$ operates in a way similar to input-displacing technology firms. They combine non-capital and capital inputs to produce technology goods $\boldsymbol{U}^\tau$. In this tree, however, technology goods are interpreted as abatement of emissions rather than intermediates. The technology goods are combined into components $\boldsymbol{C}^M$ which any firm in the sector can purchase in order to undertake end-of-pipe abatement, as indicated by the arrows in figure 3.1.

In line with the graphical representation of the nesting tree, we assume that the use of non-capital goods is inherently tied to emissions in intensities ϕ_X^M for each input.[8] Note also that while not shown in the nesting trees in figure 3.1, a separate sector produces capital inputs using a combination of non-capital and capital inputs.

To provide a better understanding of the rather complicated nesting tree, an example is warranted. The nesting level 'technologies' represents a technology as we observe them in data; thus,

[8]The assumption is that CO_2 emissions are generated directly by burning fossil fuels. Including emissions related to the production process or the final product itself can be implemented by scaling up the emission factors of the inputs used in that particular sector. Section 4 elaborates on this.

we can think of $\tau \in \mathcal{T}^{ID}$ e.g. as a specific type of heat pump. The heat pump draws on inputs (e.g. electricity and capital) to produce heating. We measure the output of τ as the *total* output of κ_τ GJs produced by the specific heat pump. Next, we can think of the nesting level 'technology goods' as being *specialized* output from the related technology. For the heat pump that produces κ_τ GJs it might be the case that 50% of this output is low temperature heating, while the rest is medium temperature heating. In this case technology τ produces two technology goods: $U^\tau_{low} + U^\tau_{medium} = \kappa_\tau$. Next, the nesting level 'components' can be thought of as markets for the specialized outputs that technology firms produce. For instance, if there are two technologies, τ_1, τ_2, that produce the specialized output 'medium temperature heating', the relevant component, C_{medium}, would be some linear combination of $(U^{\tau_1}_{medium}, U^{\tau_2}_{medium})$.[9]

While technologies' total output κ_τ, technology goods U, components C, and energy services E are subtly different, we essentially still measure all of them in the same units, e.g. GJs of heating, as in the example above. The exact nesting structure of the tree depends on the available technology data.[10] Furthermore, to be able to interpret technology goods and components as essentially the same as the energy service they contribute to, we impose an assumption, that we refer to as 'scale preservation', for all nests between technologies and energy services. A function is *scale preserving* if the sum of inputs equals the sum of outputs in optimum.

The assumption of scale preservation is furthermore intuitively appealing in our setting. Consider, for instance, two input-displacing technologies: An electric heat pump and an oil-based boiler. Suppose they are able to provide the same energy service, e.g. heating. If a firm in the economy switches from using one of the technologies to the other one, the number of final goods resulting from that should, *ceteris paribus*, not change. In other words, the productive value of the same GJ of heating does not depend on whether or not it comes from a heat pump or a boiler.

To navigate the tree structure we use a superscript to denote a node's direct ancestor/parent. For instance, C^E is a component with parent node E. When more general connections are needed, we let $\mathcal{D}_Z(Y)$ denote the subset of nodes of type Z that are descendants or ancestors of the node Y. For instance, $\mathcal{D}_U(E)$ denotes the set of all technology goods U that are descendants of the node E.[11]

Appendix B states our definition of *scale-preserving functions* (definition B.3) formally, and characteristics of this type of functions are outlined in Lemma B.1. Beyond the definition of scale-preserving functions, appendix B also includes formal definitions of *neoclassical-type functions* (B.1) and of *standard optimization problems* (B.2) that we will apply throughout.

[9] As we show later, we always include a baseline technology good for the technology firms to compete with for the component market. For this particular example, it means that C_{medium} is a linear combination of $(U_o, U^{\tau_1}_{medium}, U^{\tau_2}_{medium})$ where U_o is a technology good produced by the baseline technology.

[10] Appendix C provides an example of how to construct the nesting tree from the available technology data. We provide a program that creates the nesting tree from properly arranged technology data online https://github.com/ChampionApe/GPM_v05.

[11] We use a combination of the two notations as the composite function; for instance, the notation $\mathcal{D}_{U^\tau}(E)$ denotes the set of all technology goods that are descendants of the node E, but also the child node to the specific technology τ.

3.1 Emissions and application of end-of-pipe abatement

All firms that emit consider the trade-off between paying taxes on emissions and paying for end-of-pipe abatement to avoid said emissions. We think of emissions as being inherently tied to inputs in exogenous shares $\boldsymbol{\phi}_X$ for each $X \in \boldsymbol{X}$. Thus, emissions prior to end-of-pipe abatement in a sector ($\boldsymbol{M}_{\text{o}}$) are mechanically given by

$$\boldsymbol{M}_{\text{o}} = \sum_{X \in \boldsymbol{X}} \phi_X^M dX. \tag{1}$$

Since end-of-pipe abatement of a pollutant can be applied regardless of its source (i.e. which input caused the emissions), any firm that emits also applies end-of-pipe abatement if such technologies are available in its sector and are profitable. Technically, a firm does this by purchasing abatement services in the form of end-of-pipe components $\boldsymbol{C}^M$ from the end-of-pipe abatement nesting tree, as indicated by the arrows in figure 3.1. These abatement services are measured in units of emissions.

Consider any firm i that uses non-capital inputs resulting in emissions according to (1). The abatement (A_M) and all costs related to emissions and purchase of abatement (EC_M) of emission type M are defined by

$$\begin{aligned} A_M(i) &= \sum_{C \in \boldsymbol{C}^M} C(i) \\ EC_M(i) &= p_M \left[M_{\text{o}}(i) - A_M(i)\right] + \sum_{C \in \boldsymbol{C}^M} \left[p_C C(i) + M_{\text{o}}(i)\theta_C \; \tilde{G}\left(\frac{C(i)}{M_{\text{o}}(i)\theta_C}\right)\right], \end{aligned} \tag{2}$$

where p_M is the tax on emission type M, p_C is the equilibrium price of component C, θ_C is the share of total pre-abatement emissions that component C could abate if it was fully phased in, and $\tilde{G}(\cdot)$ measures potentially nonlinear costs of adopting the abatement equipment. The objective of firm i is to minimize costs related to emissions given the use of inputs through its purchase of abatement services.[12] We note that, from the perspective of the individual firm i, all abatement services (C) are seemingly independently offered at the cost p_C. However, as we show later, the costs of two different abatement services may be linked through technology firms that can potentially produce more than one. For example, if a single technology τ provides abatement of C^{CO_2} and C^{NO_x} , this will be reflected in the prices $p_{C^{\text{CO}_2}}, p_{C^{\text{NO}_\text{x}}}$.

The cost function $\tilde{G}(\cdot)$ is essential to the optimal adoption of abatement equipment in this formulation. Formally, we restrict the function to be monotonically increasing, continuously differentiable, and convex. Intuitively, $\tilde{G}$ can be interpreted as capturing two reasons for gradual adoption of new technologies: First, we can interpret it as a measure of heterogeneity in the costs of adoption across firms. Assume for instance that a firm j faces the decision of whether to apply an end-of-pipe technology or not. If the cost exceeds the price of emissions, it is optimal for the

[12] Formally, they minimize $\sum_{M \in \boldsymbol{M}} EC_M$ by choosing abatement (A_M) and abatement services (C) keeping the level of inputs (dX) constant. We note that the firm i later optimizes over inputs as well; however, with the assumptions made on production technologies throughout the model, this two-step optimization procedure is equivalent to the case where the firm chooses the optimal purchase of abatement and use of inputs simultaneously.

firm to adopt this technology. If we think of firm i as representing a distribution of individual firms j, the cost function $\tilde{G}$ can be interpreted as capturing heterogeneity in the costs of adopting the technology. Second, we can think of the adoption of an end-of-pipe technology as a continuous rather than a binary choice. For instance, it may be the case that an end-of-pipe technology can be applied in a number of different production facilities owned by firm i, but at varying efficiency. In this case, the function $\tilde{G}$ captures the variation in costs of applying the technology. Furthermore, the function $\tilde{G}$ serves an important technical purpose, as it allows a gradual and smoothed phasing-in of technologies. In settings where information on firm heterogeneity in costs is available this can be used to choose the functional form of $\tilde{G}$.

Minimizing emission costs (2), the first order condition for optimal adoption of end-of-pipe abatement for firm i is given by:

$$p_M = p_C + \tilde{G}'\left(\frac{C(i)}{M_0(i)\theta_C}\right),$$

and the corresponding optimal use of abatement component $C(i)$ is defined as

$$C(i) = M_0(i)\theta_C G(p_M - p_C), \qquad G(x) \equiv \left(\tilde{G}'\right)^{-1}(x).$$

The function G measures the degree of adoption of a technology as a function of the profitability of the technology $(p_M - p_C)$. We note that it can be convenient to specify the adoption function G instead of the cost function $\tilde{G}$. In particular, we emphasize that the assumption that $\tilde{G}$ is a monotonically increasing, differentiable, and convex cost function is consistent with specifying G as **any** cumulative density function with continuous support.[13] Furthermore, we note that it is convenient to assume that G is the cumulative density of a mean-zero distribution, as this ensures that the average price of adopting a technology fully is simply given by the price p_C.

Using this, abatement and the costs related to emissions and abatement when technology goods are applied optimally are given by:[14]

$$A_M(i) = M_0(i) \sum_{C \in \mathcal{D}_C(A_M)} c_M, \qquad \text{where } c_M \equiv \theta_C G(p_M - p_C) \tag{3a}$$

$$EC_M(i) = M_0(i)\left[p_M\left(1 - \sum_{C \in \mathcal{D}_C(A_M)} c_M\right) + \sum_{C \in \mathcal{D}_C(A_M)} \theta_C\left(p_C G(p_M - p_C) + \int^{p_M - p_C} \chi dG(\chi)\right)\right].$$

The first term in the square brackets measures the costs of emissions after abatement, whereas the second term measures the costs of abatement components. Using a cumulative density function to specify the adoption rate G, we can rewrite the last term in the abatement cost function using the

[13] If G is a cumulative density function it is indeed invertible (on its support), the inverted function is continuous, integratable, positive, and monotonically increasing; thus $\tilde{G}$ will indeed be a differentiable and convex function.

[14] Note that to arrive at the costs of applying abatement components, $\tilde{G}$, we integrate over the set of prices $(\chi \in (-\infty, p_M])$ in the marginal costs, instead of using the optimal $C(i)$ in the cost function $\tilde{G}$ directly. This is simply as we prefer to specify the costs in terms of the function G instead of $\tilde{G}$.

truncated expected value:[15]

$$EC_M(i) = M_0(i) \underbrace{\left[p_M \left(1 - \sum_{C \in \mathcal{D}_C(A_M)} c_M \right) + \sum_{C \in \mathcal{D}_C(A_M)} c_M \left(p_C + \mathbb{E}_G \left[\chi | \chi < p_M - p_C \right] \right) \right]}_{\equiv \hat{p}_M}. \quad \text{(3b)}$$

We note that, ultimately, the share of emissions that are abated using end-of-pipe technologies depends on the prices $p_C \in \boldsymbol{p}_C$. Recall from the nesting tree in figure 3.1 that each abatement component ultimately is produced using abatement technologies; thus, the relevant prices of abatement services are ultimately dictated by the costs of technology firms.[16]

Four important features of the end-of-pipe equations (3a) and (3b) emerge: First, total costs related to emissions and abatement are proportional to pre-abatement emissions, allowing us to introduce a price of emissions *post* end-of-pipe abatement $\hat{p}_M$. Second, $\hat{p}_M \leq p_M$; thus, abatement equipment is only used when it is profitable to do so. Third, in the limit, when prices of emissions are sufficiently high, abatement converges on its potential, in the sense that all technical abatement options are exhausted:

$$\lim_{p_M \to \infty} A_M(i) = M_0(i) \sum_{C \in \mathcal{D}_C(A_M)} \theta_C.$$

Fourth, abatement costs (the latter part of the parenthesis in $EC_M(i)$) converge to their average cost under full utilization, where $G(\cdot)$ evaluates to 1, when the price of emissions becomes sufficiently high:

$$\lim_{p_M \to \infty} \sum_{C \in \mathcal{D}_C(A_M)} c_M \left(p_C + \mathbb{E}_G \left[\chi | \chi < p_M - p_C \right] \right) = \sum_{C \in \mathcal{D}_C(A_M)} \theta_C p_C.$$

The terms related to cost heterogeneity vanish in the limit when G represents a mean-zero distribution. Finally, note that taking emission taxes and end-of-pipe abatement into account, the effective prices of non-capital inputs $\boldsymbol{X}$ are given by

$$\hat{p}_X = p_X + \sum_{M \in \boldsymbol{M}} \phi_X^M \hat{p}_M \quad (4)$$

where $\hat{p}_M$ is as defined from (3b). The implication of this is that we can capture firms optimal adoption of abatement equipment simply by replacing pure inputs prices p_X with adjusted ones $\hat{p}_X$.

[15] We use that in general the conditional expectation is given by $\mathbb{E}[x|X < k] = \int^k x g_{x|X<k}(x)dx$ where $g_{x|X<k}$ is the truncated density. The truncated density in turn is simply given by $g_{x|X<k} = g(x)/G(k)$. Using this we have:

$$\int^k x \, dG(x) = \int^k x \, g_{x|X<k} G(k) dx = G(k) \int^k x \, g_{x|X<k} dx = G(k) \mathbb{E}_G \left[x | X < k \right].$$

[16] We note that it is straightforward to include a simple mark-up on the final abatement services, to reflect that technology firms may have some market power.

3.2 End-of-pipe abatement firms

The end-of-pipe abatement sub-sector consists of two layers: The end-of-pipe technology firms (τ) and the end-of-pipe component firms (C). The end-of-pipe technology firms are modelled based on data from technology catalogues. The input structure, costs, and potential for abatement are all used to identify the way technology firms produce technology goods (U^τ) that essentially represent emissions abatement. A technology firm may produce multiple technology goods for two reasons: First, the end-of-pipe technology may be able to abate more than one type of emissions. In this case, one set of technology goods may represent abatement of NO_x emissions and another abatement of CO_2 emissions. Second, the multiple output structure is used to model competition between technologies that have overlapping potential for abatement. Consider, for instance, the case of two competing CCS technologies: Assume that 50% of large industrial plants can choose between CCS-1 and CCS-2, while the other half can only apply technology CCS-1. We model this by letting CCS-1 produce two technology goods: One that competes directly with technology CCS-2 and one that does not. Whereas technology firms represent something tangible and intuitive that we can observe in technology catalogues, the component firms are constructed to capture how overlapping technologies compete.

End-of-pipe technology firms

The technology firm $\tau \in \mathcal{T}^{\text{EOP}}$ uses non-capital and capital inputs to produce a composite technology good κ_τ using a neoclassical-type function F^{in}_τ, where in refers to *inputs*, as in definition B.1. This composite good is used to produce a vector of technology goods using a scale-preserving function F^o_τ, where o refers to *outputs*, as defined in definition B.3. This production structure allows us to control the types of inputs used in production of a technology (from specifying F^{in}), but also to allow for the technology to abate multiple types of emissions (from specifying F^o).

The end-of-pipe technology firm maximizes profits under perfect competition taking both input and output prices as given. Note that, as any other firm that relies on inputs dX, the end-of-pipe abatement firm also emit pollutants, and thus purchases end-of-pipe abatement services itself. An example could be a direct CO_2 capture technology in the cement sector requiring burning natural gas (Plaza et al. 2020). As shown in section 3.1, however, incorporating the optimal choice of these abatement services in any firm's optimization problem simply comes down to updating input prices p_X to $\hat{p}_X$ following the expression in (3b).

Given the optimal choice of end-of-pipe abatement, the technology firm τ solves

$$\max_{\boldsymbol{U}^\tau, d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)} \boldsymbol{p}_{U^\tau} \cdot \boldsymbol{U}^\tau - \hat{\boldsymbol{p}}_X \cdot d\boldsymbol{X}(\tau) - \boldsymbol{p}_K(\tau) \cdot \boldsymbol{K}(\tau), \tag{5}$$
$$\text{s.t. } \kappa_\tau = F^o_\tau\left(\boldsymbol{U}^\tau\right) \quad \text{and } \kappa_\tau = F^{in}_\tau\left(d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\right),$$

where we disregard how capital and non-capital goods are nested. Note that with technology specific adjustment costs, the capital costs are specific to the technology type, τ. The problem can be split into a two-step optimization problem: First, maximize profits given a shadow-cost variable

(p_τ)

$$\max_{\boldsymbol{U^\tau}} \boldsymbol{p}_{U^\tau} \cdot \boldsymbol{U}^\tau - p_\tau F_\tau^o\left(\boldsymbol{U}^\tau\right), \tag{6}$$

and second, minimize costs given scale (κ_τ):

$$\min_{d\boldsymbol{X}(\tau),K(\tau)} \hat{\boldsymbol{p}}_X \cdot d\boldsymbol{X}(\tau) + \boldsymbol{p}_K(\tau) \cdot \boldsymbol{K}(\tau), \qquad \text{s.t. } \kappa_\tau = F_\tau^{in}\left(d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\right).$$

where the scale κ_τ is given from the first step. Note that as both production technologies F_τ^o, F_τ^{in} are linearly homogenous, the optimization problem can be rewritten in ratios of the composite good κ_τ. We note that the first problem is a standard optimization problem (B.2) with parameters $(\boldsymbol{p}_{U^\tau}, p_\tau)$; thus, by construction, the solution to this problem can be represented by a scale-preserving aggregator $\mathcal{P}^\tau$ as defined in definition B.3. The scale-preserving aggregator is a collection of share functions $\mathcal{P}_U^\tau$ for each U^τ that return the optimal shares U^τ/κ_τ as a function of relevant prices. The optimal output split and the shadow-cost of the technology satisfies

$$U^\tau = \mathcal{P}_U^\tau(p_\tau, \boldsymbol{p}_{U^\tau})\kappa_\tau, \qquad U^\tau \in \boldsymbol{U}^\tau \tag{7}$$

$$p_\tau = \sum_{U \in \boldsymbol{U}^\tau} \mathcal{P}_U^\tau\left(p_\tau, \boldsymbol{p}_{U^\tau}\right) p_U. \tag{8}$$

The scale-preserving aggregator for the output split has the feature that, given the scale of production κ_τ, an increase in the price p_U increases the optimal share of output U^τ/κ_τ. In the case where a technology abates two types of emissions, this allows firms to focus on adoption of the technology in areas where emissions are most costly. Furthermore, the choice of aggregator $\mathcal{P}^\tau$ allows us to capture different types of technologies: If an end-of-pipe technology resembles a filter that reduces e.g. SO_2 and CO_2 emissions in fixed shares, the aggregator should resemble a Leontief-like function that yields fixed output-shares. On the other hand, it might be the case that the technology has potential to reduce both SO_2 and CO_2 emissions, because it can be applied in different facilities in an industry, such that the potential for CO_2 emission reductions are in principle relatively detached from its potential for SO_2 reductions. In this case, the aggregator should resemble a perfect substitutes-like function with highly price-responsive output-shares.

The optimal use of inputs is defined by the first order conditions

$$\frac{\partial F_\tau^{in}}{\partial Z} = \frac{\hat{p}_Z(\tau)}{p_\tau}, \qquad Z \in \{d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\}.$$

Given our assumption on the production technology F_τ^{in}, an inversion with respect to the input $Z \in \{d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\}$ is unique. Thus, we can define the optimal demand for inputs from

$$Z = H_z^\tau\left(\frac{\hat{p}_Z(\tau)}{p_\tau}\right)\kappa_\tau, \qquad Z \in \{d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\}, \tag{9}$$

where we use H_z^τ as a shorthand for the inversion of the partial derivative above. From the cost

minimization problem, we furthermore recover the price index

$$p_\tau = \sum_{Z \in \{d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\}} H_z^\tau \left(\frac{\hat{p}_Z(\tau)}{p_\tau} \right) \hat{p}_Z. \tag{10}$$

End-of-pipe component firms

The end-of-pipe component firms serve the purpose of modelling competition between overlapping technologies. Without any overlap in technologies' potential, there would be exactly one component for each technology good, and we would simply have $C = U$ and $p_C = p_U$.

The end-of-pipe component sub-sector produces abatement equipment $C \in \boldsymbol{C}$ by combining relevant technology goods with a scale-preserving production technology F_C^{in} as defined in definition B.3. For each component and related technology goods the firm then solves:

$$\max_{\boldsymbol{U}^C} p_C F_C(\boldsymbol{U}^C) - \boldsymbol{U}^C \cdot \boldsymbol{p}_U.$$

Note that this constitutes a standard optimization problem (definition B.2) with parameters $(p_C, \boldsymbol{p}_U)$. The optimal input shares U^C/C equal by construction a scale-preserving aggregator $\mathcal{P}^C$ as defined in definition B.3:

$$p_C = \sum_{U \in \boldsymbol{U}^C} \mathcal{P}_U^C \left(p_C, \boldsymbol{p}_U\right) p_U \tag{11a}$$

$$U = \mathcal{P}_U^C \left(p_C, \boldsymbol{p}_U\right) C, \qquad \forall U \in \boldsymbol{U}^C. \tag{11b}$$

The choice of aggregator $\mathcal{P}^C$ specifies how firms choose between adoption of competing end-of-pipe technologies. If two technologies essentially provide the same service, the aggregator should resemble that of a perfect substitute-like function; indeed, this is equivalent to the traditional assumption used in more technical bottom-up models (e.g. energy system models that rely on the TIMES framework as in Loulou et al. (2016)).

For convenience, appendix F.1 collects the complete set of equations and endogenous variables for the end-of-pipe abatement tree.

3.3 Input-displacing firms

The input-displacing sub-sector similarly consists of two layers: The technology firms and the component firms. In many respects, the roles of the two are similar to those of the end-of-pipe sub-sector. The technology firms are based on data from technology catalogues, and the component firms are included to capture competition of overlapping technologies. There are a couple of notable differences, however, compared to the end-of-pipe sub-sector: First, while the end-of-pipe technologies contribute to emissions abatement, the input-displacing technologies provide different energy services which are intermediates in production. Second, if a technology in the end-of-pipe case is too costly to apply, the relevant emissions are simply not abated. In the input-displacing

case, however, there is a *baseline* production of energy services even when the input-displacing technologies outlined in the catalogue are not profitable; there is some production that technologies effectively displace. To capture this, we add the notion of baseline technology firms.

Input-displacing technology firms

The technology $\tau \in \mathcal{T}^{ID}$ uses non-capital and capital inputs to produce a 'composite good' κ_τ using a neoclassical-type function F^{in}_τ. The composite good is used to produce technology goods $\boldsymbol{U}^\tau$ using a scale-preserving function F^o_τ. Formally, the problem and solution is identical to that of an end-of-pipe technology firm.

The choice of production function determines to what degree a specific technology can substitute between its inputs. For very specific technologies, such as an electrical heat pump, an assumption of zero substitution will often be the most appropriate choice. Furthermore, as with end-of-pipe technologies, the multiple output structure allows input-displacing technologies to contribute to multiple energy services and to model competition of overlapping technologies. For example, if a heat pump can supply either low or high heat, and these are represented by two distinct energy services, the heat pump will produce technology goods for both energy services.

Input-displacing component firms

As with component firms in the end-of-pipe tree, the input-displacing component firms facilitate overlap between technologies. Consider for example two heat pumps that provide similar heating services, e.g. similar temperature ranges, and cover the same percentage of energy demand. If one heat pump is fully adopted to cover this demand, the other is redundant. Therefore, the two technology goods U compete to provide the same component C.[17]

Formally, the problem and solution for input-displacing component firms is identical to that of component firms of the end-of-pipe type.

Intermediates producing firms

Intermediate goods producing firms of type E use a vector of components $\boldsymbol{C}^E$ in a scale-preserving manner as defined in definition B.3 using F^{in}_E to provide energy services. The firms thus solve:

$$\max_{\boldsymbol{C}^E} p_E F^{in}_E(\boldsymbol{C}^E) - \boldsymbol{p}_{C^E} \cdot \boldsymbol{C}^E.$$

We note that this is a standard optimization problem (definition B.2) with parameters $(p_E, \boldsymbol{p}_{C^E})$. The solution to this problem is per construction represented by a scale-preserving aggregator $\mathcal{P}^E$ as defined in definition B.3:

$$p_E = \sum_{C \in \boldsymbol{C}^E} \mathcal{P}^E_C\left(p_E, \boldsymbol{p}_{C^E}\right) p_C \tag{12a}$$

$$C = \mathcal{P}^E_C\left(p_E, \boldsymbol{p}_{C^E}\right) E, \qquad \forall C \in \boldsymbol{C}^E. \tag{12b}$$

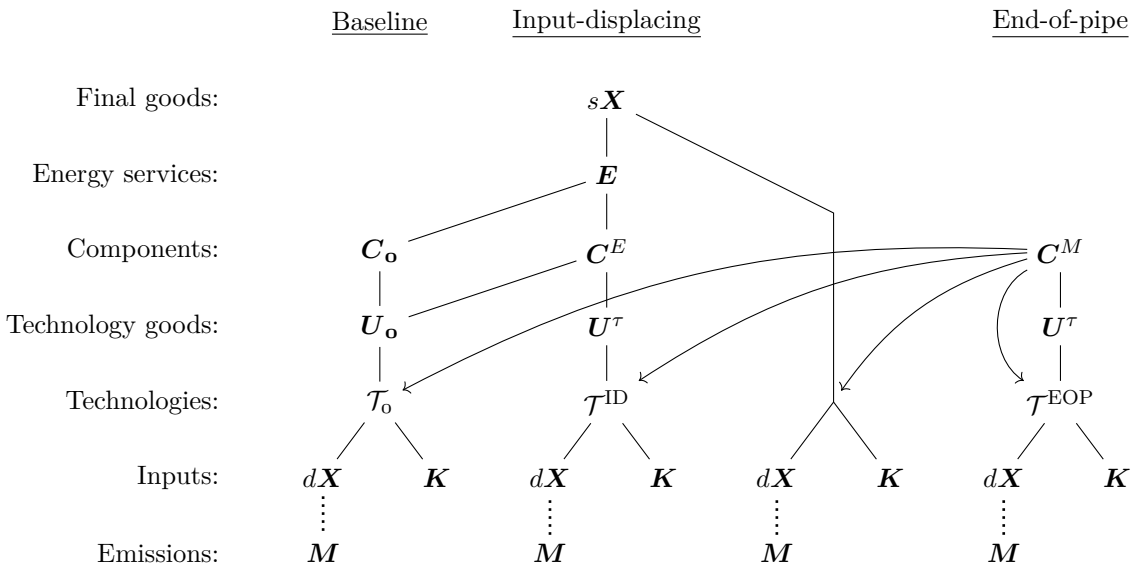
[17] We also allow partial overlap, i.e. that a technology only competes with other technologies for particular parts of its potential.

The choice of aggregator $\mathcal{P}^E$ reflects how the firms can substitute one type of energy service for another. Traditionally, in energy system models, the assumption is one of near-zero substitution between different energy services, with the argument that you cannot replace e.g. lighting with heating.[18]

Baseline technology firms

Baseline technology firms reflect the activities not accounted for by input-displacing technology firms. Thus, the organization of baseline firms depends on available technology data and input-output data, a point that we further elaborate on in section 4. Figure 3.2 illustrates the nesting structure when baseline firms are included. Baseline firms serve two purposes: First, they serve the role of firms that are eventually displaced when technology firms are phased in. For instance, assume that an electrical heat pump in a baseline year is 10% phased in. In this case, a baseline firm competes with the heat pump and produces 90% of the heat pump's potential in the baseline year. This is presented in figure 3.2 by including a line from baseline goods $\boldsymbol{U}_{\mathrm{o}}$ to components $\boldsymbol{C}^E$. Second, baseline firms are used to capture the supply of energy services that are not captured by technology firms. This is presented in figure 3.2 by including a new baseline component C_{o} for each energy service.

Figure 3.2: The organization of subsectors with baseline firms



Note: The arrows indicate flows of end-of-pipe abatement, from the end-of-pipe component firms to other firms in the industry that use energy inputs and hence, generate emissions. All firms who generate emissions are allowed to purchase end-of-pipe abatement components.

For each $E \in \boldsymbol{E}$ there is a single baseline technology that produces all baseline goods $\boldsymbol{U}^E$. For

[18] While this argument resonates with us, within a CGE model, this aggregator has to capture a multitude of effects. For instance, if firms in an industry face rapidly and sustained increases in prices on room heating, firms might reorganize so as to minimize the use of building capital going forward, and thus the need for heating services. Furthermore, it is likely that the composition of firms within this industry will adjust such that firms with less need for room heating grow relative to firms that need more of the expensive room heating. Due to these long run general equilibrium effects, we do not believe that the aggregator should necessarily reflect exactly zero possibilities for substitution.

the energy service type E let τ_{o}^E denote the relevant baseline firm. The baseline technology firm combines non-capital and capital inputs to a composite good κ_τ using a neo-classical production function F_{o}^{in}. It then splits this composite into one baseline technology good U_{o} for each component $C \in \boldsymbol{C}^E$ using a neo-classical output-allocation function F_{o}^o.

Denoting τ_{o}^E simply by τ the firm solves:

$$\begin{aligned} &\max_{\boldsymbol{U}^\tau, d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)} \boldsymbol{p}_{U^\tau} \cdot \boldsymbol{U}^\tau - \hat{\boldsymbol{p}}_X \cdot d\boldsymbol{X}(\tau) - \boldsymbol{p}_K(\tau) \cdot \boldsymbol{K}(\tau), \\ &\text{s.t. } \kappa_\tau = F_\tau^o\left(\boldsymbol{U}^\tau\right), \text{ and } \kappa_\tau = F_\tau^{in}\left(d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\right), \end{aligned} \tag{13}$$

This problem is essentially equivalent to any other technology firm; for the baseline firm, however, we specifically impose a linear transformation rate function in F_τ^o such that

$$\kappa_\tau = F_\tau^o(\boldsymbol{U}^\tau) \equiv \sum_{U \in \boldsymbol{U}^\tau} U/\gamma_U, \tag{14}$$

where γ_U measures the rate of transformation. We impose this structure for calibration purposes and because it has an interpretation appropriate for our setting: This representation is identical to a setup where separate firms each produce one of the $U \in \boldsymbol{U}^\tau$ goods using the same technology F^{in}, but with potentially different levels of total factor productivity (γ_U). Appendix B.1 shows this claim formally. This organization of baseline firms assumes that all goods that are not identified by technology data has the same input mix. Thus, if the technology data does not detail the type of inputs that the two heat pumps displace, we assume that they both displace a sector-average input mix employed by the baseline firm with which they both compete.

The set of equations that describe the solution to the problem in (13) is:

$$Z = H_z^\tau\left(\frac{\hat{p}_Z}{p_\tau}\right)\kappa_\tau, \qquad \forall Z \in \{d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\} \tag{15}$$

$$p_\tau = \sum_{Z \in \{d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\}} H_z^\tau\left(\frac{\hat{p}_Z}{p_\tau}\right)\hat{p}_Z \tag{16}$$

$$p_U = p_\tau \gamma_U, \qquad \forall U \in \boldsymbol{U}^\tau, \tag{17}$$

where the notation is similar to the one used for regular technology firms. We note that as opposed to other technology firms, we do not restrict F_τ^o to be scale-preserving. The reason is that the γ_U transformation parameters are thought of as capturing differences in productivity levels in the production function F^{in}. As appendix B.1 outlines, this set of equations can represent a setup where separate firms optimize with different levels of total factor productivity in F^{in}, and produce a single output in the trivial scale preserving manner where $\kappa = U$. Thus, while F_τ^o is not scale preserving, the general intuition of scale preservation from aggregating technology goods into energy services remains intact.

Adjusting the linear rates of transformation in calibration presents a simple way of scaling the adoption of technologies up or down: When a linear rate of transformation is increased it is directly

translated into a higher price of that particular technology good (cf. 17), increasing the adoption rate of the technologies that compete with this baseline technology good.

3.4 Final goods production

Firms in the final goods sub-sector use a vector of intermediate goods, $\boldsymbol{E}$, non-capital inputs, $d\boldsymbol{X}$, and capital inputs, $\boldsymbol{K}$, to produce a vector of outputs $s\boldsymbol{X}$. For our purposes, these goods can be arbitrarily nested; it depends on the relevant model.

For simplicity, we assume that all relevant goods are combined to produce the composite Y using a neoclassical-type function F_Y^{in}, and split onto the vector of outputs $s\boldsymbol{X}$ using a neoclassical output allocation function F_Y^o. The firm then solves

$$\begin{aligned} &\max_{s\boldsymbol{X},\boldsymbol{E},d\boldsymbol{X}} \boldsymbol{p}_X \cdot s\boldsymbol{X} - \boldsymbol{p}_E \cdot \boldsymbol{E} - \boldsymbol{p}_X d\boldsymbol{X} - \boldsymbol{p}_K \cdot \boldsymbol{K} \\ &\text{s.t. } Y = F_Y^o(s\boldsymbol{X}) \text{ and } Y = F_Y^{in}(\boldsymbol{E}, d\boldsymbol{X}, \boldsymbol{K}). \end{aligned} \tag{18}$$

Adopting the same notation as previous sections, the solution is characterized by the set of equations:

$$p_Y = \sum_{Z \in \{\boldsymbol{E}, d\boldsymbol{X}, \boldsymbol{K}\}} H_z^{Y,in}\left(\frac{\hat{p}_Z}{p_Y}\right)\hat{p}_Z \tag{19a}$$

$$p_Y = \sum_{X \in s\boldsymbol{X}} H_x^{Y,o}\left(\frac{p_X}{p_Y}\right)p_X \tag{19b}$$

$$Z = H_z^{Y,in}\left(\frac{\hat{p}_Z}{p_Y}\right)Y, \qquad Z \in \{\boldsymbol{E}, d\boldsymbol{X}, \boldsymbol{K}\} \tag{19c}$$

$$X = H_x^{Y,o}\left(\frac{p_X}{p_Y}\right)Y, \qquad X \in s\boldsymbol{X} \tag{19d}$$

Appendix F.2 includes the complete set of equations and endogenous variables of the combined input-displacing and final goods production tree.

3.5 Capital sub-sector

All sub-sectors covered so far only face static optimization problems; the dynamic aspects are left to the capital sub-sector. In the most general case, we assume that each technology, baseline, and final goods firm has a corresponding capital firm that faces adjustment costs. For each combination of firm and capital type, a firm accumulates capital and rents it out at the price $p_{K,t}$. The capital firm faces convex installation costs that depend on the level of investments, I_t, and the size of the stock K_t. The firm thus solves:

$$\max_{\{I_j, K_j\}_{j=t}^T} \sum_{j=t}^{T} R_{j,t}^{-1}\left[p_{K,j}K_j - \Psi\left(I_j, K_{j-1}\right) + \lambda_j\left(I_j + K_{j-1}(1-\delta) - K_j\right)\right] \tag{20a}$$

$$\text{s.t. } K_{t-1} \geq 0 \text{ and } f(K_T) = 0, \tag{20b}$$

where $R_{j,t}$ is the compounded interest rate factor between times j and t, Ψ is a convex installation cost function, λ_j is a shadow cost, δ is the depreciation rate of capital, T is the time horizon, and $f(K_T)$ represents a suitable transversality condition to the dynamic problem.[19] The solution to this problem yields a rental rate on capital:

$$p_{K,j} + \frac{1-\delta}{R_{j+1,j}} \frac{\partial \Psi_{j+1}}{\partial I_{j+1}} = \frac{\partial \Psi_j}{\partial I_j} + \frac{1}{R_{j+1,j}} \frac{\partial \Psi_{j+1}}{\partial K_j} \tag{20c}$$

In optimum the rental rate on capital balances a number of marginal effects: First, marginally increasing the level of capital yields a direct revenue $p_{K,j}$ and related cost of investment $\partial \Psi / \partial I_j$. Beyond this, marginally increasing the current capital stock yields a $1-\delta$ increase in the future stock of capital, which in optimum is valued at the related discounted marginal cost of investment, $\partial \Psi_{j+1} / \partial I_{j+1}$. Finally, increasing the capital stock lowers the future installation costs, $\partial \Psi_{j+1} / \partial K_j < 0$.

The technology-specific adjustment costs can be used to model a gradual phasing in of the various technologies. However, we note that making capital technology-specific is computationally expensive. A less computationally demanding alternative is to model two types of capital, one for capital demanded by technologies and one for conventional capital demanded elsewhere in the economy. Appendix H elaborates.

3.6 The aggregate sector solution

The solution to the aggregate sector, i.e. the simultaneous solution to all subsectors, is found by clearing markets for outputs $s\boldsymbol{X}$, capital $\boldsymbol{K}$, non-capital inputs $d\boldsymbol{X}$, intermediate goods $\boldsymbol{E}$, components $\boldsymbol{C}$, and technology goods $\boldsymbol{U}$, when all sub-sectors optimize as outlined throughout section 3. The full system of equations can be found in appendix F.

4 Calibration Method

The relevant model parameters are identified using two sources of information: Technology catalogues and input-output tables. The method for identifying parameters depends on how detailed the data is and can be adjusted to accommodate variations in the available data. The following outlines the case with sufficiently detailed data to utilize all the features highlighted throughout the paper. Appendix D adjusts this to the case with less detailed data. The data assumed available for the full calibration procedure is summarized in table 3.

The data on overlap between technologies is used to create the relevant nesting structure, in particular to establish the correct structure of technology components. For a detailed explanation of how to construct the tree and calibrate the model, see appendix C which does this using a simple, specific example technology catalogue. The appendix also explains how overlap information should

[19] The transversality condition depends crucially on the type of model that is employed; a finite horizon problem typically has $f(K_T) = K_T$, whereas an infinite horizon problem typically involves $f(K_T) = K_T - K_{T-1}(1+g)$ with g denoting the long run growth rate.

Table 3: Data used in calibration

Name	Symbol	Description
Technology data for each τ:[a]		
Input intensities	$\overline{\boldsymbol{x}}(\tau)$	Non-capital inputs per unit of output.
Unit cost	$\overline{c}_\tau$	Costs per relevant output. Reflects capital and non-capital costs.
Coverage potential	$\overline{\boldsymbol{\theta}}^E_\tau$ or $\overline{\boldsymbol{\theta}}^M_\tau$	Potential measured as τ's maximum attainable output share of E (input-displacing) or M_{o} (end-of-pipe).
Current application	$\overline{\boldsymbol{\vartheta}}^E_\tau$ or $\overline{\boldsymbol{\vartheta}}^M_\tau$	Current coverage measured as τ's output share of E (input-displacing) or M_{o} (end-of-pipe) in a base year.
Overlap	None	Reflects to what degree technologies compete for the same shares of E or M_{o}.[b]
Input-output data:		
Energy services	$\overline{\boldsymbol{E}}$	Quantity of energy services.
Energy inputs	$\overline{\boldsymbol{X}}^E, \overline{\boldsymbol{K}}^E$	Quantity of inputs used for each energy service.
Sector inputs	$d\overline{\boldsymbol{X}}, \overline{\boldsymbol{K}}$	Aggregate quantity of inputs used in the sector.
Sector outputs	$s\overline{\boldsymbol{X}}$	Final goods production.
Sector investments	$\overline{\boldsymbol{I}}$	Aggregate investments undertaken in the sector.
Sector emissions	$\overline{\boldsymbol{M}}$	Aggregate emissions after end-of-pipe abatement.

[a] Technology data is generally defined for all τ whereas input-output data is collected for each sector. Unit costs, input intensities and all input-output data are measured in some base year.
[b] Specifically, the overlap information is stated as a vector of fractions of coverage potential and the group of other technologies that share this part of the potential. The stated fractions refer to mutually exclusive parts of potential and hence cannot sum to more than one within an E or M. See appendix C for additional details.

be organized and used for constructing the nesting structure.[20]

Some of the model parameters are exactly identified by a counterpart in the technological data whereas others are only identified implicitly by ensuring e.g. that the observed level of fuels is used in a given industry. To perform the calibration we rely on adjustments to the various aggregators used to describe optimality throughout section 3. While we still do not take a stance on the specific choice of aggregators, production technology etc., we assume that all aggregators ($\mathcal{P}^y_x$) and demand functions (H^y_x) are equipped with a set of share parameters μ^y_x for each node $X \in \boldsymbol{X}^Y$ and a smoothing parameter σ_y for each parent node Y. The use of share and smoothing parameters are standard in CGE models that rely on nested CES functions: Share parameters govern the optimal shares $\boldsymbol{x}$ in the case where all prices in a nest is equalized, i.e. $p_x = p_y$ for all $X \in \boldsymbol{X}^Y$, whereas smoothing parameters govern how much shares react to price changes.[21]

The principles behind identification and calibration of parameters are generally relatively standard for CGE models: In most cases, share parameters are either identified one-to-one by observations in technology data or calibrated to fit one-to-one with calibration targets from input-output data. This is for instance the case for how technology- and baseline firms draw on inputs, how

[20] All programs related to this paper are collected in a Github repository at https://github.com/ChampionApe/GPM_v05.
[21] For the interested reader, we refer to appendix B that includes a formal definition (B.4) and shows examples of relevant functional forms e.g. CES, CET, d-level nested MNL, and normalized CES and CET functions.

energy service firms combine components, and how final goods firms draw on energy services and other inputs. For smoothing parameters, we generally propose to let technical data guide the choice in the technology related nests: The choice of smoothing parameters governing how technology firms use inputs, how component firms combine technology goods, and how energy service firms combine components, should be guided by the catalogue.[22] For the final goods firms that in many ways represent a 'standard' sector in a CGE model, identification may follow standard procedures, e.g. by applying a structural estimation (Kronborg et al. 2019), or use external estimations (Aguiar et al. 2019). Appendix G discusses identification of smoothing parameters using the standard methods.

In a few important ways, however, our calibration procedure is non-standard. In the input-displacing part, we leverage the information of how costly a technology is ($\bar{c}_\tau$) and how much the technology is applied in a base year ($\overline{\boldsymbol{\vartheta}}_\tau^E$) to infer the costs of the competing baseline technology firm and, relatedly, how technology and baseline firms compete when they are bundled as components. Similarly, in the end-of-pipe subtree, we can leverage the information of how costly a technology is ($\bar{c}_\tau$) and how much the technology is applied in the baseline year ($\overline{\boldsymbol{\vartheta}}_\tau^M$) to infer how large heterogeneity there is in adoption costs (the G function in section 3.1). In the following, we focus on these non-standard elements of the calibration and refer to appendix G for an overview of the remaining calibration procedure and appendix C for a complete example of how to calibrate a simple tree.

4.1 Matching current applications

Input-displacing technologies

For input-displacing technologies, we measure the degree of current coverage (ϑ_τ^E) as:

$$\vartheta_\tau^E \equiv \sum_{U \in \mathcal{D}_{U^\tau}(E)} \frac{U}{E} = \sum_{U \in \mathcal{D}_{U^\tau}(E)} \mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) \mathcal{P}_C^E\left(p_E, \boldsymbol{p}_C\right). \tag{21}$$

The first part defines the current coverage as the production that technology τ contributes to the energy service E as a share of total E. The second part, based on equations (11b) and (12b), states that this can be written in terms of how well the relevant technology good U competes in the market for component C, and how significant a part this component is compared to total E. To fit the level of current coverage of a technology to its target, we rely on two sets of parameters: The share parameters in the aggregator $\mathcal{P}^C$ and the rates of transformation for baseline technologies ($\boldsymbol{\gamma}_{U^{\tau_0}}$), see equations (11b) and (14). The former specifies how technology goods compete for market shares in the relevant C, and the latter identifies the price level for the relevant competing baseline technologies in that market.[23]

[22] See appendix G for additional information and discussion of this.

[23] Equation (21) hints that other parameters could be used to target the level of current coverage: share parameters in the aggregator $\mathcal{P}^E$ and parameters determining the price levels of technology goods. However, share parameters in the aggregator $\mathcal{P}^E$ are used to ensure the correct potential of technologies ($\overline{\boldsymbol{\theta}}_\tau^E$), and the price levels for technology firms are identified by the unit costs in data $\bar{c}_\tau$.

We note that the ex ante neutral assumption would be to fix the share parameters in the aggregator $\mathcal{P}^C$ to be uniformly distributed, as this would indicate a market where the only relevant metric for adoption of a technology is the relative price between said technology and its baseline competitor. In contrast, we have no neutral value for the rates of transformation, as we do not have any technical identification of the unit costs of the baseline technologies. Instead, it seems intuitive to infer them from the level of current coverage in a base year: If a technology is only 10% phased in, it stands to reason that the competing baseline technology good is relatively cheap.

Indeed, as the intuition above reveals, in the simple case without any overlap between technology firms, it is sufficient to rely only on these rates of transformation ($\boldsymbol{\gamma}_{U^{\tau_o}}$) to target the levels of current coverage ($\overline{\boldsymbol{\vartheta}}_\tau$). However, in the case with overlapping technologies, the market for component C potentially has a large number of technology firms competing against a single baseline technology. In this case, the single rate of transformation that identifies the price level of the baseline technology is not sufficient to ensure that the market shares for all the competing technology firms follow the technological data: In this case, we have to rely on share parameters in the aggregator $\mathcal{P}^C$ as well.

In the case where technologies can be grouped into components, where all overlap 100%, the rates of transformation of baseline technologies and the share parameters in the aggregators $\mathcal{P}^C$ are exactly sufficient to target the levels of current coverage. Consider, for instance, the case with two heat pumps that overlap 100% in technical potential for providing spatial heating, and compete against a baseline heating technology. Assume that the technology catalogue states that both heat pumps cover 10% of spatial heating in a baseline scenario ($\overline{\vartheta}_1^{sh} = \overline{\vartheta}_2^{sh} = 10\%$), but one of them is slightly more expensive than the other ($\bar{c}_1 > \bar{c}_2$). By adjusting the rate of transformation for the baseline technology, we can ensure that baseline heat is sufficiently cheap to cover 80%. The relative level of the share parameters for the two technology goods is used to ensure that both cover 10%, even though one of the technologies is slightly more expensive ($\mu_1^{sh} > \mu_2^{sh}$).

Finally, in the most general case with arbitrarily overlapping technologies, there are more free parameters in $\boldsymbol{\gamma}_{U^{\tau_o}}$ and $\boldsymbol{\mu}_U^C$ than calibration targets in $\overline{\boldsymbol{\vartheta}}_\tau^E$. The issue of having more parameters than targets arises whenever a technology competes to provide more than one component under the same E.[24] Consider again the scenario with two heat pumps that produce spatial heating (E^{sh}) and let their overlap share be less than 100 percent. Each heat pump produces spatial heating by contributing to two different components, resulting in three components in total, where one of the components represents the overlapping potential. Suppose that the first heat pump supplies spatial heating through the components C_1^{sh} and C_2^{sh} and that the latter is shared with the second heat pump. In this case, the technology data states that the first heat pump produces $\overline{\vartheta}^{sh} E^{sh}$, but not how much is contributed through each of the components C_1^{sh}, C_2^{sh}. Technically, the heat pump competes against the baseline technology firm which also produces technology goods for each of these components. Recall that component markets are only established to take the overlap information into account. Thus, intuitively, it seems reasonable that the baseline technology goods

[24] This type of nesting tree only arises when technologies overlap by less than a 100%.

produced for C_1^{sh} and C_2^{sh} enter both with roughly the same price levels.

To formalize this we rely on an identification strategy that has two core principles: First, we prefer to target current coverage levels through rates of transformation for the baseline technologies, i.e. through prices, rather than the share parameters in the aggregator $\mathcal{P}^C$. Second, technology firms that compete in multiple component markets within the same energy service should compete against roughly the same baseline technology. Formally, we minimize a criterion function,

$$\begin{aligned} &\min_{\{\gamma_C, \boldsymbol{\mu}_U^C\}_{C \in \boldsymbol{C}^{ID}}} \sum_{\tau \in \mathcal{T}^{ID}} \sum_{C \in \mathcal{D}_C(\tau)} \left(\gamma_C - \gamma_\tau^E\right)^2 + \omega_\mu^{ID} \sum_{\tau \in \mathcal{T}^{ID}} \sum_{U \in \boldsymbol{U}^\tau} \left(\mu_U^C - \overline{\mu}^C\right)^2, \\ &\text{s.t. } \vartheta_\tau^E = \overline{\vartheta}_\tau^E, \qquad \forall \tau \in \mathcal{T}^{ID},\ E \in \boldsymbol{E} \end{aligned} \tag{22}$$

where $\overline{\mu}^C$ is the level of share parameters $\boldsymbol{\mu}_U^C$ when uniformly distributed, $\omega_\mu^{ID} > 1$ is a weight that penalizes deviations from the uniform distribution, and γ_τ^E is a measure of a technology-average rate of transformation implicitly identified by solving

$$\sum_{U \in \mathcal{D}_{U^\tau}(E)} \mathcal{P}_U^C\left(p_C, \tilde{\boldsymbol{p}}_U\right) \mathcal{P}_C^E\left(p_E, \boldsymbol{p}_C\right) = \overline{\vartheta}_\tau^E, \tag{23}$$

with $\tilde{\boldsymbol{p}}_U$ being a vector of prices populated by elements:

$$\tilde{p}_U = \begin{cases} p_U, & \text{if } U \text{ is not a baseline technology good.} \\ \gamma_\tau^E p_{\tau_0^E}, & \text{if } U \text{ is a baseline technology good.} \end{cases} \tag{24}$$

The parameter γ_τ^E formalizes the second principle of the identification strategy: If a technology firm competes against the same baseline firm in all components under the relevant E, the rate of transformation γ_τ^E ensures that the current coverage of the technology firm reaches the target $\overline{\vartheta}_\tau^E$. We prefer a large value of ω_μ^{ID} to ensure that most of the adjustment is achieved through price differences, i.e. through γ parameters. However, as the parameter is not identified, robustness checks of the calibration should be carried out with different values of ω_μ^{ID}.

End-of-pipe technologies

For the end-of-pipe sub-sector, we measure the degree of current coverage (ϑ_τ^M) as:

$$\vartheta_\tau^M \equiv \sum_{U \in \mathcal{D}_{U^\tau}(M)} \frac{U}{M_0} = \sum_{U \in \mathcal{D}_{U^\tau}(M)} \mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) \theta_C G_C\left(p_M - p_C\right). \tag{25}$$

The first part defines current coverage as the sum of abatement services of M produced by τ, as a share of pre-abatement emissions M_0. The second part states that each term in the sum can be written in terms of how well the relevant technology good U competes in its respective market for component C, the potential of the component θ_C, and the degree of technology adoption $G_C(\cdot)$. Recall that the G_C function can be specified as any cumulative density function with continuous support and be interpreted as heterogeneity in the costs of adopting end-of-pipe abatement technologies. Whereas the current coverage of input-displacing technologies can mostly be attained

by adjusting cost levels of baseline technologies, the end-of-pipe technologies face no such baseline competitor. Instead, we have to rely on adjustments to the distribution function G_C.

We note that, ideally, the distribution function G_C is identified by technology data outside the model. Thus, we emphasize that calibration of G_C to fit levels of current application depends crucially on the specific distribution. In the case where G_C is not known, we propose to use a simple two-parameter distribution (μ_C^G, σ_C^G) that identifies the mean and standard deviation respectively.[25] Recall from section 3.1 that letting G_C be a mean-zero distribution ($\mu_C^G = 0$), the average cost of applying the underlying abatement technologies corresponds to the unit cost observed in technological data. Thus, it may be preferable to calibrate the model by primarily adjusting the variance of the distribution (σ_C^G). Following the same line of reasoning as in the input-displacing part, we thus propose to calibrate the end-of-pipe part by minimizing a criterion function,

$$\begin{aligned} &\min_{\{\sigma_C^G, \mu_C^G\}_{C \in \boldsymbol{C}^{EOP}}} \sum_{C \in \boldsymbol{C}^{EOP}} \left[\left(\sigma_C^G - \overline{\sigma}_C^G \right)^2 + \omega_\mu^{EOP} \left(\mu_C^G \right)^2 \right], \\ &\text{s.t. } \vartheta_\tau^M = \overline{\vartheta}_\tau^M, \qquad \forall \tau \in \mathcal{T}^{EOP},\ M \in \boldsymbol{M}, \end{aligned} \tag{26}$$

where $\overline{\sigma}_C^G$ is a target level of heterogeneity and $\omega_\mu^{EOP} > 1$ is a weight that penalizes deviations from the mean-zero distributions.

4.2 Remaining calibration issues

Emission intensities

We assume that total emissions post end-of-pipe abatement $\overline{\boldsymbol{M}}$ for the industry are available for calibration of ϕ_X^M. Calibration must ensure that emissions pre end-of-pipe abatement less abatement equals post abatement emissions:

$$\begin{aligned} \overline{M} &= M_0 - A_M \\ &= \sum_{X \in \boldsymbol{X}} \phi_X^M d\overline{X} \left[1 - \sum_{C \in \mathcal{D}_C(A_M)} \theta_C G\left(p_M - p_C\right) \right], \qquad \forall M \in \boldsymbol{M} \end{aligned} \tag{27}$$

where the contents of the bracket can be calculated from data on current applications.[26] Since end-of-pipe abatement, by definition, removes emissions irrespective of the inputs that caused them, we should not designate certain end-of-pipe abatement to certain types of inputs X. Instead, we adjust all emission intensities uniformly by defining

$$\phi_X^M = \varphi_M \overline{\phi}_X^{M,0}, \tag{28}$$

[25] For instance, the normal distribution or logit distribution.

[26] For example, if the catalogue features three end-of-pipe technologies with a total current application of 17 % for CO2 emissions, then the bracket evaluates to 1-0.17=0.83 when M = CO2.

where $\overline{\phi}_X^{M,o}$ is an estimate of the emission intensities without any abatement. Calibration then entails fixing M at the target level $\overline{M}$, and endogenizing φ_M. This methodology ensures that if end-of-pipe technologies not listed in the catalogue exist, or if the industry has process emissions (emissions not directly attributable to inputs), it will be captured by the emission intensities.

Time horizon

While not directly a calibration issue, we stress that simulations from the model should focus on periods where the technology catalogue can reasonably be relevant. In the longer run, innovation will provide new, better technologies that are naturally not covered by technology catalogues today. To remedy this, it may be more appropriate to use a more conventional CGE structure after some time horizon T. This does not imply that the long run behavior is independent of technology data. For instance, the share-parameters in the conventional long run production structure can be updated to reflect the input-shares predicted by the technology-rich model in $T-1$. Similarly, sector-specific emission coefficients can be updated to reflect the level predicted by the technology-rich model in $T-1$.

5 Simulation experiments

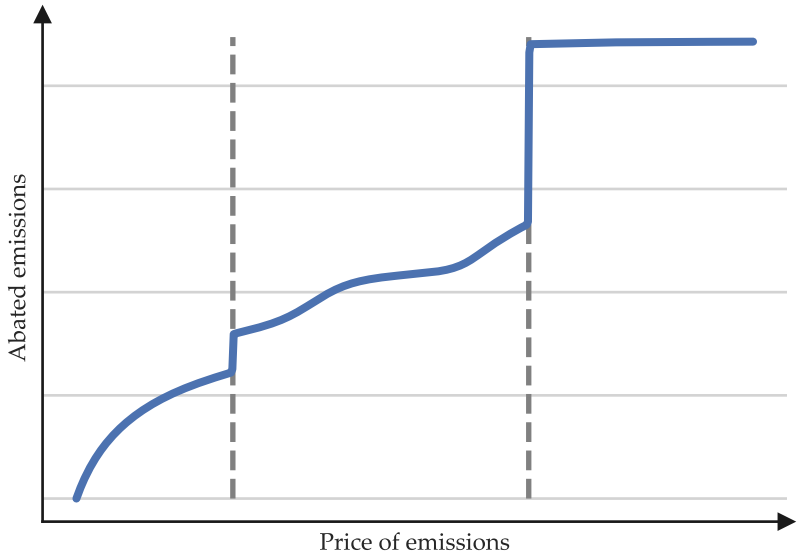
To illustrate the mechanisms in the model, we consider a static, partial equilibrium toy model with a single industry that can access four technologies: Two input-displacing and two end-of-pipe. Three technologies rely primarily on electricity while the fourth technology (input-displacing) relies on oil. The input-displacing technologies compete against an emission intensive baseline technology that relies on coal and oil. The two end-of-pipe technologies are independent, while the two input-displacing technologies are 100% overlapping. In other words, the two input-displacing technologies compete for the exact same market. Oil and coal emit CO_2 which is priced at an exogenous rate p_{CO_2}. To solve the partial equilibrium model, we fix input prices and the quantity of the single output. Appendix I contains all details related to the toy model.

The marginal abatement cost curve of the toy model is simulated in figure 5.1 by changing the price on CO_2 emissions from its initial low value and measuring abated emissions as the change compared to the initial case. The curve primarily consists of three channels: Pure input substitution, end-of-pipe abatement, and input-displacing technology abatement. When p_{CO_2} is low, there is little technological abatement; instead, the baseline technology substitutes away from emitting inputs such as coal and oil and towards cleaner inputs such as electricity. When p_{CO_2} increases, the industry adopts the cheap end-of-pipe abatement equipment, the two input-displacing technologies, and finally the more expensive end-of-pipe technology. The following sections highlight important traits of the toy model.

Endogenous costs of technologies

An important feature of the model is that the costs of technologies depend on equilibrium prices. To emulate the effect of coupling our toy model to a general equilibrium, we assume that the price

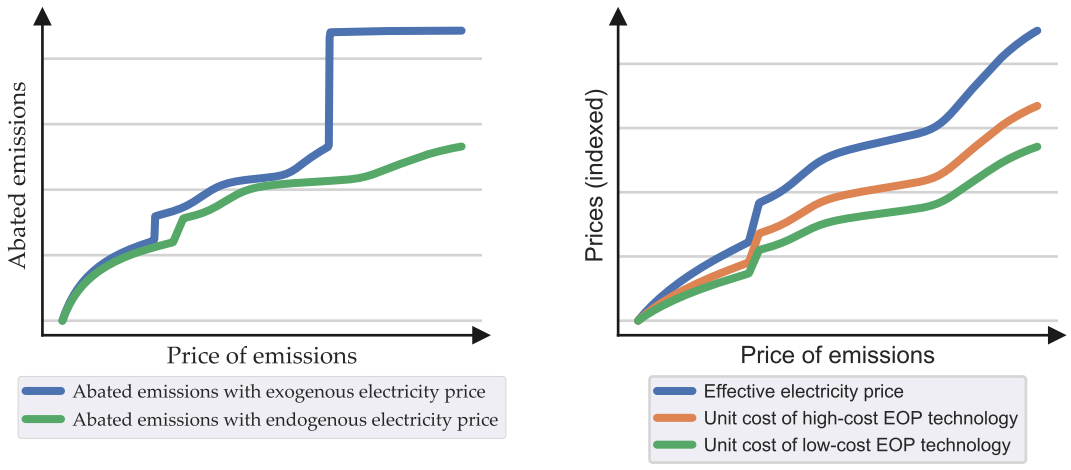
Figure 5.1: Abated emissions



Note: The vertical lines indicate the unit costs of the two end-of-pipe technologies in the catalogue.

on electricity is a linear function of demand for electricity. Thus, if abatement technologies rely heavily on electricity, there is a feedback effect of higher prices. Figure 5.2 illustrates the effect of this feedback effect on the MAC curve in the toy model. With a higher p_{CO_2} comes a higher demand for electricity and thus a higher price on electricity. The effect is that the first end-of-pipe technology is adopted at a higher emission price, and the second end-of-pipe technology is not profitable within the shown range. Figure 5.2b shows how the prices of the two end-of-pipe technologies increase as the emission tax causes demand for electricity to rise.

Figure 5.2: Endogenous costs of technologies



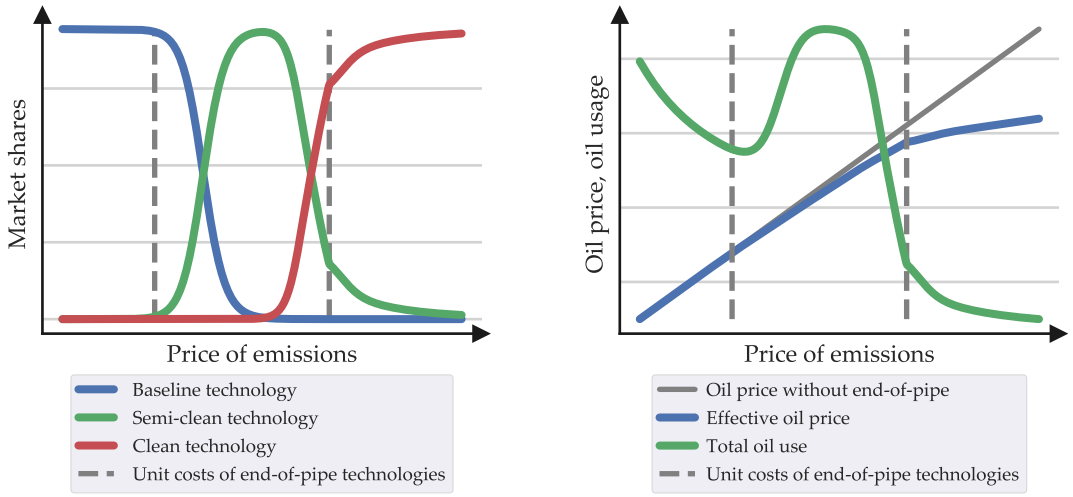
(a) Abated emissions

(b) Cost of technologies

Overlapping potentials

Another important feature of the model is the overlap in technologies' potentials. In the toy model, for instance, the two input-displacing technologies overlap 100%, implying that they compete against each other and a baseline technology to provide the same component of energy demand. Figure 5.3a illustrates the market shares of the three competing technologies as a function of the price of CO_2 emissions. For a low p_{CO_2} the baseline technology is cheapest. However, as this is also the most emission intensive, it is eventually replaced by the input-displacing technologies. In medium ranges of the price on emissions, the 'semi-clean' technology that relies on oil is adopted, and finally, for large values, the 'clean' technology that relies on electricity is adopted.

Figure 5.3: Overlapping potentials



(a) The market shares of overlapping technology goods with varying emission intensities

(b) Oil use and oil price

Note: Panel (a) depicts the quantities of three technology goods U as a share of the component C that they contribute to.

Furthermore, note that this simple technology catalogue translates into some rather complex demand functions e.g. for oil. Figure 5.3b illustrates that initially, for low emission taxes, there is a general substitution away from emission-intensive inputs as oil. At medium prices, this is reversed due to the semi-clean input-displacing technology that uses oil, and finally oil is more or less phased out at sufficiently high emission taxes. Thus, if feedback effects from the general equilibrium are significant, getting the right representation of technologies' overlap can be quite important.

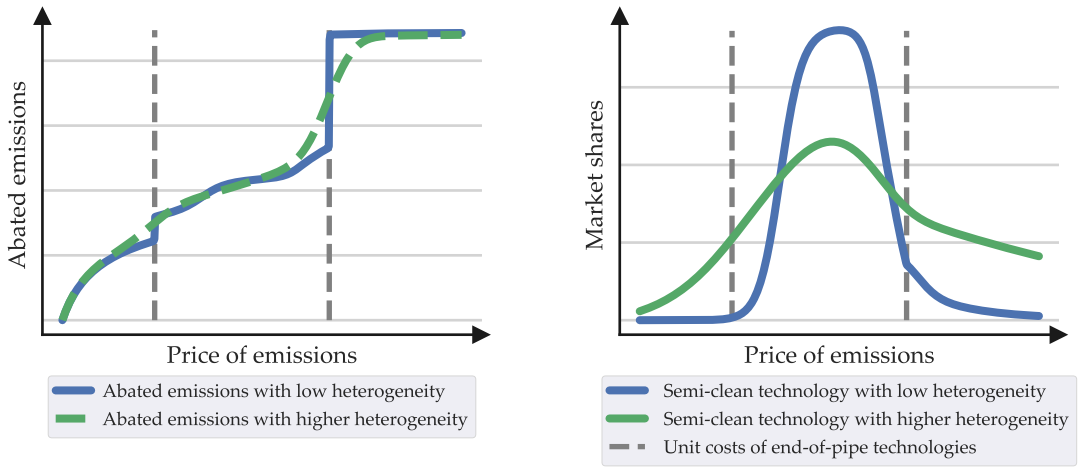
Gradual adoption

A third feature of the model is that adopting an end-of-pipe abatement technology on the industry level is not a binary decision. Instead, the costs are distributed with some degree of heterogeneity, which determines the rate of adoption. Increasing the level of heterogeneity yields a more smooth MAC curve as illustrated in figure 5.4a.

Similarly, the smoothing parameters chosen for the aggregation of technology goods into components determines how gradual the adoption of input-displacing technologies is. This is illustrated in

figure 5.4b: When we have a low degree of smoothing, the semi-clean technology is fully adopted for medium ranges of emission taxes. In the smooth case, however, it is never fully adopted. While the appropriate level of smoothing may vary, this ensures that the model does not feature the 'winner-takes-all' property that many technology rich models are criticized for (Bataille et al. 2006).

Figure 5.4: Gradual adoption



(a) Abated emissions with low and high heterogeneity/smoothing

(b) Overlapping technology goods, with increased smoothing parameter

Note: Panel (a) depicts abated emissions and panel (b) depicts the market share (of the corresponding component) of the semi-clean technology. Both sub-figures depict the base scenario against an alternative with more heterogeneity and hence smoothing.

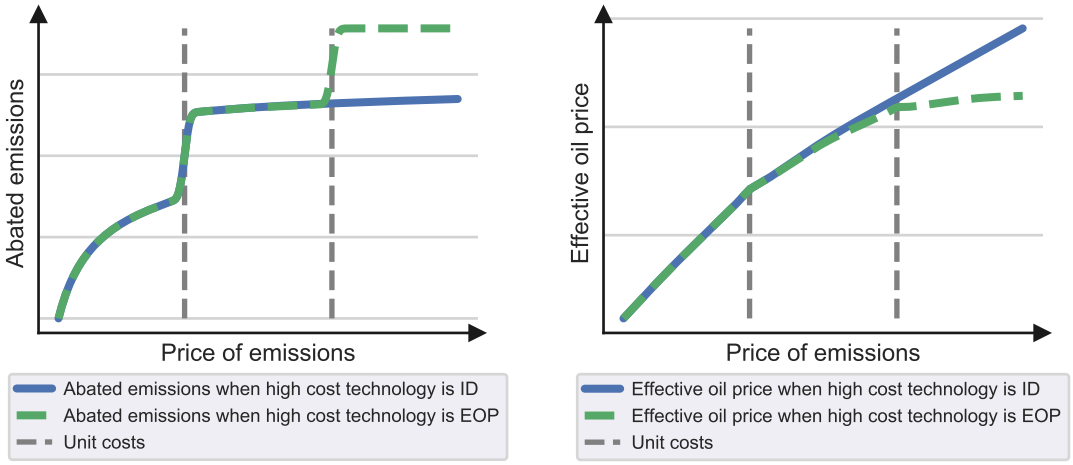
Distinction between input-displacing and end-of-pipe technologies

Finally, an important feature of the model is that we can distinguish between end-of-pipe and input-displacing technologies. In particular, we have the general result that end-of-pipe and input-displacing technologies are substitutes, in the sense that adopting one makes the other less profitable. Adoption of end-of-pipe technologies lowers the effective emission intensity of emitting inputs such as oil. This implies that adopting end-of-pipe technologies provides a discount on the price of emission intensive inputs, and thus, *ceteris paribus*, increases the profitability of more emission intensive technologies. The rate at which the clean technology is adopted in figure 5.3a and the rate at which oil is reduced in figure 5.3b both drop with the adoption of the second end-of-pipe technology. This illustrates how the interactions between clean input-displacing technologies and end-of-pipe abatement technologies cause them to be substitutes.

As end-of-pipe and input-displacing technologies are substitutes, distinguishing between the two is essential. Consider a catalogue with just two technologies: A cheap end-of-pipe technology and an expensive, input-displacing technology. Both technologies are clean, with all emissions being generated by the baseline technology. Figure 5.5 shows the effect of mistakenly categorizing both as end-of-pipe technologies. When both technologies are treated as end-of-pipe, and none of them rely on emission-intensive inputs, they are both phased in at relatively high emission

prices (5.5a). However, if the expensive technology is in fact input-displacing, adopting the end-of-pipe technology lowers the effective price on oil which makes the input-displacing technology less profitable relative to the baseline technology. Thus, even without general equilibrium feedback effects, not distinguishing between end-of-pipe and input-displacing technologies can lead to large errors in the estimation of the MAC curve.

Figure 5.5: Distinction between input-displacing and end-of-pipe technologies



(a) Abated emissions

(b) Effective oil price

Notes: Panel (a) depicts abated emissions as a function of the emission tax when the expensive technology is either input-displacing or end-of-pipe. Panel (b) depicts the effective oil prices in the same two scenarios. The vertical lines indicate the unit cost of the technologies.

6 Conclusion

We developed a methodology that utilizes detailed technology data in economic models. The method allows for a more accurate estimation of marginal abatement costs. The method represents a return to basics for economists as technology adoption is modelled as equilibria on markets for specific technology goods. This allows our framework to overcome several drawbacks of current best practices in estimating the marginal abatement costs. First, our model distinguishes between input-displacing and end-of-pipe technologies, and show that they should be treated as substitutes in abating emissions: By adopting an end-of-pipe technology, it effectively reduces the emission intensity of dirty technologies, making them more profitable. Second, by representing each technology as a firm that draws on inputs, we endogenize the cost of technologies implying e.g. that if the price of electricity rises, clean electricity-based technologies become more expensive. Third, heterogeneity in costs between firms motivates sluggish within-period adoption of technologies, and investment costs generates sluggish adoption over time. Fourth, we allow overlap between technologies: When technologies overlap they compete for the same market share and adopting one reduces the potential of the others.

We illustrated these core features in toy model simulations in a partial equilibrium setting

using a simple carbon price experiment. The simulations showed how overlap can give rise to non-monotonic effects of carbon prices on resource use in the industry. We also illustrated that the categorization of a technology as either input-displacing or end-of-pipe is important for predictions because they are substitutes in nature. When a technology is categorized as end-of-pipe rather than input-displacing, it can cause the model to predict that the technology is not adopted rather than fully adopted, *ceteris paribus*.

Identification of technologies relies on technology catalogues which are expert-based assessments of available technologies including their costs, inputs and potentials. Fundamentally, this allows the model to be based on the best available knowledge about the technologies relevant for the green transition, rather than projections based on historical data.

Our methodology is well suited for implementation in existing CGE models, because it relies on standard methods such as profit maximizing firms. Furthermore, we showed that incorporating the optimal choice of how much end-of-pipe abatement to purchase is simply a matter of updating the prices of emission intensive inputs. Applying the framework in general equilibrium allow us to capture important interaction effects between technology adoption and input prices, resulting in more realistic estimates of abatement costs.

References

Aguiar, Angel et al. (2019). "The GTAP Data Base: Version 10". *Journal of Global Economic Analysis* 4.1, pp. 1–27. ISSN: 2377-2999.

Amundsen, Eirik S, Lars Gårn Hansen, and Hans Jørgen Whitta-Jacobsen (Nov. 2018a). "Regulation of location-specific externalities". English. WorkingPaper 9/18. Department of Economics, University og Bergen.

Bataille, Chris, Frederic Ghersi, and Jean-Charles Hourcade (2006). "Hybrid Modeling: New Answers to Old Challenges - Introduction to the Special Issue of The Energy Journal". *The Energy Journal* Hybrid Modeling.

Berg, Rasmus K. and Janek Eskildsen (2021). "Electricity and District Heat systems in a Computable General Equilibrium Framework".

Böhringer, Christoph (1998). "The synthesis of bottom-up and top-down in energy policy modeling". *Energy Economics* 20.3, pp. 233–248. ISSN: 0140-9883.

Böhringer, Christoph and Thomas F. Rutherford (2009). "Integrated assessment of energy policies: Decomposing top-down and bottom-up". *Journal of Economic Dynamics and Control* 33.9, pp. 1648–1661. ISSN: 0165-1889.

Bovenberg, A. Lans, Lawrence H. Goulder, and Mark R. Jacobsen (June 2008). "Costs of alternative environmental policy instruments in the presence of industry compensation requirements". *Journal of Public Economics* 92.5-6, pp. 1236–1253.

Enkvist, P, Tomas Nauclér, and Jerker Rosander (2007). "A cost curve for greenhouse gas reduction". *McKinsey Quarterly* 1, p. 34.

Faehn, Taran and Elisabeth T. Isaksen (2016). "Diffusion of Climate Technologies in the Presence of Commitment Problems". *The Energy Journal* 0.Number 2.

Gillingham, Kenneth et al. (2013). "The rebound effect is overplayed". *Nature* 493, pp. 475–476.

Grepperud, Sverre and Ingeborg Rasmussen (2004). "A general equilibrium assessment of rebound effects". *Energy Economics* 26.2, pp. 261–282. ISSN: 0140-9883.

Hertel, Thomas et al. (2003). "How Confident Can We Be in CGE-Based Assessments of Free Trade Agreements?" Working paper.

Kiuila, O. and T.F. Rutherford (2013a). "Piecewise smooth approximation of bottom–up abatement cost curves". *Energy Economics* 40.C, pp. 734–742.

Kiuila, O. and T.F. Rutherford (2013b). "The cost of reducing CO2 emissions: Integrating abatement technologies into economic modeling". *Ecological Economics* 87.C, pp. 62–71.

Kronborg, Anders F., Christian S. Kastrup, and Peter P. Stephensen (2019). “Estimating the Constant Elasticity of Substitution when Technical Change is Time-Varying: A Kalman Filtering Approach”. Working paper.

Loulou et al. (2016). *Documentation for the TIMES Model*. iea-etsap.org/index.php/documentation. Accessed: 15-10-2020.

Plaza, Marta G, Sergio Martínez, and Fernando Rubiera (2020). "CO2 Capture, use, and Storage in the Cement Industry: State of the Art and Expectations". *Energies* 13.21, p. 5692.

Rive, Nathan (Jan. 2010). "Climate policy in Western Europe and avoided costs of air pollution control". *Economic Modelling* 27.1, pp. 103–115.

Schumacher, Katja and Ronald D Sands (2007). "Where are the industrial technologies in energy–economy models? An innovative CGE approach for steel production in Germany". eng. *Energy economics*. Energy Economics 29.4, pp. 799–825. ISSN: 0140-9883.

Weitzel, Matthias, Bert Saveyn, and Toon Vandyck (2019). "Including bottom-up emission abatement technologies in a large-scale global economic model for policy assessments". *Energy Economics* 83.C, pp. 254–263.

Appendices

A Notation

In general, boldface letters denote vectors. A lower case version of a vector indicates a vector of ratios. The non-boldface version of a letter indicates a typical or specific element from the corresponding vector. Variables with bars above them refer to their counterpart calculated from data. Table 4 collects all symbols and what they represent.

Table 4: Summary of notation

Symbol	Variable description
S	The set of sectors
s	The index for sector
$\boldsymbol{X}$	The vector of non-capital goods
n	The number of non-capital goods, i.e. the length of $\boldsymbol{X}$
$\boldsymbol{K}$	A vector of durable capital goods
n_m	The number of emission types
$\boldsymbol{M_0}$	The vector of emissions before end-of-pipe abatement
$\boldsymbol{M}$	The vector of emissions after end-of-pipe abatement
$d\boldsymbol{X}$	The vector of non-capital goods (demand)
$s\boldsymbol{X}$	The vector of non-capital goods (supply)
$\boldsymbol{x} \cdot \boldsymbol{y}$	The scalar product of two vectors x and y
F_z	The production function for producing z.
$\mathcal{D}_Z(Y)$	The subset of descendants or ancestors of type Z to the node Y
$\boldsymbol{E}$	Intermediate goods in final production, 'energy services'
τ	A technology of either the input-displacing or the end-of-pipe type
κ_τ	Technology composite produced by τ
$\boldsymbol{U}^\tau$	Vector of technology goods produced by technology τ
F_τ^{in}	Production function to combine inputs into composite good κ_τ
F_τ^{o}	Production function to split composite good κ_τ into outputs $\boldsymbol{U}^\tau$
$\mathcal{T}^{\text{ID}}$	The set of input-displacing technologies
$\mathcal{T}^{\text{EOP}}$	The set of end-of-pipe technologies
$\boldsymbol{A}$	The vector of abatement
ϕ_X	Number of emissions generated by each unit used of input X
$\boldsymbol{C}$	Vector of components
p_X	The price of X
θ_C	End-of-pipe abatement cost parameter
$\tilde{G}(\cdot)$	Function governing the firm-specific element of end-of-pipe abatement cost
$EC(i)$	Total costs related to emissions and abatement for firm i
i	The index for firms
$\widehat{p}_X$	Input price corrected for abatement costs and emission taxes
$\mathcal{P}(\cdot)$	Scale-preserving aggregator taking prices as inputs and returns optimal shares $\frac{X^*}{Y^*} \in [0, 1]$
$\bar{c}_\tau$	Unit cost of technology τ stated in technology data
$\bar{x}(\tau)$	Non-capital input intensities of technology τ
R	Rental rate of capital
t	The index for time
δ	Depreciation rate of capital
λ	Shadow price for the law of motion for capital
Ψ^τ	Convex installation costs function
H_z^τ	Function determining the optimal use of inputs from prices

B Definitions

Definition B.1. *Neoclassical-type functions.*

i. Production function: We define a neoclassical-type production function as a mapping $F : \mathbb{R}^n_+ \to \mathbb{R}_+$ *for* $n \in \mathbb{N}$*, where* F *is (i) increasing and concave, (ii) twice continuously differentiable, and (iii) exhibits constant returns to scale.*

ii. Output-allocation function: We define a neoclassical-type output-allocation function as a convex version of the neoclassical-type production function.

Production functions are usually used with the notation $F = F^{in}$ to reflect that the function takes inputs (to produce some output). Output-allocation functions are usually used with the notation $F = F^o$ to reflect that the function takes a vector of outputs (which must equal some composite good).

Definition B.2. *Standard optimization problem.*

Let $F : \mathcal{X} \to \mathcal{Y}$ *denote a neoclassical-type function. We define the standard optimization problem* $P(F;\theta)$ *as the static profit maximization problem with parameters* $\boldsymbol{\theta} = (\theta^y, \boldsymbol{\theta}^x)$ *as one of two problems: (i)* $\max_{\boldsymbol{X}\in\mathcal{X}} \theta^y F(\boldsymbol{X}) - \boldsymbol{\theta}^x \cdot \boldsymbol{X}$ *when* $F(\boldsymbol{X})$ *is a neoclassical-type production function, or (ii)* $\max_{\boldsymbol{X}\in\mathcal{X}} \boldsymbol{\theta}^x \cdot \boldsymbol{X} - \theta^y F(\boldsymbol{X})$ *when* $F(\boldsymbol{X})$ *is a neoclassical-type output-allocation function.*

Definition B.3. *Scale-preserving functions.*

Let the pair (F, P) *denote a neoclassical-type function* $F : \mathcal{X} \to \mathcal{Y}$ *and an optimization problem* $P(F;\boldsymbol{\theta})$ *respectively. Let* $(\boldsymbol{X}^*, Y^*) \in (\mathcal{X}, \mathcal{Y})$ *denote the optimum for the problem* $P(F;\boldsymbol{\theta})$*.*

i. Scale-preserving function: We say that the function F *is a scale-preserving function with respect to* P *if* $\boldsymbol{x}^* \in \Delta^n$*, where* $x \equiv X/Y$ *and* Δ^n *denotes the unit-simplex.*

ii. Scale-preserving aggregator (for production functions): If F *is a scale-preserving neoclassical-type production function, we define the corresponding scale-preserving aggregator as the mapping* $\mathcal{P} : \mathbb{R}^{n+1} \to \Delta^n$ *from parameters* $\boldsymbol{\theta}$ *to the optimal shares* $\boldsymbol{x}^*$*.*

iii. Scale-preserving aggregator (for output-allocation functions): If F *is a scale-preserving neoclassical-type output-allocation function, we define the corresponding scale-preserving aggregator as the mapping* $\mathcal{P} : \mathbb{R}^{n+1} \to \Delta^n$ *from parameters* $\boldsymbol{\theta}$ *to the optimal shares* $\boldsymbol{x}^*$*.*

Lemma B.1. *Characteristics of scale-preserving functions.*

i. A scale-preserving aggregator (Definition B.3) with respect to a standard optimization problem (Definition B.2) exhibits the following characteristics:

$$\sum_{x\in\boldsymbol{x}} \mathcal{P}_x(\boldsymbol{\theta}) = 1, \qquad \frac{\partial \mathcal{P}_x}{\partial \theta^x} \leq 0, \qquad \mathcal{P}_x \geq 0. \tag{B.29}$$

ii. A scale-preserving output-allocation function (Definition B.3) with respect to a standard optimization problem (Definition B.2) exhibits the following characteristics:

$$\sum_{x\in\boldsymbol{x}} \mathcal{P}_x(\boldsymbol{\theta}) = 1, \qquad \frac{\partial \mathcal{P}_x}{\partial \theta^x} \geq 0, \qquad \mathcal{P}_x \geq 0. \tag{B.30}$$

In some cases, it will prove useful to go to directly specifying the scale-preserving aggregator ($\mathcal{P}$) instead of the underlying function F. Finally, for brevity, when referring to scale-preserving aggregators/functions, it will be with respect to a standard optimization problem unless otherwise specified; thus the '*with respect to*'-qualification is generally omitted.

Definition B.4. *Share parameters and smoothing parameters.*

Consider a standard optimization problem as in definition B.2 with neoclassical-type function F and parameters $\boldsymbol{\theta} \equiv (p_y, \boldsymbol{p}_x)$. Let $\mathcal{I}_x$ denote the interval of feasible values for $x \equiv X/Y$. Let $(\boldsymbol{\mu}_x^y, \sigma_y)$ denote parameters in F, and let $\mathcal{I}_\sigma$ denote the interval that σ_y belongs to. Let the infima of the intervals be denoted $\underline{I}_\sigma$ and $\underline{I}_x$, and similarly the suprema $\overline{I}_\sigma$ and $\overline{I}_x$.

i. μ_x^y is a share-parameter iff in optimum $x = \mu_x^y$ whenever $p_x = p_y$ for all $x \in \boldsymbol{x}$.

ii. σ_y is a smoothing-parameter iff in optimum

$$\lim_{\sigma_y \to \overline{I}_\sigma} (x) = \mu_x^y$$

and

- *If F is a production function:*

$$\lim_{\sigma_y \to \underline{I}_\sigma} (x) = \begin{cases} \underline{I}_x, & \text{if } p_x > p_y \\ \overline{I}_x, & \text{if } p_x \leq p_y \end{cases}$$

- *If F is an output-allocation function:*

$$\lim_{\sigma_y \to \underline{I}_\sigma} (x) = \begin{cases} \underline{I}_x, & \text{if } p_x < p_y \\ \overline{I}_x, & \text{if } p_x \geq p_y \end{cases}$$

We note that in many instances, it is convenient to have share-parameters as defined in definition B.4. However, share-parameters are often leveraged to target a baseline vector $\overline{\boldsymbol{x}}$ given a set of prices $\boldsymbol{p}_x, p_y$. Thus, to gauge how flexible and useful the share-parameters are, the relevant metric is rather the feasible values of $\boldsymbol{x}$ that can be induced by altering the set of share parameters. Define the parameter function $G : \mathcal{M} \to \mathcal{X}$ as the mapping of optimal shares $\boldsymbol{x}^* \in \mathcal{X}$ as a function of parameters $\boldsymbol{\mu}_x^y \in \mathcal{M}$. The relevant metric of how flexible the share-parameters are in the calibration procedure is the range of G ($= \mathcal{X}$).

Among the functions that naturally have both smoothing and share parameters are constant elasticity of substitution (CES) and constant elasticity of transformation (CET) functions (nor-

malized to be scale-preserving as in definition B.3 and regular), and various nested multinomial-logit-type functions (MNL, both output- and input type functions). A few examples of the demand functions for some input x derived from these:

$$x = \mu_x^y \left(\frac{p_x}{p_y}\right)^{-\frac{1}{\sigma_y}}, \qquad \mu_x^y > 0,\ \sigma_y > 0 \tag{CES}$$

$$x = \mu_x^y \left(\frac{p_x}{p_y}\right)^{\frac{1}{\sigma_y}}, \qquad \mu_x^y > 0,\ \sigma_y > 0 \tag{CET}$$

$$x = \frac{\mu_x^y \left(\frac{p_x}{p_y}\right)^{-\frac{1}{\sigma_y}}}{\sum_{x' \in \boldsymbol{x}} \mu_{x'}^y \left(\frac{p_{x'}}{p_y}\right)^{-\frac{1}{\sigma_y}}}, \qquad \boldsymbol{\mu}_x^y \in \Delta^n,\ \mu_x^y > 0,\ \sigma_y > 0, \tag{NCES}$$

$$x = \frac{\mu_x^y \left(\frac{p_x}{p_y}\right)^{\frac{1}{\sigma_y}}}{\sum_{x' \in \boldsymbol{x}} \mu_{x'}^y \left(\frac{p_{x'}}{p_y}\right)^{\frac{1}{\sigma_y}}}, \qquad \boldsymbol{\mu}_x^y \in \Delta^n,\ \mu_x^y > 0,\ \sigma_y > 0, \tag{NCET}$$

$$x = \frac{\mu_x^y \exp\left[(p_y - p_x)/\sigma_y\right]}{\sum_{x' \in \boldsymbol{x}} \mu_{x'}^y \exp\left[(p_y - p_{x'})/\sigma_y\right]}, \qquad \boldsymbol{\mu}_x^y \in \Delta^n,\ \mu_x^y > 0,\ \sigma_y > 0, \tag{MNL}$$

$$x = \frac{\mu_x^y \exp\left[(p_x - p_y)/\sigma_y\right]}{\sum_{x' \in \boldsymbol{x}} \mu_{x'}^y \exp\left[(p_{x'} - p_y)/\sigma_y\right]}, \qquad \boldsymbol{\mu}_x^y \in \Delta^n,\ \mu_x^y > 0,\ \sigma_y > 0, \tag{MNL-out}$$

It is straightforward to see that any nested MNL-like function can similarly be applied. Furthermore, we note that these are all examples that have very flexible share-parameters: In the CES and CET instances, we can basically target any value of x, independent of what the prices are, by adjusting μ_x^y appropriately. In the normalized functions (NCES,NCET,MNL,MNL-out), we can similarly target any vector $\boldsymbol{x} \in \Delta^n$ by appropriately adjusting $\boldsymbol{\mu}_x^y$, independent of what the prices may be.

B.1 A tale of N firms: Equivalence of multiple firms with identical production technologies and a single firm with multiple outputs

Consider $i \in I \equiv \{1, ..., N\}$ firms that produce different goods κ_i and employs identical production technologies *except* for a difference in total factor productivity $A_i > 0$:

$$\kappa_i = A_i F(\boldsymbol{X}), \qquad i \in I. \tag{B.31}$$

Assuming that F adheres to the assumptions in definition B.1 the firms face a standard static optimization problem as in definition B.2. Normalizing the problem by $\kappa_i A_i$ and letting $f(\boldsymbol{x}) \equiv F(\boldsymbol{X}/A_i\kappa_i)$ the solutions of the firms' problems are defined for each firm i:

$$X_i = H_x^f\left(\frac{p_X}{A_i p_i}\right)\kappa_i A_i, \qquad \forall X \in \boldsymbol{X} \tag{B.32}$$

$$A_i p_i = \sum_{X \in \boldsymbol{X}} H_x^f\left(\frac{p_X}{A_i p_i}\right) p_X, \tag{B.33}$$

where H denotes the inverse of the marginal product of f with respect to x. Note that as the production technology functions are identical the solutions coincide with (see Lemma B.2 below):

$$\begin{aligned}\frac{X_i}{\kappa_i A_i} &= \frac{X_j}{\kappa_j A_j} = x, \qquad \forall (i,j,X) \in (I \times I \times \boldsymbol{X}) \\ A_i p_i &= A_j p_j = p, \qquad \forall (i,j) \in I^2.\end{aligned}$$

Using this we have:

$$X \equiv \sum_i X_i = H_x^f \left(\frac{p_X}{p}\right) \kappa, \qquad \kappa \equiv \sum_i \kappa_i A_i, \qquad \forall X \in \boldsymbol{X} \tag{B.34}$$

$$p = \sum_{X \in \boldsymbol{X}} H_x^f \left(\frac{p_X}{p}\right) p_x \tag{B.35}$$

$$A_i p_i = p, \qquad \forall i \in I. \tag{B.36}$$

This is identical to the solution to an optimization problem of a single firm that produces the composite good κ using the production technology F, and then transforms κ into specific goods κ_i at linear rates of transformation defined by A_i.

Lemma B.2. *A small useful Lemma*

Consider the standard optimization problem (cf. definition B.2) with $F(\boldsymbol{X}) = Y$ denoting the neoclassical type production function and $\boldsymbol{p} = (p_y, \boldsymbol{p}_x)$ the price vectors. The first order conditions of an optimum requires:

$$\frac{\partial F}{\partial x}(\boldsymbol{x}) = \frac{p_x}{p_y}, \qquad \forall x \in \boldsymbol{x}, \tag{B.37}$$

where $x \equiv X/Y$ denotes the scaled input variable. This set of first order conditions identifies a unique vector $\boldsymbol{x}^*$ for a given vector of prices $\boldsymbol{p}$ (by monotonicity and concavity of F). Applying the envelope theorem, we can further add the average cost restriction (derivative of the expenditure function wrt. scale, Y):

$$p_y \equiv \boldsymbol{p}_x \cdot \boldsymbol{x}^*(\boldsymbol{p}). \tag{B.38}$$

The implication is that any vector of prices $\boldsymbol{p}_x$ maps to a unique combination of $(p_y, \boldsymbol{x})$ from first order conditions and the cost price index.

C Constructing and calibrating a specific tree

This appendix explains the general calibration procedure in detail and supplements it by constructing and calibrating an actual tree based on a very simply catalogue featuring just two input-displacing technologies, denoted τ_1 and τ_2. The simple catalogue is shown in table 5, and features the input intensities, the unit cost, the coverage potentials, overlap information as well as the

current applications. The overlap information states a list of other technologies that a technology shares a fraction of its potential with and states also the magnitude of this specific fraction. If more than one part of its potential is shared with other technologies, more pairs of information containing a list of technologies and a fraction is added. In the example shown in table 5, there is only one such pair: τ_1 shares 50 % of its potential with τ_2. Likewise, τ_2 shares 25 % of its potential with τ_1, such that the potential that they share for providing that E is 5 %.

Table 5: Simple technology catalogue example

τ	$\overline{x}_{elec}(\tau)$	$\overline{x}_{oil}(\tau)$	$\overline{c}_\tau$	$\overline{\theta}_\tau^{E_{heat}}$	Overlap share, E_{heat}	Overlapping τ	$\overline{\vartheta}_\tau^{E_{heat}}$
τ_1	0.75	0	700 \$	10 %	50 %	τ_2	3 %
τ_2	0	1.2	500 \$	20 %	25%	τ_1	4 %

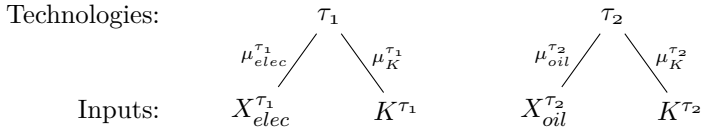
Note: The table is an example of a simple technology catalogue featuring just two input-displacing technologies.

In the construction and calibration of the tree, the first step is to construct the sets of input-displacing and end-of-pipe technologies. In the simple example from the catalogue in table 5, this means defining $\mathcal{T}^{\text{ID}} = \{\tau_1, \tau_2\}$.

Then for each $\tau \in \mathcal{T}^{\text{ID}} \cup \mathcal{T}^{\text{EOP}}$, the non-zero input intensities from the catalogue determine the branches for non-capital goods $d\boldsymbol{X}$ below each technology. In the simple example, this means we create a branch for $X_{elec}^{\tau_1}$ below τ_1 and a branch $X_{oil}^{\tau_2}$ below τ_2. We assume that all technologies require capital investments, so we include a branch for each $K \in \boldsymbol{K}$ underneath each τ. For the sake of simplicity, we assume for the simple catalogue example that only one type of capital K exists, so only one capital branch underneath τ_1 and τ_2 is created.[27] The simple tree is depicted in figure C.1.

In the general case, the input intensities from the catalogue, $\overline{x}(\tau)$, act as targets in calibration. Their model counterparts are the corresponding $X(\tau)/\kappa_\tau$ for all input and technology types. Therefore, the calibration procedure adjusts the share parameters μ_X^τ such that $\overline{x}(\tau) = X(\tau)/\kappa_\tau$. Table 6 summarizes targets, model counterparts and the adjusted parameters. For the simple catalogue, the calibration so far results in the tree illustrated in figure C.1 with parameters $\mu_{elec}^{\tau_1}$ and $\mu_{oil}^{\tau_2}$ calibrated.[28] Note that in the case where the production function of a technology is Leontief, the corresponding share parameters can be set directly to equal the input intensities stated in the catalogue, e.g. $\mu_{elec}^{\tau_1} = \overline{x}_{elec}(\tau_1) = 0.75$.

Figure C.1: A simple input-displacing subtree, part one



Note: The figure depicts the two input-displacing technologies from table 5 and the inputs they rely on.

The next information in the catalogue is the unit cost of each technology $\overline{c}_\tau$, see table 5. This

[27] The capital intensity will be calibrated, and could in principle be zero such that the capital branch drops out.
[28] Formally, we do not calibrate parameters sequentially, because calibration requires running the entire model. Hence, the ordering introduced here only serves the purpose of easing explanation.

Table 6: Calibration targets, model counterparts and parameters to adjust in calibration

Target				
Name	**Sym.**	**Model counterpart**	**Domain**	**Par.**
Potential (input-displacing)	$\bar{\theta}_C$	C^E / E	$C \in \boldsymbol{C}^{\text{ID}}$	μ_C^E
Potential (end-of-pipe)	$\bar{\theta}_C$	θ_C	$C \in \boldsymbol{C}^{\text{EOP}}$	θ_C
Unit cost	$\bar{c}_\tau$	p_τ	$\tau \in \{\mathcal{T}^{\text{ID}}, \mathcal{T}^{\text{EOP}}\}$	μ_K^τ
Input intensities	$\bar{x}(\tau)$	$X(\tau)/\kappa_\tau$	$X \in \boldsymbol{X}, \tau \in \{\mathcal{T}^{\text{ID}}, \mathcal{T}^{\text{EOP}}\}$	$\mu_X^{in,\tau}$
Input use for energy services	$\bar{Z}^E$,	$Z(\tau_0^E) + \sum_{\tau \in \mathcal{D}_\tau(E)} Z(\tau) s_\tau^E$	$Z \in \{\boldsymbol{X}, \boldsymbol{K}\}, E \in \boldsymbol{E}$	$\mu_X^{\tau_0^E}$
Energy service use	$\bar{E}$	E	$E \in \boldsymbol{E}$	$\mu_E^{in,Y}$
Final good production	$s\bar{X}$	sX	$X \in \boldsymbol{X}$	$\mu_X^{o,Y}$
Aggregate use of inputs	$\overline{Z}$	$\sum_\tau Z(\tau) + Z^{in,Y}$	$Z \in \{\boldsymbol{X}, \boldsymbol{K}\}$	$\mu_Z^{in,Y}$
Current coverage (input-displacing)	$\bar{\vartheta}_\tau^E$	$\sum_{U \in \mathcal{D}_{U^\tau}(E)} U / E$	$\tau \in \mathcal{T}^{\text{ID}}, E \in \boldsymbol{E}$	γ_C, μ_U^C
Current coverage (end-of-pipe)	$\bar{\vartheta}_\tau^M$	$\sum_{U \in \mathcal{D}_{U^\tau}(M)} U / M_0$	$\tau \in \mathcal{T}^{\text{EOP}}, M \in \boldsymbol{M}$	σ^G, μ^G

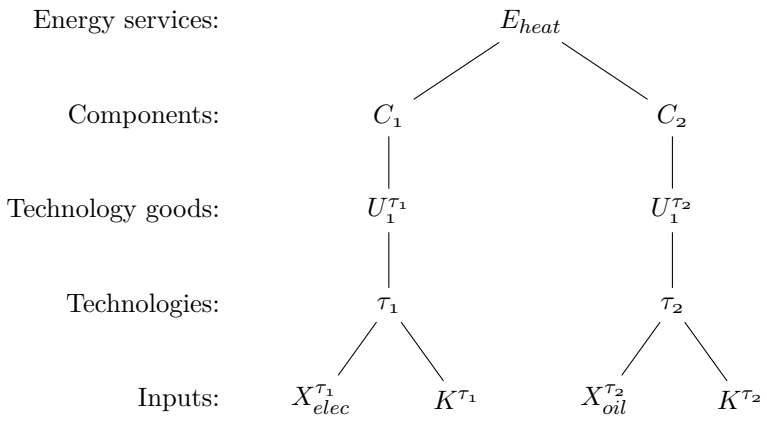
The variable s_τ^E identifies the share of outputs that technology τ contributes to energy service E relative to the technology's total output measured as $\sum_{U \in \mathcal{D}_{U^\tau}(E)} U/\kappa_\tau$.
For calibrating the current applications, we use minimization approaches since we have more parameters than targets. For input-displacing, see (22) and for end-of-pipe, see (26).

is also a target in calibration, the model counterpart is simply the price of the technology, p_τ, see table 6. Since this price is closely tied to the technology's inputs, we can set $p_\tau = \bar{c}_\tau$ by adjusting the capital intensity through the share parameters on capital types, μ_K^τ for $K \in \boldsymbol{K}$. When there are multiple types of capital, we assume that their relative intensities are identified in some other way. Fitting unit costs by adjusting the capital intensity effectively assumes that if e.g. two technologies have different unit costs, but the same input intensities $\overline{\boldsymbol{x}}(\tau)$, it must be that the costlier technology is more capital intensive. This is why we do not need data on capital intensities; they are identified from unit costs. Note that the unit cost stated in the catalogue determines the cost of a technology at a given set of input prices. For the simple example with just two technologies, this procedure identifies the remaining two parameters in figure C.1, $\mu_K^{\tau_1}$ and $\mu_K^{\tau_1}$.

The coverage potentials, $\overline{\theta}_\tau^E$ or $\overline{\theta}_\tau^M$, measure the share of a particular E that a technology can cover or the share of a particular M_0 a technology can abate, if fully adopted. The next step in constructing the tree is, for each $\tau \in \mathcal{T}^{\text{ID}}$, to create a technology good U^τ for each of the energy services $E \in \boldsymbol{E}$ where τ has non-zero coverage potential. For the simple example catalogue, this means adding $U_1^{\tau_1}$ above τ_1 and $U_1^{\tau_2}$ above τ_2. No more technology goods would have to be constructed if these two technologies did not overlap. If this was the case (and we ignore baseline

technologies for now), the resulting tree would be as depicted in figure C.2: Each of the two technology goods would connect to separate components, C_1 and C_2, which then connect to E_{heat}, the only energy service.

Figure C.2: A simple input-displacing subtree, part two



Note: Relative to figure C.1, this adds technology goods, components and energy services.

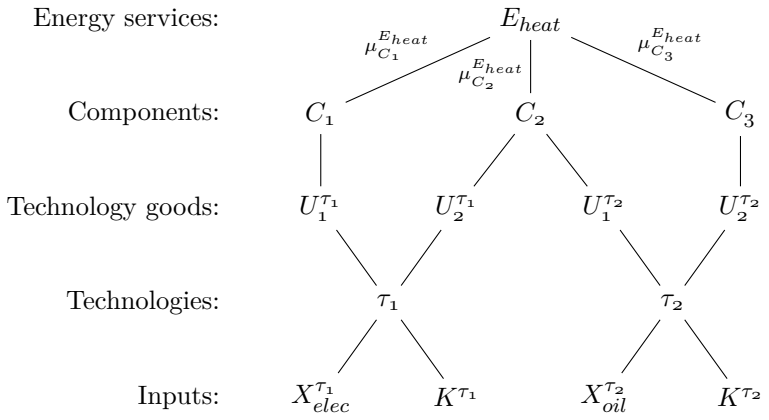
When technologies overlap, the structure of the tree must reflect it. If a technology's potential within some E or M overlaps with other technologies, in the sense that it has non-zero overlap shares stated in the technology catalogue, the number of technology goods is increased. Overlap shares divide a technology's potential into mutually exclusive parts, and the technology produces one technology good for each of these parts.[29] All sets of overlapping technology goods are then connected to separate components which are finally connected to the relevant E or M.

In the simple catalogue, τ_1 has a coverage potential of E_{heat} of 10 %, but 50 percent of that potential overlaps with τ_2. Since τ_2 itself has a coverage potential for E_{heat} of 20 %, its overlap share with τ_1 must be 25 %, such that the overlapping potential between them is 5 %. In this way, the overlap information results in potentials distributed between components; $\overline{\theta}_{C_1} = 0.05$, $\overline{\theta}_{C_2} = 0.05$ and $\overline{\theta}_{C_3} = 0.15$, two of them without overlap, i.e. where only one explicit technology contributes to it, and one where the two technologies compete. The overlap information means τ_1 must produce two technology goods, $U_1^{\tau_1}$ and $U_2^{\tau_1}$, and τ_2 must similarly produce two technology goods. Due to overlap, $U_2^{\tau_1}$ and $U_1^{\tau_2}$ compete to provide C_2. The resulting tree is shown in figure C.3.

To reach a technology's full coverage potential in some component C, the corresponding technology good U^τ must be fully adopted. For input-displacing technologies, a particular technology good U^τ is fully adopted when it is the sole contributor to the component C above it. To see this, note that when this is the case, we have $U^\tau = C$ (remember that the aggregation from technology goods to components is scale-preserving). When this holds, we have $U^\tau/E = C^E/E$. Calibration must therefore ensure that $C^E/E = \overline{\theta}_C$ since this means that when the technology good U^τ is

[29] If the overlap shares sum to 1, the number of technology goods for τ under a particular E will be equal to the number of overlap parts. If they sum to less than one, the number of technology goods will be equal to the number of overlap parts plus one, where the plus one reflects the part of potential which does not overlap with other technologies.

Figure C.3: A simple input-displacing subtree, part three



Note: Relative to figure C.2, this figure adds additional technology goods and a component, to incorporate overlap between the two technologies.

fully adopted, its share of E equals $\overline{\theta}_C$. The calibration target $\overline{\theta}_C$ therefore has the model counterpart C^E/E, see table 6. The parameters to adjust in calibration are the share parameters of the production functions F_E^{in} for each E, denoted μ_C^E. In our simple tree with two input-displacing technologies, targeting the coverage potentials therefore identifies the share parameters $\mu_{C_1}^{E_{heat}}$, $\mu_{C_2}^{E_{heat}}$ and $\mu_{C_3}^{E_{heat}}$, see figure C.3.

For end-of-pipe technologies, the calibration procedure for coverage potentials is slightly different. For a given component C and a technology U^τ contributing to it, the calibration target is still $\overline{\theta}_C$ and U^τ still equals C when fully adopted. Recall also that $C = M_0\theta_C G(p_M - p_C)$ where $G(\cdot)$ measures the degree to which C is phased in. When it is fully phased in, i.e. when the coverage potential of the component is fully exhausted, we have $G(\cdot) = 1$, so that $\frac{C}{M_0} = \theta_C$. Since $U^\tau = C$ under full adoption, the calibration target $\overline{\theta}_C$, simply has the model counterpart θ_C. Hence, calibration simply requires setting $\theta_C = \overline{\theta}_C$, see table 6.

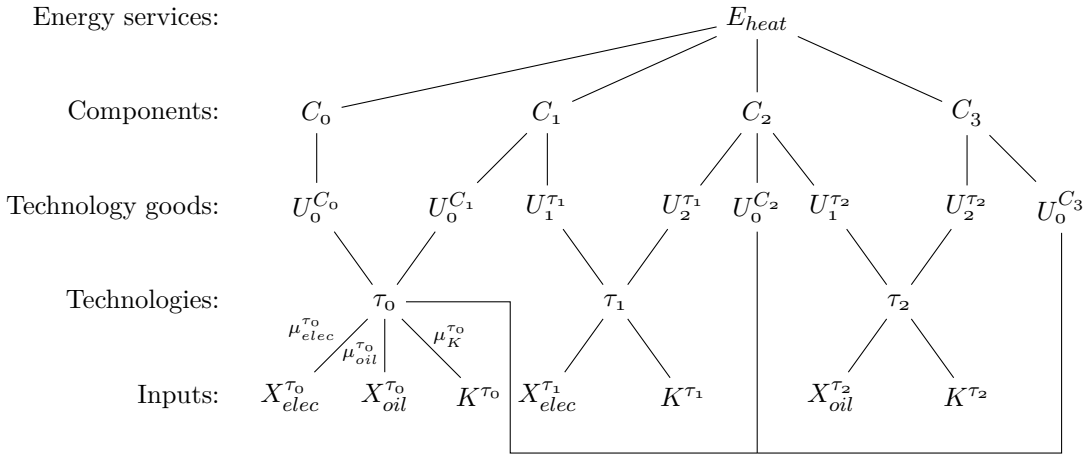
Note that a technology that has no clear upper boundary on its coverage potential for some E or M simply implies that $\overline{\theta}_\tau^E = 1$ or $\overline{\theta}_\tau^M = 1$, i.e. that the technology has the potential to fully cover energy demand for E or fully abate all emissions of M. Note also that we do not offer a method to impose a coverage potential formulated in absolute quantities. To use technologies formulated in this way, one needs to divide this reduction potential by E or M_0, and accept the implication that the coverage potential in absolute terms can change with the sizes of E or M_0.

The final step of the tree construction is to add baseline technologies in the input-displacing subtree (the end-of-pipe subtree has no baseline technologies). There is one baseline technology, τ_0^E, for each energy service. As noted in the main text, input-displacing technology catalogues are rarely completely exhaustive; they do not realistically list all existing technologies and so there will be coverage potential for some E unaccounted for in the catalogue. In other words, the sum of $\overline{\theta}_C$ across all $C \in \boldsymbol{C}^E$ does not equal 1. To account for this, we introduce a baseline component C_0^E to each E with potential equal to the difference between 1 and the measured potentials, $1 - \sum_{C \in \boldsymbol{C}^E} \overline{\theta}_C$ and calibrate the corresponding share parameter $\mu_{C_0}^E$ as explained above.

The baseline technology τ_0^E produces the technology good $U_0^{C_0}$ which by construction is the only technology good contributing to C_0, so due to scale preservation, $U_0^{C_0}$ always equals C_0. For our simple catalogue example, this means, apart from introducing the baseline technology $\tau_0^{E_{heat}}$ itself, adding the baseline component $C_0^{E_{heat}}$ and its baseline technology good $U_0^{C_0}$.

The baseline technology also serves the purpose of competing against the input-displacing technologies from the catalogue. To understand the need for this, consider again our simple example, and specifically the node C_1 in figure C.3. As the tree is drawn there, and with the assumption of scale preservation, we always have $U_1^{\tau_1} = C_1$ such that this technology good by construction is always fully adopted. To avoid that, we add a baseline technology good $U_0^{C_1}$ against which the technology good $U_1^{\tau_1}$ competes to provide C_1. With similar reasoning, we add baseline technology goods to all other components allowing them to be less than fully adopted. In the simple catalogue example, adding the baseline technology and its inputs, the baseline component and the vector of baseline technology goods results in the tree illustrated in figure C.4.

Figure C.4: A simple input-displacing subtree, part four



Note: Relative to figure C.3, this figure adds the baseline technology, its technology goods and the baseline component.

Apart from representing the status quo which technologies from the catalogue compete against, baseline technologies serve another important purpose. They allow our model to reproduce national accounts data for the industry. We assume that data on energy service use, $\overline{\boldsymbol{E}}$, final goods production $s\overline{\boldsymbol{X}}$, input use across different energy services, $d\overline{\boldsymbol{X}}^E$ and $\overline{\boldsymbol{K}}^E$, and finally total input use for the industry, $d\overline{\boldsymbol{X}}$ and $\overline{\boldsymbol{K}}$, is available.[30]

The model counterpart for the input use across different energy services, $d\overline{\boldsymbol{X}}^E$, consists of two parts; the input use of the baseline technology under that energy service $d\boldsymbol{X}(\tau_0^E)$ and the input use for other technologies under that E. Because one technology τ can contribute to multiple $E \in \boldsymbol{E}$ at the same time however, for calibration purposes, we distribute its input use across the relevant E. We do so by calculating the share of a technology's output ($\boldsymbol{U}^\tau$) that contributes to different E, i.e. we measure $\sum_{U \in \mathcal{D}_{U^\tau}(E)} U/\kappa_\tau$ for each E. These shares sum to one by construction due to the assumption of scale preservation. Using these shares to distribute input use across different

[30] See appendix D for suggested solutions when data varies from this.

E, calibration must ensure that the model counterpart is fitted to the target:

$$d\overline{X}^E = X(\tau_o^E) + \sum_{\tau \in \mathcal{D}_\tau(E)} \sum_{U \in \mathcal{D}_{U^\tau}(E)} \frac{U}{\kappa_\tau} dX(\tau), \qquad \forall X \in \boldsymbol{X}$$

To hit this target, and the corresponding ones for capital use $\overline{\boldsymbol{K}}^E$, we adjust the share parameters of the production function for the baseline technology τ_o^E e.g. $\mu_X^{\tau_o}$ and $\mu_K^{\tau_o}$, see table 6. For our simple example catalogue in figure C.4, these input use targets identify the share parameters $\mu_{elec}^{\tau_o}$, $\mu_{oil}^{\tau_o}$ and $\mu_K^{\tau_o}$.

The next two targets, final goods production $s\overline{\boldsymbol{X}}$ and energy service use $\overline{\boldsymbol{E}}$, requires extending the tree to include the production of final goods from the composite Y which is produced by energy services $\boldsymbol{E}$ and inputs $d\boldsymbol{X}$. The resulting tree structure for the simple catalogue example is illustrated in figure C.5, where we assume that the industry produces two outputs, white and grey cement.

The targets of $s\overline{\boldsymbol{X}}$ and $\overline{\boldsymbol{E}}$ have straightforward model counterparts in $s\boldsymbol{X}$ and $\boldsymbol{E}$ respectively. To fit the final goods production to its target, we adjust the share parameters of the output-allocation function F_Y^o, e.g. $\mu_X^{o,Y}$. To fit energy service use to its target, we adjust the share parameters of the production function F_Y^{in}, e.g. $\mu_E^{in,Y}$, see table 6. In the simple catalogue example, $\overline{\boldsymbol{E}}$ provides identification of $\mu_{E_{heat}}^{in,Y}$ and $s\overline{\boldsymbol{X}}$ provides identification of $\mu_{white}^{o,Y}$ and $\mu_{grey}^{o,Y}$, see figure C.5.

The last target from national accounts data is the sector's total use of inputs. The model counterpart for this is the sum of two parts: Input use for technologies, baseline, input-displacing and end-of-pipe, and input use directly used for final goods production. Therefore, calibration must ensure

$$d\overline{X} = \sum_\tau dX(\tau) + dX^{in,Y} \qquad \forall X \in \boldsymbol{X} \tag{C.39}$$

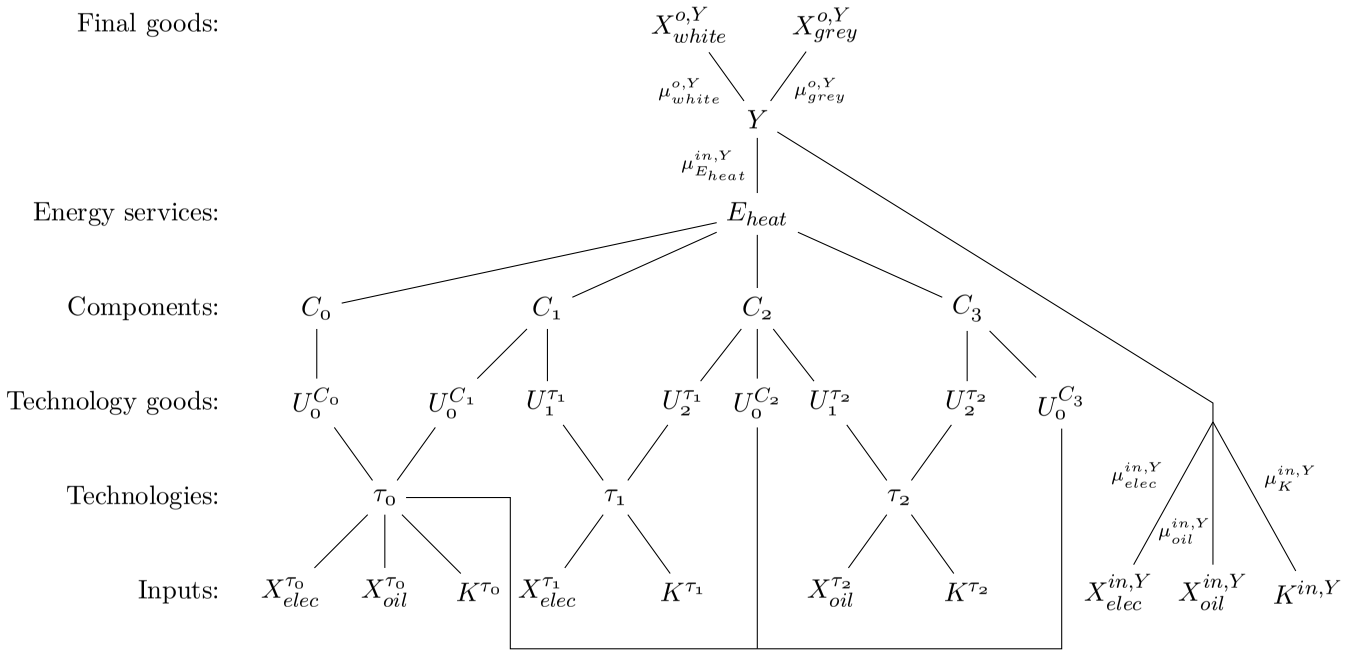
and likewise for all capital types. To fit the model counterpart to its target, we adjust the share parameters for inputs in the final goods production function F_Y^{in}, e.g. $\mu_X^{in,Y}$. In the simple catalogue example which features no end-of-pipe technologies, this means $d\overline{\boldsymbol{X}}$ identifies $\mu_{elec}^{in,Y}$, $\mu_{oil}^{in,Y}$ and $\mu_K^{in,Y}$, see figure C.5.

The last set of targets are the current applications, $\overline{\vartheta}_\tau^E$ and $\overline{\vartheta}_\tau^M$, i.e. the extent to which the potentials across different E or M_o are exhausted in some base year. For the simple catalogue example in table 5, the current application for τ_1 in E_{heat} is 3%. Using the tree in figure C.5, calibration must ensure $\frac{U_1^{\tau_1}+U_2^{\tau_1}}{E_{heat}} = 0.03$ at base year prices.

Section 4.1 of the main text explains the calibration methodology that we use to fit current applications to their targets. For input-displacing technologies, we rely on two sets of parameters: The share parameters in the aggregator $\mathcal{P}^C$ and the rates of transformation for baseline technologies $(\boldsymbol{\gamma}_{U^{\tau_o}})$.

For the simple catalogue example, this set of parameters consists of four rates of transformation, γ_{C_o}, γ_{C_1}, γ_{C_2} and γ_{C_3}, and eight share parameters. Since γ_{C_o} is not identified it is simply set to 1, so only three transformation rates can be adjusted in calibration. Since all production functions

Figure C.5: A simple input-displacing subtree, part five



Note: Relative to figure C.4, this figure adds final goods production including the input use outside of energy service production.

for component firms are scale preserving, only $n-1$ share parameters per F_C^{in}, where n is the length of $\mathcal{D}_U(C)$, are free. Therefore, we simply normalize by setting all share parameters related to baseline technology goods equal to 1. For the simple example catalogue, this results in figure C.6, where four remaining share parameters are free to be adjusted in calibration.

By minimizing the criterion function of (22) using the seven free parameters, the current applications of 3 % and 4 % for τ_1 and τ_2 are targeted. Since the simple catalog here does not include any end-of-pipe technologies, we need not solve the minimization problem related hereto. This concludes the calibration procedure for the catalog shown in table 5.

D Variations in available data

The data available for calibration can vary from what is assumed in section 4. This appendix covers a variety of such cases.

D.1 Missing data on overlap

Without information on technologies' overlaps, assumptions have to be made depending on the application. If the technologies represented in the catalog are sufficiently different one could assume zero overlap. This will not always be possible, since the technologies' combined potential could reach a value above 1 which is not possible. In other applications, assuming full overlap between all technologies under each E or M could also be appropriate. The better choice will depend on the setting and the catalogue used.

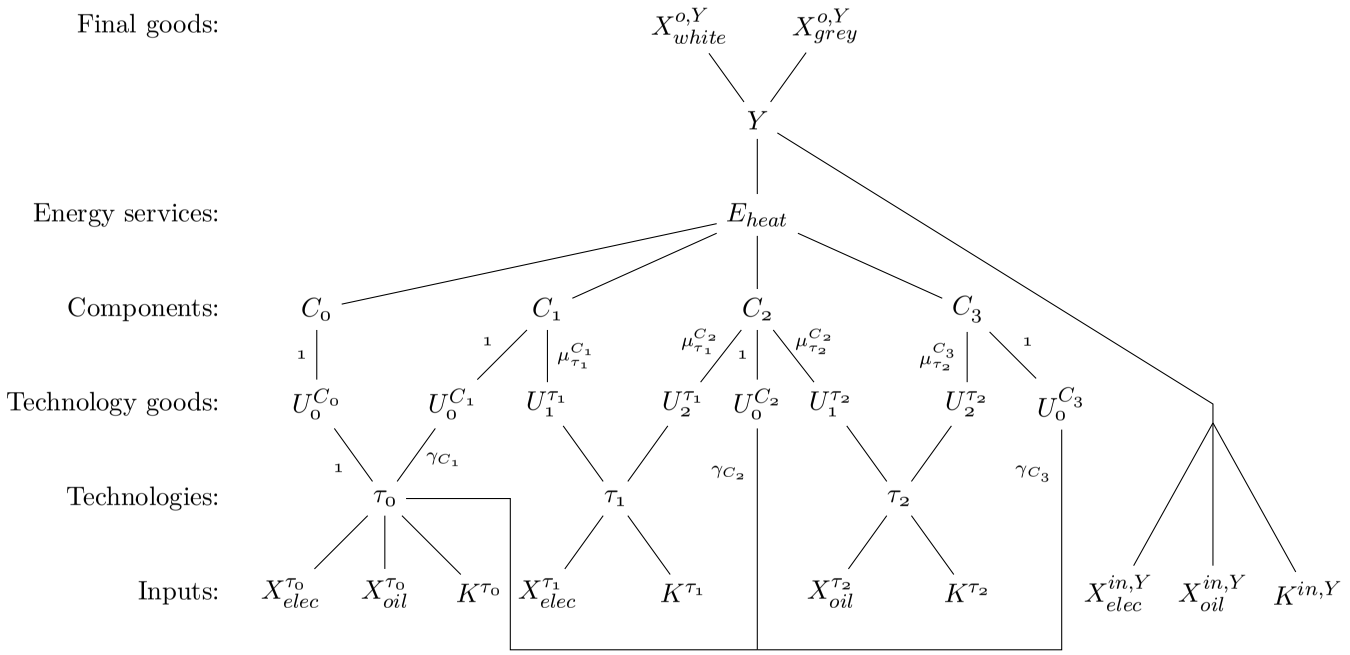
D.2 Less detailed input-output data

The most common case of less detailed input-output data is one where input uses underneath energy services are not known, but only the total input uses in the sector, $\overline{\boldsymbol{X}}, \overline{\boldsymbol{K}}$ are known. This means we cannot calibrate separate input structures for baseline firms across different E nor for the part of final good production outside energy services. Instead, we assume that only one baseline firm for the entire industry exists, and that this produces all the baseline technology goods in the model. The input use outside of energy services is also replaced by a technology good produced by the single baseline firm. Then, the total input use in the industry is fit by calibrating the share parameters in the production function for the baseline firm.

D.3 Data on competing technology, but not current application

If we do not have information on the current applications of the input-displacing technologies, it suffices to know the average cost of the main competing technology. In this case, we can calibrate the rates of transformation (γ_U) to target the price $p_u = \gamma_U p_{\tau_0^E}$ for the baseline firm, and let current application be endogenous. In the case where even the average cost of the main competing

Figure C.6: A simple input-displacing subtree, part six



Note: Relative to figure C.5, this figure does not add additional branches, but highlights the parameters relevant for calibration of current applications.

technology is also unknown, a solution could be to simply set all linear rates of transformation equal to 1 and keep current applications endogenous.

D.4 Data on input-displacing technologies arranged as MAC curve

If the data on input-displacing technologies is arranged as a marginal abatement cost curve, that is, with potentials $\overline{\theta}_{\tau}^{M}$ measuring a fraction of emissions M_{o} rather than a fraction of the relevant E and the unit cost c_{τ}^{M} measured as 'per unit of emission abated' rather than 'per unit of energy service produced', a conversion of these values is necessary. To convert potentials, one calculates first the pre-abatement emissions generated by production of E in the base year by using emission coefficients ϕ_{X}^{M} as well as input use under that energy service, $d\overline{\boldsymbol{X}}^{E}$. From this measure of emissions, M_{o}^{E}, the correct potential, $\overline{\theta}_{\tau}^{E}$, must satisfy the equation

$$M_{\mathrm{o}}\overline{\theta}_{\tau}^{M} = M_{\mathrm{o}}^{E}\overline{\theta}_{\tau}^{E} \quad \Leftrightarrow \quad \overline{\theta}_{\tau}^{E} = \frac{M_{\mathrm{o}}}{M_{\mathrm{o}}^{E}}\overline{\theta}_{\tau}^{M}. \tag{D.40}$$

This says is that if a full adoption of an input-displacing technology will remove 25 % of emissions and it contributes to an E that generates 50 % of emissions, its coverage potential measured as a share of that E must be 50 %. In cases where the technology contributes to more than one E, the potential to reduce emissions must be split between them.

Similarly, the unit cost c_{τ}^{M} must be converted to 'cost per unit of energy service produced', c_{τ}^{E}. Hence, if E measures spatial heating in GJs, the unit cost of an input-displacing technology contributing to spatial heating must be measured in USD per GJ. To perform this conversion, note that we have an estimate of the total cost of fully adopting the technology from the marginal abatement cost curve data:

$$TC^{M} \equiv c_{\tau}^{M} M_{\mathrm{o}}\overline{\theta}_{\tau}^{M}, \tag{D.41}$$

i.e. the product of the cost per unit of emissions and the total amount of emissions that the technology can abate (measured in the base year where M_{o} is known). Any conversion of costs from per M terms to per E terms must preseve the total cost i.e. the unit cost $\overline{c}_{\tau}^{E}$ must satisfy

$$TC^{M} = TC^{E} \equiv c_{\tau}^{E} E\overline{\theta}_{\tau}^{E} \quad \Leftrightarrow \quad \overline{c}_{\tau}^{E} = \frac{TC^{M}}{E\overline{\theta}_{\tau}^{E}}, \tag{D.42}$$

which we use to calculate the parameter $\overline{c}_{\tau}^{E}$. In cases whre a technology contributes to only one E, we can simply define this as the unit cost of the technology, $\overline{c}_{\tau}$.

D.5 Data on end-of-pipe technologies' potential *beyond* current application

If we do not have current application for end-of-pipe technologies and the potential for these technologies describe how much can be added on top of the (unknown) current application: Let current application be endogenous. Note that current application of a component is defined as the

share $\theta_C G(p_M - p_C)$. Thus, let $\overline{\theta}_C$ define the extra potential that a technology can abate. The level of θ_C is then defined as $\theta_C = \overline{\theta_C}/(1 - G(p_M - p_C))$. Otherwise, things can work as usual. Note that if the technology catalogue has unrealistically low costs, the potential for technologies becomes unrealistically large - potentially larger than 1.

E Simplifying assumptions in Schumacher and Sands (2007)

Our framework can reproduce the model from Schumacher and Sands (2007) with appropriate simplifications and functional form choices. Their catalogue features five input-displacing technologies which all produce crude steel, which then becomes the only element of the intermediate good vector $\boldsymbol{E}$. The catalogue is assumed exhaustive, so we exclude baseline technologies. Furthermore, the industry has no input use outside the production of $\boldsymbol{E}$, so we have by construction that $\boldsymbol{E} = E = Y = sX$ which are all scalars referring to crude steel production.

All five input-displacing technologies are assumed to be Leontief. Four technologies overlap 100 percent pairwise, resulting in three components (two with overlaps, one without). With one technology good for each technology, we have $\kappa_\tau = U^\tau$ for each technology. The functional forms for the aggregators that combine technology goods $\boldsymbol{U}$ into components $\boldsymbol{C}$ and these components into intermediate goods $\boldsymbol{E}$ are both assumed to be of the standard logit form (which is scale-preserving by construction), with more substitution in the former and less in the latter.

Their calibration procedure adjusts share parameters in the Leontief production functions of the technologies to match data on input use. The share parameters in the aggregation from $\boldsymbol{U}$ to $\boldsymbol{E}$ are adjusted to fit base year production of each technology.

F The full system of equations

F.1 End-of-pipe abatement module

This appendix summarizes the relevant equations and variables to solve for, when applying the end-of-pipe module as outlined in sections 3.1-3.2.

The general case: In the most general case the relevant price indices are given by:

$$\widehat{p}_M = p_M \left(1 - \sum_{C \in \mathcal{D}_C(A_M)} c_M\right) + \sum_{C \in \mathcal{D}_C(A_M)} c_M \left(p_C + \mathbb{E}_G\left[\chi | \chi < p_M - p_C\right]\right), \qquad \forall M \in \boldsymbol{M} \tag{F.43a}$$

$$\text{where} \quad c_M = \theta_C G\left(p_M - p_C\right),$$

$$p_C = \sum_{U \in \boldsymbol{U}^C} \mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) p_U, \qquad \forall C \in \boldsymbol{C}^M,\ \forall M \in \boldsymbol{M} \tag{F.43b}$$

$$p_\tau = \sum_{U \in \boldsymbol{U}^\tau} \mathcal{P}_U^\tau\left(p_\tau, \boldsymbol{p}_{U^\tau}\right) p_U, \qquad \forall \tau \in \mathcal{T}^{\text{EOP}} \tag{F.43c}$$

$$p_\tau = \sum_{Z\in\{d\boldsymbol{X}(\tau),\boldsymbol{K}(\tau)\}} H_z^\tau\left(\frac{\hat{p}_Z}{p_\tau}\right)\hat{p}_Z, \qquad \forall\tau\in\mathcal{T}^{\text{EOP}} \tag{F.43d}$$

$$\hat{p}_Z = p_Z + \sum_{M\in\boldsymbol{M}} \phi_Z^M \hat{p}_M, \qquad \forall Z\in\{d\boldsymbol{X},\boldsymbol{K}\}, \tag{F.43e}$$

and where input prices (p_z) and emission prices (p_M) are considered exogenous in the partial equilibrium. In the general case, this system of price equations cannot be decoupled from quantities.[31] Thus we add the quantity equations:

$$U = \mathcal{P}_U^C\left(p_C,\boldsymbol{p}_U\right)C, \qquad \forall U\in\boldsymbol{U}^C,\ \forall C\in\boldsymbol{C}^M,\ \forall M\in\boldsymbol{M} \tag{F.43f}$$

$$U^\tau = \mathcal{P}_U^\tau\left(p_\tau,\boldsymbol{p}_{U^\tau}\right)\kappa_\tau, \qquad \forall U^\tau\in\boldsymbol{U}^\tau,\ \forall\tau\in\mathcal{T}^{\text{EOP}} \tag{F.43g}$$

$$C = M_{\text{o}} c_M, \qquad \forall C\in\boldsymbol{C}^M,\ \forall M\in\boldsymbol{M}, \tag{F.43h}$$

$$Z = H_z^\tau\left(\frac{\hat{p}_Z(\tau)}{p_\tau}\right)\kappa_\tau, \qquad \forall Z\in\{d\boldsymbol{X}(\tau),\boldsymbol{K}(\tau)\},\ \forall\tau\in\mathcal{T}^{\text{EOP}}, \tag{F.43i}$$

$$M_{\text{o}} = \bar{M}_{\text{o}} + \sum_{\tau\in\mathcal{T}^{\text{EOP}}}\sum_{Z\in\{d\boldsymbol{X},\boldsymbol{K}\}} Z(\tau)\phi_Z^M, \qquad \forall M\in\boldsymbol{M}, \tag{F.43j}$$

where $\bar{M}_{\text{o}}$ are defined as emissions before abatement for all other firms than end-of-pipe technology firms. These are determined outside the end-of-pipe abatement module. We say that the system of equations in (F.43) is square in the sense that given $(\bar{M}_{\text{o}}, p_X, p_K, p_M)$ and parameters there is the same number of endogenous variables as there is equations. The endogenous variables are: $(\hat{p}_M, M_{\text{o}}, p_C, C, p_\tau, \kappa_\tau, p_{U^\tau}, U^\tau, \hat{p}_X, dX(\tau), \hat{p}_K, K(\tau))$, so the number of endogenous variables and equations are given by:

$$\#Equations = 2(n_M + n_{\mathcal{T}^{\text{EOP}}} + n_U + n_C + n_X + n_K).$$

By combining equations, the quantity system can be reduced:

$$\mathcal{P}_U^\tau\left(p_\tau,\boldsymbol{p}_{U^\tau}\right) = \mathcal{P}_U^C\left(p_C,\boldsymbol{p}_U\right)\frac{M_{\text{o}} c_M}{\kappa_\tau} \qquad \forall U^\tau\in\boldsymbol{U}^\tau,\ \forall\tau\in\mathcal{T}^{\text{EOP}} \tag{F.44a}$$

$$Z = H_z^\tau\left(\frac{\hat{p}_Z(\tau)}{p_\tau}\right)\kappa_\tau, \qquad \forall Z\in\{d\boldsymbol{X}(\tau),\boldsymbol{K}(\tau)\},\ \forall\tau\in\mathcal{T}^{\text{EOP}}, \tag{F.44b}$$

$$M_{\text{o}} = \bar{M}_{\text{o}} + \sum_{\tau\in\mathcal{T}^{\text{EOP}}}\sum_{Z\in\{d\boldsymbol{X},\boldsymbol{K}\}} Z(\tau)\phi_Z^M, \quad \forall M\in\boldsymbol{M}, \tag{F.44c}$$

thus dropping the variables (U, C). The number of equations is then reduced to $2(n_M + n_{\mathcal{T}^{\text{EOP}}} + n_X + n_K) + n_U + n_C$.

F.2 Input-displacing technology module

This appendix summarizes the relevant equations and variables when applying the input-displacing technology module, including final goods production, as outlined in sections 3.3-3.4.

[31] For each $\tau\in\mathcal{T}^{\text{EOP}}$ we need $n_{U^\tau}-1$ additional equations for the system to be square.

The general case

In the most general case the system of equations is given by:

$$p_Y = \sum_{x \in s\boldsymbol{X}} H_x^{Y,o}\left(\frac{p_X}{p_Y}\right) p_X \tag{F.45a}$$

$$p_Y = \sum_{Z \in \{\boldsymbol{E}, d\boldsymbol{X}, \boldsymbol{K}\}} H_z^{Y,in}\left(\frac{\hat{p}_Z}{p_Y}\right) \hat{p}_Z \tag{F.45b}$$

$$X = H_x^{Y,o}\left(\frac{p_X}{p_Y}\right) Y, \qquad \forall X \in s\boldsymbol{X} \tag{F.45c}$$

$$Z = H_z^{Y,in}\left(\frac{\hat{p}_Z}{p_Y}\right) Y, \qquad \forall Z \in \{\boldsymbol{E}, d\boldsymbol{X}, \boldsymbol{K}\} \tag{F.45d}$$

$$p_E = \sum_{C \in \boldsymbol{C}^E} \mathcal{P}_C^E\left(p_E, \boldsymbol{p}_{C^E}\right) p_C, \qquad \forall E \in \boldsymbol{E} \tag{F.45e}$$

$$C = \mathcal{P}_C^E\left(p_E, \boldsymbol{p}_{C^E}\right) E, \qquad \forall C \in \boldsymbol{C}^E,\ \forall E \in \boldsymbol{E} \tag{F.45f}$$

$$p_C = \sum_{U \in \boldsymbol{U}^C} \mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) p_U, \qquad \forall C \in \boldsymbol{C}^E,\ \forall E \in \boldsymbol{E} \tag{F.45g}$$

$$U = \mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) C, \qquad \forall U \in \boldsymbol{U}^C,\ \forall C \in \boldsymbol{C}^E,\ \forall E \in \boldsymbol{E} \tag{F.45h}$$

$$p_\tau = \sum_{Z \in \{\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\}} H_z^\tau\left(\frac{\hat{p}_Z}{p_\tau}\right) \hat{p}_Z, \qquad \forall \tau \in \{\mathcal{T}^{\mathrm{ID}}, \mathcal{T}_{\mathrm{o}}\} \tag{F.45i}$$

$$p_\tau = \sum_{U \in \boldsymbol{U}^\tau} \mathcal{P}_U^\tau\left(p_\tau, \boldsymbol{p}_{U^\tau}\right) p_U, \qquad \forall \tau \in \mathcal{T}^{\mathrm{ID}} \tag{F.45j}$$

$$U^\tau = \mathcal{P}_U^\tau\left(p_\tau, \boldsymbol{p}_{U^\tau}\right) \kappa_\tau, \qquad \forall U^\tau \in \boldsymbol{U}^\tau,\ \forall \tau \in \mathcal{T}^{\mathrm{ID}} \tag{F.45k}$$

$$Z = H_z^\tau\left(\frac{\hat{p}_Z}{p_\tau}\right) \kappa_\tau, \qquad \forall Z \in \{d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\},\ \forall \tau \in \{\mathcal{T}^{\mathrm{ID}}, \mathcal{T}_{\mathrm{o}}\} \tag{F.45l}$$

$$\kappa_\tau = \sum_{U \in \boldsymbol{U}^\tau} U / \gamma_U, \qquad \forall \tau \in \mathcal{T}_{\mathrm{o}} \tag{F.45m}$$

$$p_U = p_\tau \gamma_U, \qquad \forall U \in \boldsymbol{U}^\tau, \forall \tau \in \mathcal{T}_{\mathrm{o}} \tag{F.45n}$$

where $\mathcal{T}_{\mathrm{o}}$ denotes the set of baseline technologies.

To reduce this system of equations, note that for any non-baseline technology $\tau \in \mathcal{T}^{\mathrm{ID}}$, we can combine equations to write

$$\kappa_\tau = \frac{U}{\mathcal{P}_U^\tau\left(p_\tau, \boldsymbol{p}_{U^\tau}\right)} = \frac{\mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) C}{\mathcal{P}_U^\tau\left(p_\tau, \boldsymbol{p}_{U^\tau}\right)} = \frac{\mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) \mathcal{P}_C^E\left(p_E, \boldsymbol{p}_{C^E}\right)}{\mathcal{P}_U^\tau\left(p_\tau, \boldsymbol{p}_{U^\tau}\right)} E. \tag{F.46}$$

Note that equation F.46 has to hold for every technology good U^τ produced by input-displacing technologies, i.e. it determines the price index p_{U^τ}. For baseline technologies $\tau \in \mathcal{T}_{\mathrm{o}}$, we can likewise write

$$\kappa_\tau = \sum_{U \in \boldsymbol{U}^\tau} \frac{U}{\gamma_U} = \sum_{U \in \boldsymbol{U}^\tau} \frac{\mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) C}{\gamma_U} = \sum_{U \in \boldsymbol{U}^\tau} \frac{\mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) \mathcal{P}_C^E\left(p_E, \boldsymbol{p}_{C^E}\right) E}{\gamma_U}. \tag{F.47}$$

Using (F.46) and (F.47), we can drop variables (C, U) and reduce the system of equations to:

$$p_Y = \sum_{x \in s\boldsymbol{X}} H_x^{Y,o}\left(\frac{p_X}{p_Y}\right) p_X \tag{F.48a}$$

$$p_Y = \sum_{Z \in \{\boldsymbol{E}, d\boldsymbol{X}, \boldsymbol{K}\}} H_z^{Y,in}\left(\frac{\widehat{p}_Z}{p_Y}\right) \widehat{p}_Z \tag{F.48b}$$

$$X = H_x^{Y,o}\left(\frac{p_X}{p_Y}\right) Y, \qquad \forall X \in s\boldsymbol{X} \tag{F.48c}$$

$$Z = H_z^{Y,in}\left(\frac{\widehat{p}_Z}{p_Y}\right) Y, \qquad \forall Z \in \{\boldsymbol{E}, d\boldsymbol{X}, \boldsymbol{K}\} \tag{F.48d}$$

$$p_E = \sum_{C \in \boldsymbol{C}^E} \mathcal{P}_C^E\left(p_E, \boldsymbol{p}_{C^E}\right) p_C, \qquad \forall E \in \boldsymbol{E} \tag{F.48e}$$

$$p_C = \sum_{U \in \boldsymbol{U}^C} \mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) p_U, \qquad \forall C \in \boldsymbol{C}^E,\ \forall E \in \boldsymbol{E} \tag{F.48f}$$

$$p_\tau = \sum_{Z \in \{\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\}} H_z^{\tau}\left(\frac{\widehat{p}_Z}{p_\tau}\right) \widehat{p}_Z, \qquad \forall \tau \in \{\mathcal{T}^{\text{ID}}, \mathcal{T}_{\text{o}}\} \tag{F.48g}$$

$$p_\tau = \sum_{U \in \boldsymbol{U}^\tau} \mathcal{P}_U^\tau\left(p_\tau, \boldsymbol{p}_{U^\tau}\right) p_U, \qquad \forall \tau \in \mathcal{T}^{\text{ID}} \tag{F.48h}$$

$$\kappa_\tau = \frac{\mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) \mathcal{P}_C^E\left(p_E, \boldsymbol{p}_{C^E}\right)}{\mathcal{P}_U^\tau\left(p_\tau, \boldsymbol{p}_{U^\tau}\right)} E, \qquad \forall U \in \boldsymbol{U}^C \cap \mathcal{D}_U(\mathcal{T}^{\text{ID}}),\ \forall C \in \boldsymbol{C}^E,\ \forall E \in \boldsymbol{E} \tag{F.48i}$$

$$Z = H_z^{\tau}\left(\frac{\widehat{p}_Z}{p_\tau}\right) \kappa_\tau, \qquad \forall Z \in \{d\boldsymbol{X}(\tau), \boldsymbol{K}(\tau)\},\ \forall \tau \in \{\mathcal{T}^{\text{ID}}, \mathcal{T}_{\text{o}}\} \tag{F.48j}$$

$$\kappa_\tau = \sum_{U \in \boldsymbol{U}^\tau} \frac{\mathcal{P}_U^C\left(p_C, \boldsymbol{p}_U\right) \mathcal{P}_C^E\left(p_E, \boldsymbol{p}_{C^E}\right) E}{\gamma_U}, \qquad \forall \tau \in \mathcal{T}_{\text{o}} \tag{F.48k}$$

$$p_U = p_\tau \gamma_U, \qquad \forall U \in \boldsymbol{U}^\tau, \forall \tau \in \mathcal{T}_{\text{o}} \tag{F.48l}$$

where the set $\boldsymbol{U}^C \cap \mathcal{D}_U(\mathcal{T}^{\text{ID}})$ refers to non-baseline technology goods used in production of component C. This system of equations is of size

$$\#\text{Equations} = 2(1 + n_E + n_{\mathcal{T}^{\text{ID}}} + n_{\mathcal{T}_{\text{o}}}) + n_{sX} + n_C + n_U + n_X + n_K$$

where e.g. n_X reflects demand for all X for baseline technology firms, input-displacing technology firms and final goods firms combined. In a partial equilibrium, input prices, $\widehat{p}_X, \widehat{p}_K$, and output quantities, $s\boldsymbol{X}$, are taken as given (exogenous). In general equilibrium these are endogenized. The system of equations is thus square with the following endogenous variables: $(Y, p_Y, E, p_E, \kappa_\tau, p_\tau, p_C, p_U)$ and (p_X) for $X \in s\boldsymbol{X}$ and $(X(\tau), K(\tau))$ for $\tau \in \{\mathcal{T}^{\text{ID}}, \mathcal{T}_{\text{o}}\}$ and (X, K) for final goods producers.

G Calibration of share and smoothing parameters

Technology firms

All technology firms are described by share- and smoothing parameters in the demand for inputs

(H_z^τ) and similarly for the output-aggregator $\mathcal{P}^\tau$. In the demand for inputs, share-parameters on non-capital inputs ($\boldsymbol{\mu}_x^\tau$) are identified by the data on non-capital input intensities. The share-parameters on capital inputs ($\boldsymbol{\mu}_k^\tau$) are identified by the level of unit costs of the technology ($\bar{c}_\tau$).[32] The smoothing parameter in technology firms' demand function (H_z^τ) is not as easily measured as input-intensities, but should, nonetheless, be guided by technology data. In many instances it is appropriate to use a Leontief function that results in fixed input intensities.[33] In this case, the share-parameters $\boldsymbol{\mu}_x^\tau$ can be observed directly from technology data.

The identification of the share parameters ($\boldsymbol{\mu}_u^\tau$) in the output-aggregator $\mathcal{P}^\tau$ depends on the corresponding relevant smoothing parameter: If the aggregator is Leontief-like with fixed output shares, the share parameters is defined by the relevant coverage potentials ($\boldsymbol{\theta}_\tau^E$). The identification of the smoothing parameter should be guided by the technology catalogue. Consider for instance the case of a heat pump technology that contributes to two energy-services (e.g. spatial heating and dewatering). If the pump always produces the two in fixed ratios, the smoothing parameter is approximately zero and the share parameters reflect the relative output ratios. If the pump can more freely adjust between the two, the smoothing parameter is positive and the share parameters are chosen to be uniform.

Component firms

Component firms are described by share- and smoothing parameters in the aggregator $\mathcal{P}^C$ that govern how overlapping technology goods and baseline technologies compete. The choice of smoothing parameter should be guided by the relevant technology catalogue. In general, however, detailed information on competing technologies is reflected by an aggregator that resembles perfect substitutes, i.e. a smoothing parameter that is close to zero. With the interpretation that competing technology and baseline goods are roughly providing the same service, the natural assumption is to have uniformly distributed share parameters: In this case, only the relative price levels determines to what degree one technology good is preferred over another. As we show in section 4.1, however, we have to depart from this assumption of uniform share parameters in order to match data on current application when we have overlapping technologies.

Energy service firms

Energy service firms are described by share- and smoothing parameters in the aggregator $\mathcal{P}^E$. Share parameters are identified by the underlying relevant technologies' potential in combination with overlap information. In the simple case, consider again a heat pump that produces a single good U^C that is used in a component C^E eventually used to produce an energy service E (e.g. spatial heating). In this case, the share parameter for μ_C^E is defined by the technical potential for the heat pump $\bar{\theta}_\tau^E$.[34]

[32] Note that we have a single data observation $\bar{c}_\tau$ and a vector of parameters $\boldsymbol{\mu}_k^\tau$ to identify. We assume that the relative size share of different types of capital are identified in some other way.

[33] This corresponds to the case of a CES production function with zero substitution, or equivalently, a smoothing parameter that tends to ∞, cf. section B.

[34] In the case where technologies are overlapping, this is slightly more complicated, but the intuition is the same. Appendix C covers the general case.

The choice of smoothing parameter should be guided by the technology catalogue. The traditional bottom-up assumption in the energy service scenario is that of a Leontief-like aggregator with fixed shares C/E. In this case, the share parameters $\boldsymbol{\mu}_C^E$ can be observed directly from technology data.

Baseline technology firms

Baseline technology firms are described by share- and smoothing parameters in the demand for inputs ($H_z^{\tau_o}$), and linear rates of transformation ($\boldsymbol{\gamma}_U^{\tau_o}$) in the output nest. Each baseline technology firm represents the activity within a certain E branch that is not explicitly described by the technology catalogue. As explained in section 4.1, the linear rates of transformation are adjusted to fit the current applications of technologies. The share parameters $\boldsymbol{\mu}_x^{\tau_o}, \boldsymbol{\mu}_k^{\tau_o}$ are identified by matching the inputs that can be assigned to each E branch $\overline{\boldsymbol{X}}^E, \overline{\boldsymbol{K}}^E$.

Finally, for the smoothing parameter in the aggregator $F^{in}_{\tau_o^E}$, we suggest to rely on existing estimation methods. In conventional CGE models, the common way to estimate such smoothing parameters governing the substitutability between inputs (or outputs) is to use time variation in said inputs and their prices, see e.g. Kronborg et al. (2019). Although the method is the same, an important caveat is introduced when applied in our framework. Suppose that one obtains an estimate of a smoothing parameter from historical data. Because the estimate incorporates the substitution made possible by technological adoption over time, using it directly in a model with technologies explicitly featured will overestimate the overall substitution ability of the industry: The explicit technology adoption provides the industry with additional substitution possibilities. For the model to be properly calibrated, the smoothing parameter must reflect less substitution than would be inferred directly from historical data. To accommodate this problem, we suggest one uses only the years for which the technologies in the catalogue existed and estimate the smoother parameter from historical data with those technologies explicitly included in the model. That way, some of the substitution that happens between inputs will be captured by technology adoption, and only the remaining substitution will be captured in the smoothing parameter estimates.

Final goods firms

Final goods firms are described by share- and smoothing parameters in the demand for inputs (H_z^Y) and similarly for the output-aggregator $\mathcal{P}^Y$. The share parameters $\boldsymbol{\mu}_x^Y, \boldsymbol{\mu}_k^Y$ are identified by matching the total use of inputs in the industry $d\overline{\boldsymbol{X}}, d\overline{\boldsymbol{K}}$. Similarly, the share parameters $\boldsymbol{\mu}_E^Y$ are identified from the data on energy services, $\overline{\boldsymbol{E}}$. We suggest to estimate the smoothing parameters for demand and the output allocator using conventional methods, as discussed above for baseline firms.

H A more detailed version of capital

The description of the capital sub-sector is the simplest case of a dynamic version of the model, where we only include one firm that produces capital and rents out K for the various sectors.

In this appendix we introduce a more general version of the capital formation sector: This version allows for (1) sector-specific adjustment costs and (2) applying the national accounts information on the inputs on capital formation. Consider the case where the capital component covers machines and buildings. In the traditional way of modelling capital formation, each sector can increase its holdings of machine/buildings by foregoing some profits (I_t). However, the capital is also inherently produced somewhere in the economy. The national accounts data on investment from/to sectors contain information on this capital formation.

H.1 Capital formation sector

As with any other sector of the economy, the capital formation sector demands final goods from the other sectors and combines them to produce the aggregate investment good output I:

$$\max_{s\boldsymbol{X},d\boldsymbol{X}} \boldsymbol{p}_X \cdot s\boldsymbol{X} - \boldsymbol{p}_X \cdot d\boldsymbol{X}, \qquad \text{s.t. } Y = F_Y^o(s\boldsymbol{X}), \text{ and } Y = F_Y^{in}(d\boldsymbol{X}), \tag{H.49}$$

where $s\boldsymbol{X} = I$ in the simple case with only one durable good. In optimum the solution to this system is simply defined as

$$p_I = \sum_{X \in \boldsymbol{X}} H_x^{Y,in}\left(\frac{p_X}{p_I}\right) p_X \tag{H.50a}$$

$$X = H_X^{Y,in}\left(\frac{p_X}{p_I}\right) p_X. \tag{H.50b}$$

This defines the price on the investment good p_I. Also, note that by drawing on inputs of the general economy, we can compute the emission intensity of the investment goods I. In this way, the cost of capital may also be affected by emission taxes and adoption of end-of-pipe abatement equipment.

Furthermore, while we may account for installation costs for very specific types of capital, we only have data on the capital formation sector on the national accounts level. Thus, the price of investments for different types of capital is only identified for national accounts types of capital (e.g. buildings, machines), and not on technology-specific capital (e.g. 'capital' used for CCS abatement).

H.2 Adjustment costs and application of capital

Every other sector demands this investment good and, after incurring convex installation goods, each sector produces the capital good K as a function of investments I, and previously installed capital. The value of the firm is given by:

$$V_t = \max_{\{d\boldsymbol{X}_s, K_s, I_s\}_{s=t}^T} \mathbb{E}_t \left\{ \sum_{s=t}^{T} R_{t,s}^{-1} \left[\boldsymbol{p}_{X,s} \cdot \left(\boldsymbol{X}_s^s - \boldsymbol{X}_s^d \right) - p_{I,s} I_s - \Psi\left(I_s, K_{s-1}\right) \right] \right\}, \tag{H.51}$$

subject to production technologies and capital accumulation:

$$F_Y^o\left(\boldsymbol{X}_t^s\right) = F_Y^{in}\left(\boldsymbol{X}_t^d, K_{t-1}\right), \qquad K_t = (1-\delta)K_{t-1} + I_t. \tag{H.52}$$

where Ψ is the installation cost function. Optimization wrt. $\boldsymbol{X}^s, \boldsymbol{X}^d$ are straightforward, and yields the solution outlined in the sections above. Optimization wrt. I_t, K_t yields the first order conditions:

$$\lambda_t = R_{t,t+1}^{-1}\left[\mu_{t+1}\frac{\partial F_Y^{in}}{\partial K_t} - \frac{\partial \Psi_{t+1}}{\partial K_t} + \lambda_{t+1}\left(1-\delta\right)\right]$$
$$\lambda_t = p_{I,t} + \frac{\partial \Psi_t}{\partial I_t},$$

where μ_t is shadow cost of the technology constraint, and λ_t the shadow value of capital. The first order condition for optimal demand of a non-capital good is defined as:

$$p_{X,t} = \mu_t \frac{\partial F_Y^{in}}{\partial X_t^d}.$$

If the firm applies more than one type of capital, the relevant equations are simply repeated for each type. Using the first order conditions for I_t, K_t, the demand for K_t can be written in a similar form as non-capital inputs with the capital price:

$$p_{K,t} = \mu_{t+1}\frac{\partial F_Y^{in}}{\partial K_t} \tag{H.53}$$

$$p_{K,t} \equiv R_{t+1}\left(p_{I,t} + \frac{\partial \Psi_t}{\partial I_t}\right) + \frac{\partial \Psi_{t+1}}{\partial K_t} - \left(1-\delta\right)\left(p_{I,t+1} + \frac{\partial \Psi_{t+1}}{\partial I_{t+1}}\right). \tag{H.54}$$

To model gradual adoption of technology-specific capital, it requires that the adjustment costs outlined above are defined for each technology firm. We note that for the aggregate sector to be calibrated to national accounts data, we ultimately have to map the technology-specific capital to the types of capital in the national accounts data. For instance, we can model the installation costs related to 'capital used in CCS abatement', but this type of capital has to be accounted for in the national accounts capital types that usually cover e.g. machines and buildings.

Reducing the number of dynamic equations

Letting capital be technology-specific allows us to model a gradual increase of all technologies. However, this requires that the model includes dynamic forward looking equations as in (H.54) for all technology types. In this general case, the rate of adoption of e.g. CCS abatement equipment depends solely on equilibrium prices and how much of the CCS equipment has been installed already.

To reduce the number of intertemporal problems, however, we may assume that installation costs accrue on the level of total installed level technology capital in each sector. This inherently means that the rate of adoption for e.g. CCS abatement equipment becomes tied to how fast other abatement technologies are phased in. In this case the system of equations correspond to (H.54)

where equations for $(p_{K,t}, K_t)$ are not technology-specific.

I Background information for the toy model

Section 5 presents simulations from a simple toy model. This appendix outlines explicitly the catalogue and functional forms used as well as chosen parameter values and other details. The catalogue as well as the IO data being calibrated are constructed solely to allow illustration of main features of the model. The model represents a single industry producing a single output. To produce, firms rely on four inputs: capital, electricity, oil and coal. The only energy service required for production is heating. Usage of oil and coal emits CO2 with emission intensities of 0.5 and 1.75. As explained in the main text, the catalogue consists of two input-displacing technologies and two end-of-pipe technologies, listed in table 7 here.

Table 7: Technology catalogue behind the toy model

τ	$\overline{x}_{elec}(\tau)$	$\overline{x}_{oil}(\tau)$	$\overline{x}_{coal}(\tau)$	$\overline{c}_\tau$	$\overline{\theta}_\tau$	Overlap share	Overlapping τ
Input-displacing technologies:							
τ_{semi}	1.5	0.6	0	3.5 \$	35 %	100 %	τ_{clean}
τ_{clean}	3	0	0	7 \$	35 %	100 %	τ_{semi}
End-of-pipe technologies:							
τ_{low}	3	0	0	5 \$	10 %	0 %	-
τ_{high}	10	0	0	13.5 \$	70 %	0 %	-

Note: We do not use the current applications of the technologies for the toy model, so these are not stated. The potentials, $\overline{\theta}_\tau$ are stated as a fraction of the single energy service E for the input-displacing technologies, and as a fraction of the only emission type CO2 for the end-of-pipe technologies.

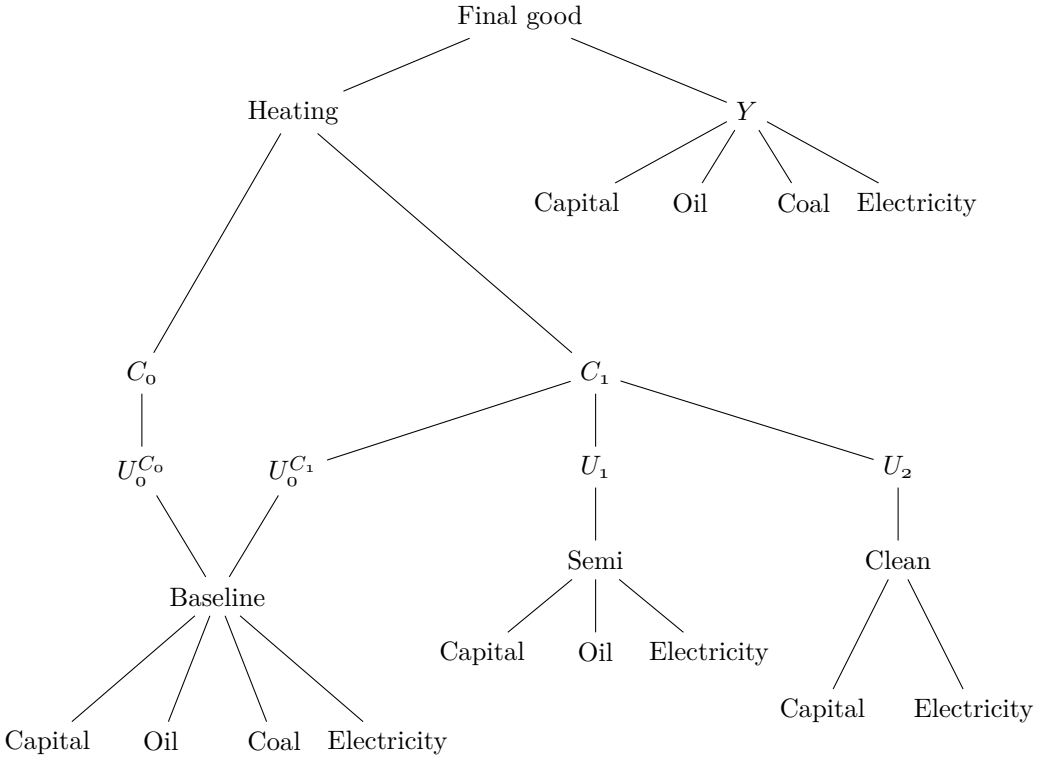
The technology catalogue gives rise to the input-displacing subtree and the end-of-pipe subtree depicted in figures I.1 and I.2 respectively. For the input use outside energy services $\boldsymbol{E}$, we simply assume that these are combined into an aggregate Y which is then combined together with the single E to produce the final good.

To solve the model, we must assume specific functional forms in all nests and specify the smoothing parameters of these. Table 8 summarizes these choices.

The toy model is calibrated to a set a fictitious input-output data, stated in table 9. We do not calibrate the toy model to current applications, but otherwise the calibration follows the general calibration procedure of section 4. Since we do not target current applications, the parameters that would otherwise have been adjusted are simply set to their respective neutral values, see equations 22 and 26. The only exception to this are the linear rates of transformation, these are set to 1 rather than γ^E_τ.

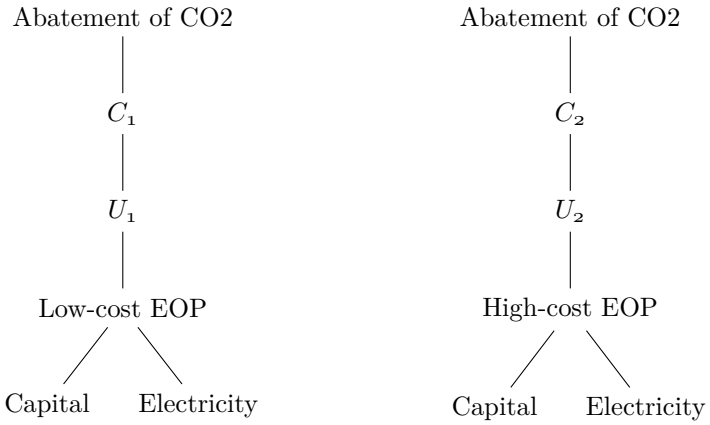
To solve the model in partial equilibrium, we fix pure input prices of capital, electricity, oil and coal to 1 as well as exogenize the quantity of final goods produced (to 108, as the IO data prescribes). The emission tax is set at an initial value of 0.5 and then gradually rises.

Figure I.1: The nested production tree of the toy model, input-displacing subtree



Note: The figure depicts the input-displacing subtree of the toy model.

Figure I.2: The nested production tree of the toy model, end-of-pipe subtree



Note: The figure depicts the end-of-pipe subtree of the toy model.

Table 8: Functional forms and smoothing parameters chosen for the toy model

Nest	**Functional form**	**Smoothing parameter value**
Heating and $Y \longrightarrow$ Final good	CES	∞
Inputs $\longrightarrow Y$	CES	2
Components $\boldsymbol{C} \longrightarrow$ Heating	NCES	∞
Technology goods $\boldsymbol{U} \longrightarrow \boldsymbol{C}$	MNL	0.1
Baseline technology $\longrightarrow$ Baseline technology goods $\boldsymbol{U}_0$	Linear	-
Inputs $\boldsymbol{X} \longrightarrow$ Technologies	CES	∞

The aggregator from $\boldsymbol{U}$ to $\boldsymbol{C}$ as well as from inputs $\boldsymbol{X}$ into technologies is the same for both input-displacing and end-of-pipe. The function splitting technologies' composite into technology goods need not be specified here since no non-baseline technology produces more than one technology good.
For an overview of the demands resulting from the chosen functional forms, see appendix B.

Table 9: Input-output data behind toy model

Data	**Value**
Input use for heating	
Electricity	40
Oil	16
Coal	17
Capital	15
Aggregate input use	
Electricity	65
Oil	28
Coal	20
Capital	25
Aggregate heating use	55
Final good production	108

Heating refers to the only energy service E in the model.