Electricity Generation, Carbon Pricing, and Air Quality: Evidence from the EU ETS*

Nico Lukas Kasparetz†

Ulrich J. Wagner‡

March 26, 2026

Abstract

Cap-and-trade systems such as the European Union Emissions Trading System (EU ETS) are designed to reduce greenhouse gas emissions cost-effectively, but their implications for co-emitted air pollutants and associated health outcomes remain ambiguous. We develop an integrated framework that combines econometric modeling of electricity dispatch, emissions estimation, atmospheric transport modeling, and epidemiological analysis to quantify these effects. Using hourly data from fossil-fuel fired power plants in 11 European countries, we estimate the impact of carbon pricing on unit-level generation and associated emissions of CO_2 and precursors of $PM_{2.5}$. These emissions are translated into changes in ambient $PM_{2.5}$ concentrations and premature mortality. Comparing observed outcomes with counterfactual scenarios without carbon pricing, we find that the EU ETS reduced $PM_{2.5}$-related mortality by 6.9% between 2016 and 2021. A counterfactual carbon price floor of €50 would have yielded an additional 8.7% reduction. Monetizing health impacts reveals substantial net welfare gains, with co-benefits exceeding increased generation costs by €2.2 billion. However, marginal co-benefits decline at higher carbon prices, suggesting diminishing returns. Overall, our findings indicate that carbon pricing in the power sector can deliver significant health co-benefits that reinforce its role as an efficient environmental policy instrument.

*Still preliminary. We are indebted to Laure de Preux for generous help with computing pollution intensities from EPRTR data and for many helpful discussions. We are grateful to Yixuan Gu and Daven Henze for graciously providing us with adjoint sensitivities from Gu et al. (2023) and to Siyu Zhang for help with processing those. We thank Leo Launhardt, Michael Indenkämpen and Yuqi Wang for research assistance with entity linking across ENTSOE, EPRTR and LCP data. We have benefitted from valuable comments and suggestions from participants at the first Alpine Workshop on "Energy Economics and Policy" in Airolo in September 2025. Support by the German Research Foundation (DFG) through CRC TR 224 (Project B07) is gratefully acknowledged. This project received financial support from the European Research Council under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No. 865181). Declarations of interest: none.

†Department of Economics, University of Mannheim. nico.kasparetz@uni-mannheim.de

‡Department of Economics, University of Mannheim, and CEPR. ulrich.wagner@uni-mannheim.de

1 Introduction

To avoid dangerous levels of global warming beyond 1.5°C, greenhouse gas (GHG) emissions must drop to net zero within the next thirty years (IPCC, 2023). The electricity sector plays a critical role for achieving this. On one hand, electricity generation accounts for about 40% of global carbon dioxide (CO_2) emissions per year (International Energy Agency, 2026). On the other hand, emissions reductions by sectors like transportation, buildings and manufacturing increasingly rely on electrification. The combination of ambitious abatement targets, a heavy dependence on fossil fuels and increasing power demand means all point in the direction of high costs of decarbonizing the electricity sector. Since 2005, several of the world's largest electricity markets such as the European Union (EU), California, and China have been regulated with cap-and-trade policies. Two aspects make this instrument especially well-suited to reducing power sector emissions. First, electricity firms have the incentives and the ability to factor carbon prices into their production decisions. Second, because electricity is traded on an integrated grid, emissions monitoring and carbon leakage pose smaller challenges than they do for industrial emitters. Both these features strongly favor cost-effective carbon abatement, yet additional efficiency implications may arise as CO_2 is jointly released with air pollutants that pose direct risk to human health.

While uniform emissions reductions in CO_2 and co-pollution might create ancillary health benefits, this cannot be taken for granted with market-based climate policies in the electricity sector. In the short and medium run, carbon pricing reduces CO_2 emissions by shifting generation to less carbon-intensive power plants—both within and between interconnected electricity markets. Whether such reallocation increases or decreases total population exposure to air pollution depends not only on plants' pollution intensities and proximity to population centers but also on weather and atmospheric chemistry, all of which matter for pollution transport and the formation of secondary air pollutants like fine dust ($PM_{2.5}$). If health externalities are not fully internalized by existing regulations, this creates regulatory trade-offs or complementarities which ought to be considered when designing mutually consistent climate and air pollution control policies for the power sector.

This paper develops a new framework for measuring such trade-offs in the context of by electricity firms regulated under the EU Emissions Trading System (ETS), a cap-and-trade system for CO_2 and other GHG emissions. To overcome the challenges presented by this endeavor, we draw on elements from program evaluation, integrated assessment and atmospheric chemistry. To begin, we estimate the impact of carbon pricing on short-run production decisions by fitting a machine-learning model to unit-level hourly generation data from eleven countries. The model accounts for economic determinants of dispatch,

such as the merit-order principle, grid congestion, and electricity trade. Augmenting the model with unit- and plant-level emissions intensities—estimated from data collected under the auspices of EU air quality directives—allows us to predict emissions of CO_2 and $PM_{2.5}$ precursor pollutants for each generating unit. To track pollution from these sources to the place of impact, we use a custom-built chemical transport model capable of mapping changes in generation into changes in ambient concentrations of $PM_{2.5}$ on a grid for the European domain. Our final health outcome, total premature deaths due to $PM_{2.5}$, is estimated by combining data on population and ambient pollution concentrations in each grid cell with dose-response functions from the Global Burden of Disease project.

We identify the causal impact of the ETS on mortality and other outcomes by comparing simulated equilibria on the electricity market for both observed and counterfactual carbon prices. Our results indicate that ETS-induced fuel switching reduced $PM_{2.5}$-related mortality by 6.9 percent over the period 2016–2021, compared to a zero carbon price. This finding highlights substantial health co-benefits of carbon pricing in the form of improved public health. In further analysis, we evaluate the additional health gains from introducing a carbon price floor of €50 per allowance. We find that such a policy would have generated a further reduction in $PM_{2.5}$-related mortality of 8.7 percent. Overall, our results suggest that higher and more stable carbon prices, in addition to reducing CO_2 emissions that drive global climate change, have delivered significant local benefits in terms of improved air quality and reduced mortality from population exposure to fine dust.

An important caveat is that such ancillary benefits increase welfare only if they have not already been fully internalized by pre-existing local air quality regulations. Since we cannot fully observe such regulations, we resort to a sufficient-statistic approach. For co-pollution reductions to increase welfare, the incremental health benefits must outweigh the incremental cost of electricity generation due to carbon pricing. We evaluate this difference for the historical time series of ETS permit prices between 2016 and 2021 and using a conservative value of statistical life at €3 million. We find that carbon pricing raised total generation costs by €588 million while inducing health co-benefits worth €2.8 billion; the resulting net benefits amount to €2.2 billion. Thus, we conclude with some confidence that carbon trading among European power plants did indeed enhance efficiency by internalizing air pollution externalities.

We can attribute this result to co-pollution benefits per tonne of CO_2 abated being almost five times larger than the average abatement cost of €4.3, at the prevailing carbon prices. When we compute net co-benefits under counterfactual permit prices of €50 or higher, the scope for net co-benefits shrinks as the higher carbon price justifies more costly fuel switching which is not offset by similar increases in co-benefits. Therefore, co-benefits in

this scenario are still positive but drop from €16 to €10 per tonne abated. In light of these results, we conclude that decentralized responses to carbon pricing by electricity firms have reinforced—rather counteracted —the beneficial effects of other policies targeted at reducing the external effects of air pollution emissions from the European power sector. As rising carbon prices will rise in the future, so will marginal abatement costs, with the consequence that marginal net co-benefits of fuel switching will ultimately vanish. At that point, the realization of additional health co-benefits will critically depend on broader changes to the energy system which are beyond the scope of this study.

Related Literature Our study provides the first micro-founded evaluation of the ancillary health benefits of climate policy, focusing on power plants operating in the world's first and most stringent carbon market. In so doing, our paper builds on prior research on the EU ETS, on the external costs of electricity generation, and on integrated assessment.

An early literature in the tradition of integrated assessment models (IAM) has sought to quantify the ancillary benefits of climate policy in terms of avoided pollution by simulating emissions paths with and without the policy (Aunan et al., 2007; Burtraw et al., 2003; Ekins, 1996; Ekin, 1996; Pittel and Rübbelke, 2008; Rypdal et al., 2007; van Vuuren et al., 2006; Driscoll et al., 2015). Given the uncertainty about future regulations, technology, economic development, demography, and natural activities (Davis et al., 2000), it is very difficult to ascertain the path of emissions *ex ante* with confidence. To relax these assumptions, we train a machine learning model on ex-post data to predict how generation by hundreds of point sources varies with the carbon price. The granularity of our approach is particularly well-suited to our question because we want to model decentralized pollution releases that are characteristic for compliance with market-based regulatory instruments.

Another related literature is on the ex-post evaluation of emissions trading schemes (Carlson et al., 2000; Fowlie et al., 2012), and in particular, the EU ETS (Martin et al., 2016; Colmer et al., 2024). Previous research on carbon pricing in the power sector has focused on CO_2 abatement (Abrell et al., 2022; Zaklan, 2023) and carbon cost pass-through (Fabra and Reguant, 2014; Hintermann, 2016) in the power sector. Our paper adds to this literature by also measuring the impacts on ancillary health effects of abating CO_2 emissions which, to the best of our knowledge, have not been studied yet in the context of the EU ETS.

While heterogeneity in abatement costs is what renders cap-and-trade an efficient policy, there is increasing evidence that heterogeneity in environmental damages across air sheds can take away from that efficiency, unless appropriate exchange rates are applied for permit trades between air sheds (Mauzerall et al., 2005; Muller and Mendelsohn, 2009). In a similar vein,

cap-and-trade programs have been under scrutiny with regard to the question of whether they shift pollution emissions – and the environmental damages associated with them – towards minorities and low-income households (Fowlie et al., 2012; Grainger and Ruangmas, 2018). Recent research in this area has analyzed the environmental justice implications of the Californian carbon market for four prominent air pollutants (Cushing et al., 2018; Hernandez-Cortes and Meng, 2023). We contribute to this literature by measuring and comparing CO2 abatement costs and monetized health benefits in the context of the world's first and largest carbon market.

Our paper is closely related to empirical research on the social cost of the electricity generation. Compared to Borenstein and Bushnell (2022), who estimate social costs of US electricity generation, we focus on the EU. Our approach to modeling generation and pollution transport is more structural. Following previous research on the UK carbon price floor (Abrell et al., 2022) and Germany's nuclear exit (Jarvis et al., 2022), we rely on machine learning to estimate causal impacts of policy on electricity generation. While Abrell et al. (2022) also focus on carbon pricing, our machine learning model builds on Jarvis et al. (2022) which we extend by (i) covering a large set of interconnected electricity markets and the electricity trade between them, (ii) accounting for carbon pricing, and (iii) predicting co-emissions using empirically estimated pollution intensities.

In sum, our paper brings to bear new data, improved methods and a novel link to atmospheric chemistry to bear on analyzing the ancillary health co-benfits of the EU's flagship climate policy instrument.

2 Policy Background

2.1 The EU Emissions Trading Scheme

The EU ETS is a cap-and-trade system for CO_2 emissions, initially focused on power generation and energy-intensive manufacturing industries. Launched in 2005 by the Emissions Trading Directive[1], the policy has been covering more than 2 billion tons of CO_2 emitted at roughly 12,000 installations in 31 countries. Participation in the EU ETS is mandatory for all combustion installations with a rated thermal input exceeding 20 megawatts (MW), which includes conventional power plants. In addition, the scheme covers large installations engaging in activities that are intensive in CO_2 emissions such as mineral oil refineries, coke ovens, iron and steel, and factories producing cement, glass,

[1]Directive 2009/29/EC of the European Parliament and of the Council amending Directive 2003/87/EC so as to improve and extend the greenhouse gas emission allowance trading scheme of the Community (2009) OJ L 140, 5.6.2009, p. 63-87 (Emissions Trading Directive).

lime, bricks, ceramics, and pulp and paper. Since 2013, the system also include airlines and aluminium manufacturing installations, as well several greenhouse gases other than CO_2.

Participating installations are required to surrender a pollution permit, known as an EU allowance (EUA), for each metric ton of CO_2 emitted per year. These allowances are distributed to firms in part for free or through an auctioning system. The number of permits - the "cap" - is chosen to limit the total amount of pollution below the levels that would otherwise arise - "business-as-usual" (BAU) emissions. The cap on total emissions was set to decline at a rate of 1.74% per year up until 2020 and has since been decreasing by 2.2% per year, so as to ensure the EU interim target for 2030 of a CO_2 emissions reduction by 43% from 2005 levels.

The declining cap induces scarcity of pollution permits and puts upward pressure on the price for emission permits as firms trade on the carbon market. In this way, the EU ETS has established a uniform carbon price that signals the opportunity cost of emitting CO_2 to all market participants. The CO_2 price provides each of them with an incentive to reduce emissions up until the point at which they are indifferent between buying one permit at the market price and paying the cost of reducing its emissions by one additional ton of carbon.

By creating a multi-billion Euro market for emission permits, the EU ETS represents the largest and arguably most ambitious carbon pricing policy thus far. It is the cornerstone of the EU's policy roadmap towards carbon neutrality. It is also regarded as a prototype for a future global carbon market.[2]

2.2 Air Pollution Control Policies

Pollution emissions from fossil fuel–fired power plants have long been subject to both national and European regulations.

EU regulations A series of increasingly stringent EU regulations has targeted large combustion plants (LCPs), i.e., industrial combustion plants with a total rated thermal input equal to or greater than 50MW. The central legal instrument during the early years of the EU ETS was the Large Combustion Plant (LCP) Directive.[3] It imposed emission limit values for key pollutants such as SO_2, NO_x, and PM. The LCP directive also required older plants either to retrofit pollution-control technologies (e.g. flue gas desulfurization) or to opt out and close after limited operating hours by 2015.

[2] A comprehensive review of the history and structure of the EU ETS can be found in Ellerman et al. (2016).

[3] European Union, Directive 2001/80/EC on the Limitation of Emissions of Certain Pollutants into the Air from Large Combustion Plants, OJ L309 (2001): 1–21.

From 2016 onwards, the LCP was effectively replaced by the Industrial Emissions Directive (IED).[4] The IED required permits based on "best available techniques" (BAT), with stricter and more harmonized emission limit values. For power plants, this meant tighter operational conditions and a stronger link between environmental performance and permitting. From around 2015 onward, operators had to comply with increasingly strict limits derived from BAT reference documents, which further accelerated emission reductions and technological upgrades.

The Pollutant Release and Transfer Register (PRTR) Regulation complements the LCP and Industrial Emissions Directives by requiring operators of power plants and large industrial facilities to report annual emissions of a broad range of air pollutants (subject to reporting thresholds).[5] Unlike the two directives, which set emission limits and permitting conditions, the PRTR focuses on transparency and public access to data through an EU-wide database, henceforth referred to as the E-PRTR. The more integrated regulatory framework introduced by the IED has established standardization, monitoring and verification requirements that have benefited data quality in the E-PRTR.

Based on these data, LCP emissions have fallen dramatically since the onset of the EU ETS. From 2004 until 2024, SO_2 and dust emissions fell by 94% and NO_x emissions by 74% (Agency, 2026). The main reasons were technological changes and a reduction in the use of coal and other fossil fuels. Against this background, it is important to empirically distinguish the direct effects of air pollution control policies on mortality from the ancillary effects of carbon pricing on copollution emissions—a formidable task. We address this issue by focusing on the short-run effects of carbon pricing on pollution emissions which are driven by fuel switching in response to daily variation in carbon prices. We take as given any longer-term decisions on plant closures and new plant openings, which might be driven by retrofits or BAT regulations stipulated in LCP and IED. This is conservative in the sense that it ignores such extensive-margin decisions even if they were indeed driven by carbon pricing.

National regulations National policies to limit SO_2 and NO_x emissions from power plants in European countries have generally operated alongside EU directives, but often go further through stricter standards, economic instruments, and structural energy policies.

First, when implementing EU directives through environmental permits, national authorities often impose stricter conditions or faster compliance timelines. Second, countries like Sweden have been pricing pollution via taxes on SO_2 and NO_x emissions. Third, many countries adopted technology- and fuel-based measures to reduce emissions. These include

[4]European Union, Directive 2010/75/EU on Industrial Emissions, OJ L334 (2010): 17–119.

[5]European Parliament and Council, Regulation (EC) No 166/2006 concerning the establishment of a European Pollutant Release and Transfer Register, OJ L33 (2006): 1–17.

mandatory installation of flue gas desulfurisation for SO_2 control and selective catalytic reduction (SCR) for NO_x, as well as fuel switching from coal to gas or renewables. Fourth, since the 2010s, national strategies increasingly rely on structural and economic measures, such as coal phase-outs, emissions ceilings, and climate-energy policies. Overall, most national policies have tended to promote cleaner technologies and accelerate the shift away from high-emission fossil fuels; however, examples exist of such policies working in the other direction (e.g., the German nuclear exit, cf. Jarvis et al., 2022).

3 Electricity Markets in Europe

We study the short-run impact of carbon pricing on electricity generation and emission in 11 European countries. This section provides background on liberalized wholesale electricity markets in Europe, with a focus on the merit-order principle that determines dispatch in day-ahead markets. We begin by describing the available data required to build empirical merit-order curves.

3.1 Data on Electricity Generation

The main data source for this study is the Central Information Transparency Platform of the European Network of Transmission System Operators for Electricity (ENTSO-E). The Transparency Platform offers free access to essential data and insights on electricity generation, transmission, and consumption across the pan-European wholesale energy market. ENTSO-E provides hourly data on the electricity market for all countries of the European Economic Area (EEA) plus Albania, Bosnia and Herzegovina, North Macedonia, Montenegro, Serbia, Switzerland, and Ukraine. Turkey and Moldova have observer status. The United Kingdom was a member until June 15, 2021. Data submission for England, Wales and Scotland ceased on that day. Because Northern Ireland is part of the Single Electricity Market (SEM) on the island of Ireland, the Northern Irish system operator has remained an ENTSO-E member and continues to submit data on electricity generation to ENTSO-E.

Actual Generation per Generation Unit This data set [16.1.A] is the centerpiece of this project. It contains hourly data on electricity generation on the unit level. A unit refers to an individual unit in a fossil fuel power plant. The data set contains the unit's bidding zone, a unit identifier, the unit's production type (lignite, hard coal, natural gas, hydro power, wind, etc.), the installed capacity, and the actual generation in the given hour. The units can be further matched to the plants they belong to, allowing us to identify sets

of units within one plant. EU regulation requires all operators to report a unit's hourly electricity generation to ENTSO-E if the unit has at least 100 MW of installed capacity. However, this threshold appears to be applied somewhat inconsistently. Hirth et al. (2018) point out that some units with less than 100 MW are included while other generating units are completely missing. In the case of solar parks and wind parks, some operators apply the threshold to the individual PV panels or turbines and therefore do not report any unit-level data (e.g. German operators) whereas others consider their entire park to be one unit and therefore do report (e.g. British operators until June 15, 2021).

Actual Total Load and Actual Generation per Production Type The first data set [6.1.A] contains gross electricity demand on the bidding zone level. Gross electricity demand is called "Total Load". The second data set [16.1.B] contains total electricity generation by production type at the level of the bidding zone. In both data sets, most bidding zones provide data at the hourly level. Some bidding zones (Hungary, Belgium, the Netherlands, Germany, Austria, and Luxembourg) report total load and generation per type in 15-minute intervals. The British and Irish bidding zones report in half-hourly intervals. The information from both data sets is combined to obtain an estimate for the share of total load that is relevant for fossil fuel units.[6]

Further ENTSO-E data sets The data set Cross-Border Physical Flows [12.1.G] provides hourly data on physical flows of electricity between bidding zones. Day-ahead electricity prices are provided in the data set Day-ahead Prices [12.1.D]. These are the prices that result from the merit order principle using day-ahead forecasts of total load and that generally determine dispatch in the energy market, with the exception of short-term dispatch that responds to unexpectedly high demand. These day-ahead prices vary at the hourly level as the merit curve is established separately for every hour of the day. ENTSO-E also makes available data on unavailability of units due to maintenance work or due to forced outages in Planned Unavailability of Generation Units [15.1.A] and Changes in Actual Availability of Generation Units [15.1.B].

Other data sources Other price data come from LSEG Datastream, formerly known as Refinitiv Datastream. This source provides daily data on the API2 coal index. This is the standard benchmark for European steam coal used primarily in electricity generation. According to S&P global, this price "reflects the tradable, repeatable spot market value"

[6] Assessing an early version of the ENTSO-E data in comparison to other sources, Hirth et al. (2018) find that reporting gap were relatively stable over time.

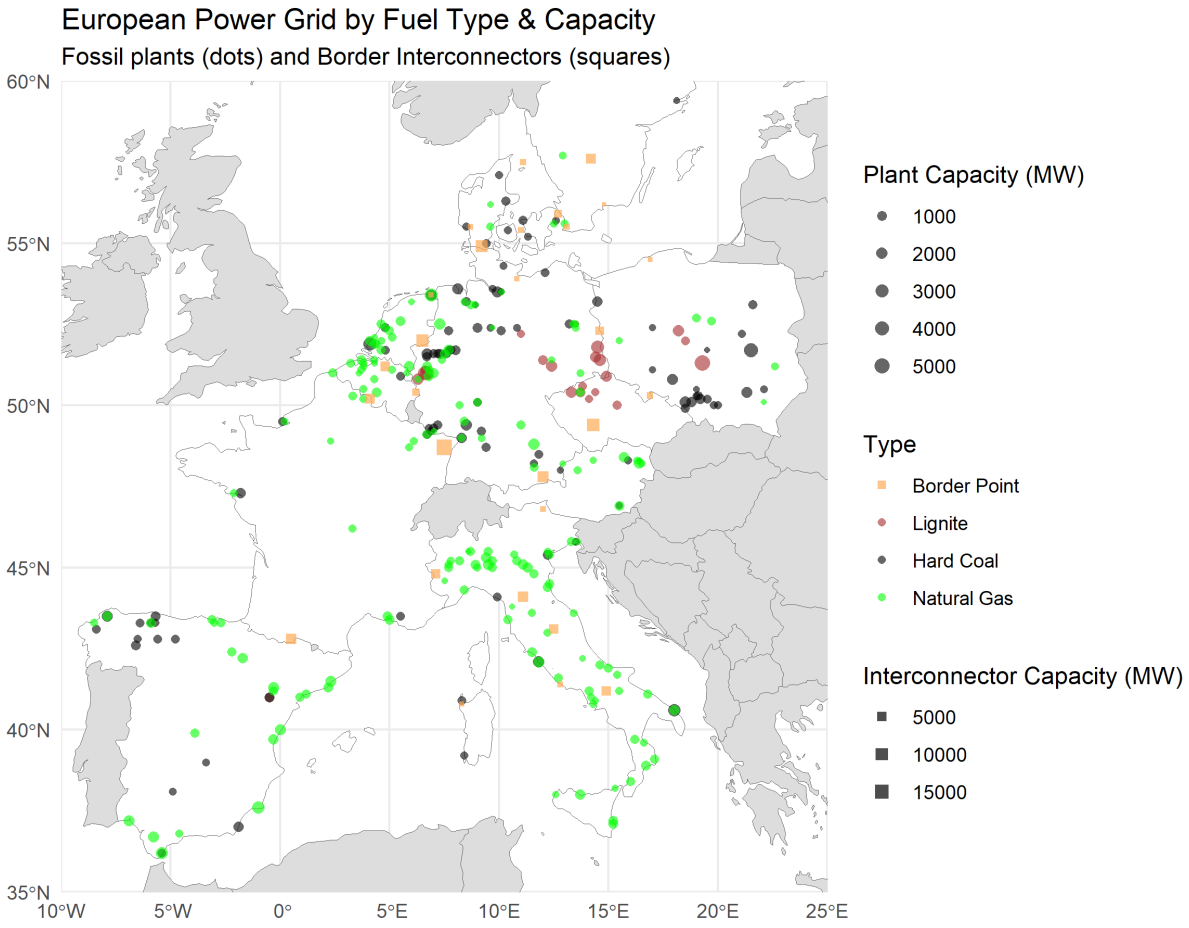
of hard coal. LSEG Datastream also provides data on the daily settlement price of EU ETS emission allowances. This price is the official end-of-day price on the European Energy Exchange (EEX), a major trading venue for both spot and derivatives. There is also data on natural gas prices by natural gas trading hub. So far, these include the PSV hub for Italy, MIGBAS for Spain and Portugal, POLPX for Poland, THE for Germany, TTF for the Netherlands, and VTP for Austria and Czechia. Additionally, we impute natural gas prices for Denmark and (Southern) Sweden with the THE hub, and for Belgium using the TTF hub.

Several other data sources expand the unit-level data from ENTSO-E. The JRC Open Power Plants Database from the Joint Research Centre of the European Commission provides unit-level information on minimum stable load, ramping capabilities, and net electric efficiency. This data is valuable to calculate marginal costs. "Minimum stable load" is the lowest fraction of its maximum capacity at which it can run continuously. Coal plants, for instance, need to run at a minimum load in order to ensure combustion stability and produce enough steam to maintain required temperature and pressure. The "ramping capability" is the speed at which units can power up or down. Fossil fuel units generally cannot go from being shut off to maximum generation in an instant but need time to power up. In the case of lignite units, this can take several hours. Lastly, "net electric efficiency" denotes the proportion of a fuel's inherent energy that is converted into electricity. Conventional coal- or gas-fired power stations typically exhibit net electric efficiencies of only 30–40 percent, indicating that merely 30–40 percent of the combustion energy is converted into electricity, while the remainder is dissipated as waste heat. The Open Power Plants Database also provides latitude and longitude information for units as well as information on possible decommissioning.[7]

Data on transportation cost for hard coal in Germany comes from the Open Source Electricity Model for Germany ELMOD-DE, maintained by DIW German Institute for Economic Research (Egerer, 2016). Information on a unit's technology type (steam turbine, gas turbine, or combined cycle) as well as on whether it is a combined heat-and-power (CHP) unit comes from three sources: First, Jarvis et al. (2022) provide this information for German power units. Second, Open Power System Data (2020) provides data on the technology type and CHP status of fossil fuel power units in 15 EEA countries plus the United Kingdom and Switzerland. Third, gaps in the data have been manually filled by using Global Energy Monitor's GEM.wiki.

[7]Occasionally, units stay in the ENTSO-E register and continue to report zero generation even if they were already decommissioned.

Figure 1: Location Of Fossil Generating Units In The Sample; Aggregated To Power Plant Level



Notes: The map shows all fossil-fuel power electricity generation units in our dataset, aggregated to the plant level for readability, as well as border in points. The map shows lignite plants ("CB", brown), hard coal plants ("CH", black), natural gas plants ("G", green), and border in points ("BI", orange). The location of border in points is calculated as follows: when the border between two bidding zones follows country borders, I take the geographic average of their border as the geographic location of the border point. For underseas cables, I use the point at which the cable ends. For within-country biddingzone borders, I use the geographic average of the borders. In the case of Czechia, which bordered the joint German-Austrian-Luxembourgian bidding zone until September 2018, I take the geographic average of the Czech border to both Austria and Germany.

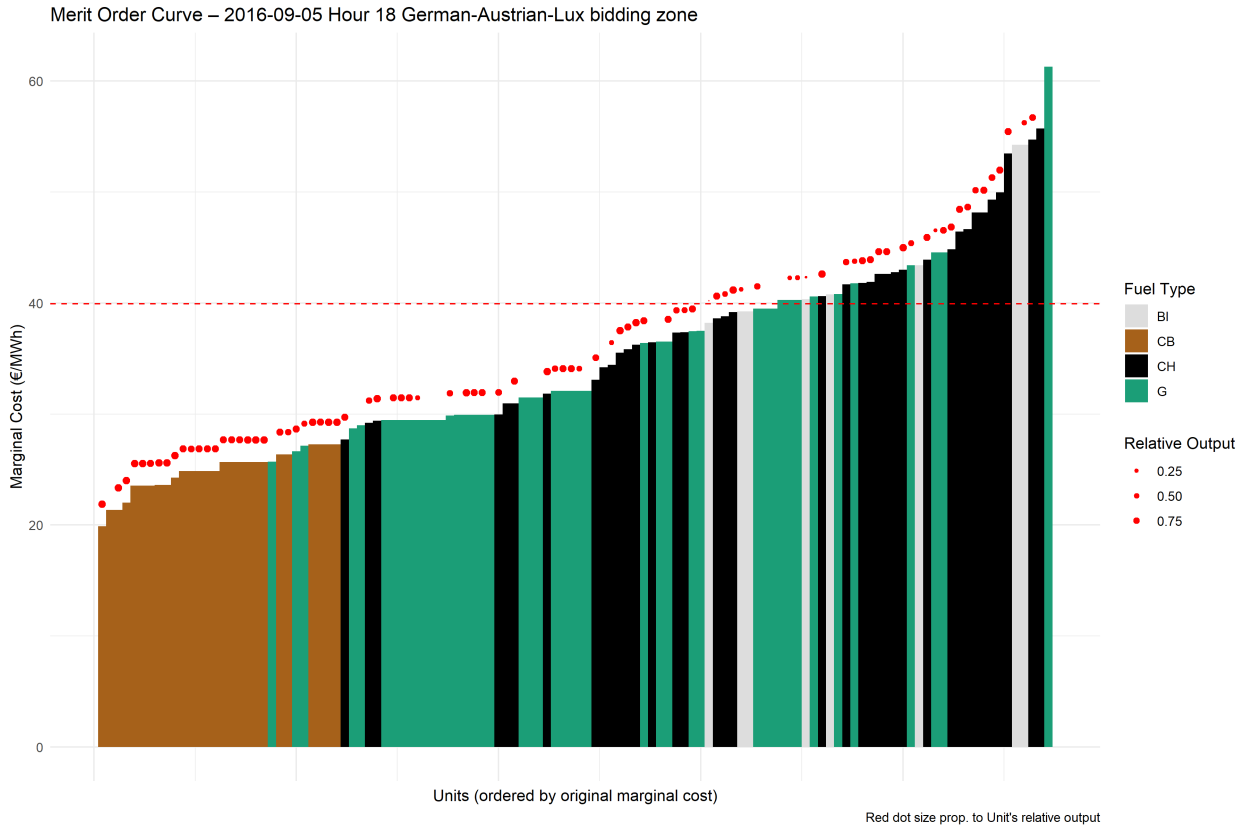
3.2 The Merit Order

Liberalized electricity markets in the EU follow the merit-order principle which ensures that electricity generation occurs at lowest cost. A wholesale market is characterized by electricity generators that supply electricity and retail firms that demand electricity. Within so-called bidding zones –well-defined geographic areas with a strong capacity to transmit electricity from generators to retail firms– the bulk of electricity trade is organized in the form of a day-ahead market: Every day before noon, plant operators place bids to disclose how much electricity they are willing to supply with each of their power units at which price. This is done separately for every hour of the following day. Once the bidding zone's power exchange has collected all the bids, it aggregates them into a stacked supply curve, ordering the bids of energy suppliers by increasing price. At the same time, it also converts buy bids by retail firms into a demand curve. This mechanism yields a candidate equilibrium price in the wholesale market for electricity for every hour and bidding zone.

Since bidding zones are connected via transmission lines, this candidate equilibrium price does not clear the day-ahead market, however. The power exchanges transmit their merit curves to the European Market Coupling Company, which then runs them through their EUPHEMIA algorithm to match supply and demand in every bidding zone accounting for cross-border flows. This allows for lowest-cost generation across bidding zones. For instance, if the no-trade equilibrium prices were €10/MWh in bidding zone A and €20/MWh in bidding zone B, then additional units in A should operate and generate electricity to be exported to B. This raises hypothetical electricity prices in A and lowers them in B until either prices between A and B have equalized or transmission capacities are at their maximum. This mechanism eventually establishes the cross-border merit orders and day-ahead equilibrium wholesale prices.

Figure 2 shows one of our estimated merit curves. It corresponds to the German-Austrian-Luxembourgian bidding zone on January 7, 2019 at the 18^{th} hour of the day. Units are ordered by their marginal costs in €/MWh. Units run on either lignite (CB), hard coal (CH), natural gas (G), or border points (BI). Strictly speaking, this is only the right tail of the entire merit curve. Renewable and nuclear units lie to the left of the fossil fuel section of the merit curve. The former are generally subject to priority dispatch. EU regulation mandates that power exchanges and system operators must schedule the output of renewable energy units ahead of fossil-fired (and nuclear) plants. This is often implemented in practice through feed-in tariffs or premia coupled with priority grid access rules. For instance, Germany's renewable energy law (EEG) guarantees priority dispatch for wind, solar, biomass, hydro, and geothermal power units. Accordingly, renewable energy supply

Figure 2: The Merit-Order Curve for Fossil Power Plants



Note: The graph shows the estimated merit curve on September 5, 2016, at the 18th hour of the day, in the German-Austrian-Luxembourgian bidding zone. Units are ordered by their estimated marginal costs in €/MWh. The dashed red line indicates the equilibrium day-ahead market price for electricity in the bidding zone in the given hour, which was €39.92/MWh. Red dots indicate units that are not shut off in the given hour. Units are either lignite units (CB), hard coal units (CH), natural gas units (G), or border points (BI).

always comes first in the merit curve. Nuclear energy units are also dispatched before fossil-fuel powered units, in practice, due to their low running costs. Therefore, these units are not relevant for establishing the equilibrium electricity price as long as residual energy demand, after accounting for renewable and nuclear production, is still positive.

On January 7, 2019, the day-ahead ETS allowance price was low at €4.09. Accordingly, lignite units at this low carbon price are the cheapest fossil fuel units in the market due to the low fuel costs of lignite relative to hard coal and natural gas, and run near full capacity. Next, there is a block of highly efficient natural gas units at marginal costs of around €30/MWh. However, not all of them are active. Notably, there are several active units at the upper end of the merit curve, above the equilibrium day-ahead price. This may be due to several factors that the estimated merit curve does not account for. An important one is short-term dispatch. The merit curve displays what suppliers bid on the day-ahead market based on day-ahead estimates of marginal costs, and the dashed red line indicates the resulting equilibrium day-ahead electricity price of in this case €39.92/MWh. However, this is based on expected demand and expected availability of units. In the given hour, actual total load was about 1,000 MWh higher than the day-ahead load forecast on which the day-ahead electricity price is based. In such a case, suppliers of fossil fuel energy can bid extra capacity on the intraday market, which allows for last-minute adjustments to balance supply with realized demand. If actual load exceeds forecasted load, prices in the intraday market will exceed day-ahead prices so that units with prohibitive marginal cost in the day-ahead market may now place short-term bids and still produce. In addition to that, some suppliers may be unable to deliver on their day-ahead bids, for instance due to forced outages. In this case, high-cost units may also step in and fill the supply gap in the intraday market.

Ramping constraints also play an important role in plant operation schedules. In particular lignite and coal units are inert and may take several hours to fully launch or shut down. Letting them run at current levels may be more economical if adverse market conditions are expected to last only for a short time span. Lastly, the merit curve assumes that the bidding zone is a copperplate on which any unit can satisfy any demand. Transmission within bidding zones, however, is also subject to certain constraints. If there are isolated demand peaks in some parts of the bidding zone, then only units near these demand peaks may be able to deliver the required electricity, which will in general not be in line with the merit order. This is an important concern in large bidding zones with constrained internal transmission capacities. In Spain, for instance, electricity demand is generally concentrated in a few metropolitan areas, with vast areas with little demand inbetween. Actual generation therefore tends to differ substantially from predicted generation under the merit order.

3.3 Determinants of the Merit Order

This subsection discusses the key variables that determine a generating unit's position in the merit order and whether it actually gets to dispatch. We shall include those variables in the random forest algorithm to predict hourly electricity generation at the unit level, discussed below.

Utilization rate To study unit-level electricity dispatch decisions, we focus on units' utilization rate which takes values between 0 and 1. The utilization rate measures how much of the unit's capacity is used in a given hour. For example, if a unit has a capacity of 100 MW, generation of 50 MWh in one hour corresponds to a utilization rate of 0.5.

Net load Demand (also referred to as load) is the single-most important determinant of electricity generation. Specifically, fossil units serve the residual demand, or net load, calculated as total load minus electricity generation by renewables, nuclear energy, and biomass. According to the merit-order principle, fossil units will only generate electricity if net load is positive, i.e., whenever generation from renewables, biomass and nuclear plants does not suffice to meet total load.

Marginal costs Competitive bidding in the day-ahead market implies that units supply electricity at (or just above) their marginal cost. Hence, marginal costs are critical as they determine a unit's position in the merit curve. We construct marginal cost estimates for fossil fuel units (in €/MWh) as the sum of (i) variable operating costs, (ii) fuel costs, (iii) costs of ETS allowances needed to cover emissions, and (iv) transportation costs. Variable operating costs of steam turbines, gas turbines, and combined cycle turbines are taken from Lundy (2019). Fuel costs for natural gas and lignite are directly given in €/MWh. We assume a lignite price of €6/MWh as in Egerer (2016) and Jarvis et al. (2022).[8] The hard coal market price is converted from euro per metric ton to €/MWh using conversion factors as in Jarvis et al. (2022). Costs to acquire emission allowances per MWh are calculated using conversion factors from Jarvis et al. (2022). In the case of hard coal, we also add transportation costs in €/MWh. The sum of these components yields the cost of one MWh of primary energy. To convert this into the cost of generating one MWh of electricity, we multiply it with the inverse of the unit's net electric efficiency, which is provided by the JRC Open Power Plants Database. This gives us marginal cost estimates for all fossil-fuel powered units, and provides the basis for constructing empirical merit curves.

[8]There is no market price for lignite due to the prevailing vertical integration of lignite mining and electricity generation.

Capacity gap The capacity gap serves as a normalized metric that quantifies the relative position of a unit in the supply curve. For any given unit, the capacity gap is defined as the difference between the aggregate capacity of all other units that have lower marginal costs and the net load, divided by the net load. A unit with a negative capacity gap should be active as the aggregate capacity of all units with lower marginal costs falls short of residual demand, whereas a unit with a positive capacity gap should not be. For example, the first unit in the merit curve is the one with the lowest estimated marginal cost and has a capacity gap of -1. As the least-cost producer in that hour, the unit should be switched on and generate electricity. The second unit in the merit curve serves net load not yet served by the first unit, so its capacity gap is closer to zero. As we move up the merit curve, the capacity gap decreases in absolute terms and, eventually, switches sign from negative to positive at some unit. If all previous cheaper units actually run at full capacity, this unit will be the marginal unit required to meet net load. The capacity gap is helpful because it summarizes both whether or not a plant is "in the money", whether or not it is marginal, and how far it is from being marginal.

Border points If there was no cross-border transmission, the marginal cost of the last unit to satisfy demand would establish the wholesale electricity price. Since we have to account for cross-border transmission, we add border points to our sample and treat them as separate units. Consider the transmission lines between bidding zone A and B. We model the transmission lines from A to B as an additional power unit in bidding zone B because it can supply electricity to B. Likewise, we model transmission from B to A as an additional power unit in bidding zone A. The marginal cost of A's border point is determined by the hypothetical pre-trade equilibrium price in bidding zone B, and vice versa. If electricity in B was cheaper than in A without cross-border transmission, the border point in A would supply electricity because the market clearing mechanism would activate additional units in B to produce electricity for A and shut off some units in A, which must be more expensive than those in B, lowering overall electricity costs.

Because hypothetical pre-trade prices are unobserved, we do the following to calculate marginal costs at border points. First, we take only a bidding zone's domestic units, estimate marginal costs, build the merit curve, and then calculate the capacity gap. We then look for the unit that is sits at the margin, that is the unit that is the first with a positive capacity gap, and consider its marginal cost as our estimate for the hypothetical pre-trade price. We then impute this hypothetical price as the marginal cost for all the border points that direct electricity from this bidding zone to a neighboring zone. Once we have imputed marginal costs for all border points, we reconstruct the merit curves in every hour of sample in every

bidding zone again, but this time with the border points included. In other words, we treat border points like domestic units now. The same ranking of units, including border points, then provides the estimate for the merit curves in bidding zones with cross-border flows included, which is what the European Market Coupling Company actually calculates and what determines market outcomes.

Output gap Analogous to the capacity gap, we estimate an output gap for each unit, defined as the gap between net load and cumulative generation by all units with lower marginal cost (i.e., those to the left of a given unit in the merit curve). Again, units with a negative output gap should be active because the generation of cheaper units is not sufficient to serve net load. Based on actual output rather than installed capacity, this variable is closer to actual market conditions than the capacity gap which is predicated on the strong assumption that units either run at full capacity or not at all, which is at odds with observed generation levels. Nonetheless, the capacity gap is useful because it is calculated using exogenous factors only, namely net load and a unit's capacity. If discrepancies between capacity and actual output are approximately constant across bidding zones and across different market conditions, the capacity gap still yields important information about a unit's propensity to generate in a given hour.

3.4 Analysis Sample and Summary Statistics

Our analysis sample comprises 610 electricity-generating units in 11 countries (cf. Table 1). To keep track of bidirectional flows between any two adjacent and connected bidding zones, we additionally model 26 border points.[9] Figure 1 shows the geographic location of all units on the map, along with their capacity and existing border points. Because reporting is mandatory only for units with at least 100 MW of installed capacity, our sample comprises about 80% of all fossil electricity generation (82.7% for lignite, 83.7% for hard coal, and 77.5% for natural gas).

Table 2 reports descriptive statistics of the utilization rate at every hour over the sample period from 2016-2021. Figure 3 illustrates how a unit's utilization rate (generated electricity relative to installed capacity) varies with the capacity gap (a) and the output gap (b), differentiated by fuel types lignite, hard coal, and natural gas. A common feature across fuel types in both types is that the utilization rate declines with both the capacity gap as well as the output gap. This is in line with the merit-order principle since higher values of

[9]We model every border point twice, once per direction of flow, and have some additional observations due to the split-up of the German-Austrian-Luxembourgian bidding zone into an Austrian and a German-Luxembourgian bidding zone in October 2018.

Table 1: Power Plants and Generation Units, by Country and Fuel Type

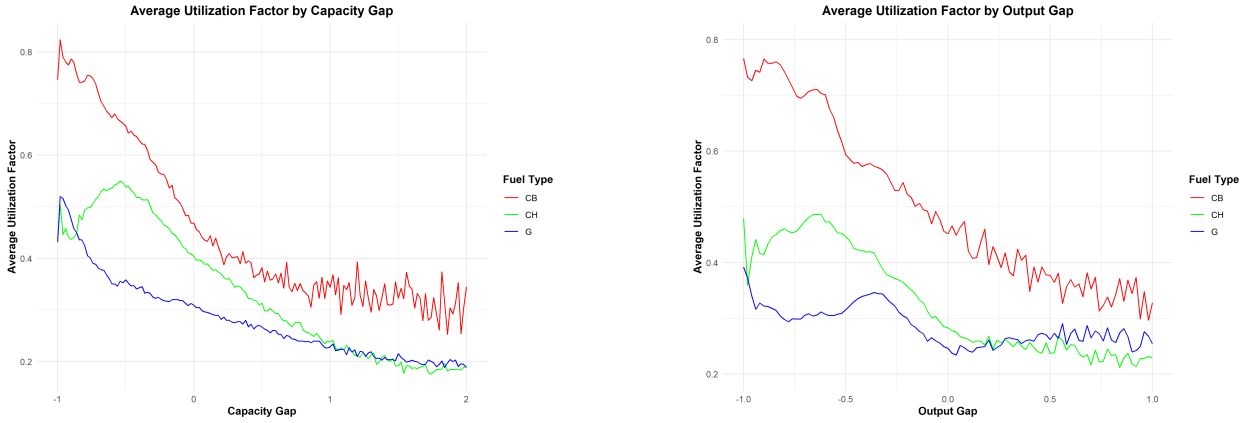
	Generating units					Power plants				
Country	**All**	Lignite	Coal	Gas	Other†	**All**	Lignite	Coal	Gas	Other*
Austria	17		2	14	1	11		2	8	1
Belgium	18		2	16		13		2	11	
Czechia	28	23	4	1		9	7	1	1	
Denmark	13		10	3		11		8	3	
Germany	141	38	53	49	1	78	10	34	28	6
France	24		5	19		21		5	16	
Italy	110		18	92		64		6	57	1
Netherlands	41		8	33		28		6	21	1
Poland	115	31	79	5		32	5	22	5	
Sweden	4		1	3		3		1	2	
Spain	99		24	72	3	49		16	30	3
Total	610	92	206	307	5	319	22	103	182	12

Notes: †"Other" denotes units that switch fuel during the observation period. One Austria unit switches from coal to gas, one German unit switches from lignite to gas, and three Spanish units switch from lignite to coal.
*"Other" refers to one Italian, one Dutch, two Spanish, and four German plants that operate both coal and gas units, and one German plant that operates both lignite and gas units. Additionally, one Austrian, one German, and one Spanish plant contain units that switch fuels (the "Other" units in the left part of the table).

Table 2: Utilization Rate by Fuel: Descriptive Statistics

	Mean	SD	Min	p25	Median	p75	Max
Total	0.410	0.386	0	0	0.428	0.804	1
Lignite	0.696	0.319	0	0.542	0.833	0.939	1
Hard Coal	0.430	0.372	0	0	0.506	0.775	1
Natural Gas	0.358	0.383	0	0	0.217	0.765	1
Border In Points	0.279	0.341	0	0	0.099	0.524	1

Figure 3: Utilization Rate, Capacity Gap, and Output Gap



(a) Utilization rate and capacity gap by fuel type

(b) Utilization rate and output gap

Note: The figures plot the average utilization rates at lignite (CB), hard coal (CH), and natural gas units against the unit's capacity gap (a) or output gap (b). The capacity gap of a unit is the difference between aggregate installed capacity at fossil fuel units with lower estimated marginal cost and net load, divided by net load. A negative capacity gap indicates that net load is larger than what all cheaper units could generate at a maximum so that the unit should run. A positive capacity gap indicates that net load is smaller than what all cheaper units could generate at a maximum. Net load could be met without the unit so that it may not run. The cheapest fossil fuel unit in a given hour has a capacity gap of minus one. At a capacity gap of zero, the installed capacity at cheaper units equals net load. The output gap is defined analogously for actual output instead of installed capacity. A negative output gap indicates that the cheaper units do not generate enough electricity to satisfy net load so that the unit should run. A positive output gap indicates that the cheaper units generate more electricity than net load so that the unit should not run.

either gap correspond to the unit being more expensive relative to other units. However, the utilization rate does not drop all the way to zero, even for high values of the gap. This strongly suggests that the merit-order principle is not the unique determinant of generation decisions. A qualitative difference is that lignite-fired units tend to have higher utilization rates than the other fuels. Since lignite units are among the cheapest fossil fuel units in the market, they tend to have utilization rates of more than 70%, often operating near their full capacity when they are in the money. This suggests that lignite units serve base load and are not frequently ramped up and down. Ramping costs also matter for hard coal units, which could explain why the utilization rate is not maximized under price conditions that push those units to the extreme left of the merit order.

In the random forest estimation, we include both the capacity gap and the output gap for a very practical reason. When predicting a unit's utilization rate under counterfactual marginal costs, the counterfactual output gap is unknown because it follows from the very object we want to estimate. In contrast, the counterfactual capacity gap can always be estimated because it only relies on counterfactual marginal costs and installed capacity, both

of which we know. We still imply the output gap, starting with the observed values at actual allowance prices and marginal costs, and then iteratively apply the random forest model, using the implied counterfactual output gap from the last iteration's predicted utilization rates as a predictor for the next iteration. In practice, three iterations suffice to achieve a correlation coefficient above 0.99 between the final two output gap predictions.

In a nutshell, the utilization rate is a function of the marginal cost structure of the unit as well as of other units in the market. These marginal costs critically depend on fuel and allowance prices. The merit curve is established on the day-ahead market so that, if operators take today's prices as the best predictor for tomorrow's prices, their bids for tomorrow must be based on today's prices. However, there is also short-term dispatch that is particularly relevant when net load is unexpectedly high, outstripping yesterday's expectation for today's net load. Since short-term dispatch will depend on contemporaneous prices, we include those in the model as well. We also include dummies for hour-of-the-day, day-of-the-week, and month-of-the-year. Year fixed effects would broadly cover common trends across all countries such as the roll-out of renewable energies, however, this is already covered through net load. We drop observations during which a unit is labeled as unavailable in the ENTSO-E data, but keep those during which a unit is only partly unavailable and account for this through a dummy variable for such partial unavailability. Finally, we include a factor variable for production type (lignite, hard coal, or natural gas), a factor variable for bidding zone, a dummy for whether a unit is a combined heat-and-power unit, a factor variable for technology type, and a unit's net electric efficiency. Border units receive their own production type, their own technology type, and a net electric efficiency of one. Table 3 provides an overview of all the predictors that we use in the Quantile Regression Random Forest to predict the utilization rate.

4 A Model of Electricity Dispatch

We propose an empirical model of European electricity markets that captures both merit-order driven dispatch and the limitations to it, which derive from transmission constraints within and between bidding zones, re-dispatch, intermittent renewable energy supply, and other sources of congestion. Our approach is to embed the structural determinants in a machine learning model capable of predicting hourly electricity generation at each unit. In what follows, we present the quantile regression forest (QRF) model and assess its predictive power. We also explain how the model is used to predict generation in the counterfactual experiments we are interested in.

Table 3: Predictors Used in the Random Forest Algorithm

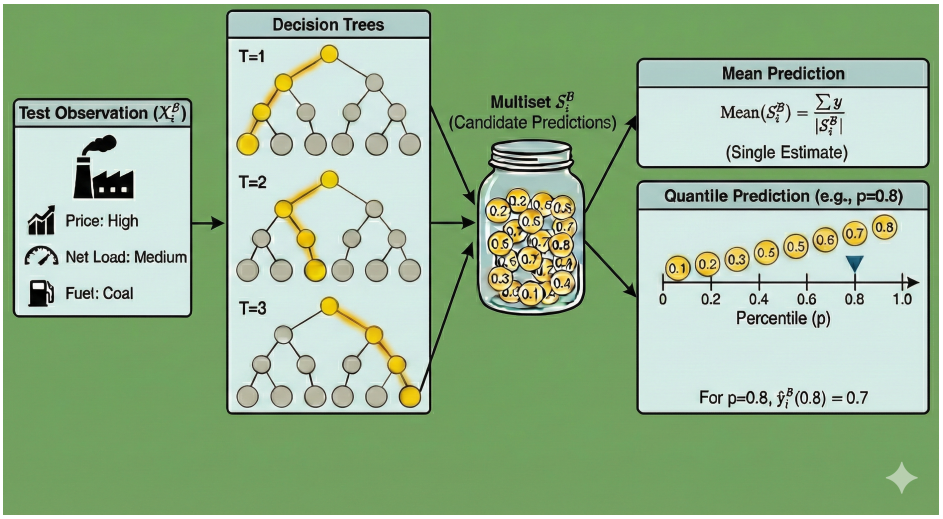
Variable Name	Explanation	Type
Hour-of-the-day	0, 1, 2, ..., 23	Factor
Day-of-the-week	Mo-Su	Factor
Month-of-the-year	Jan-Dec	Factor
Net load	(Total load) - (Generation by renewable, nuclear, and biomass units)	Numeric
Price of hard coal	API2 Coal Index	Numeric
Price of natural gas	Price at national trading hub	Numeric
ETS allowance price	Daily settlement price	Numeric
Brownspread	Price of hard coal minus price of natural gas	Numeric
Production type	Lignite, hard coal, natural gas, or border in point	Factor
Bidding zone	A geographic region with a uniform wholesale electricity clearing price	Factor
Net electric efficiency	Conversion rate: units of electric energy per unit of primary energy	Numeric
Technology	Steam turbine, gas turbine, combined cycle, or unknown	Factor
CHP	Yes, No, or unknown	Dummy
Marginal cost	Marginal cost estimate of unit	Numeric
Capacity gap	(Cum. capacity of units with lower marginal costs / net load) - 1	Numeric
Output gap	(Cum. generation by units with lower marginal costs / net load) - 1	Numeric
Partly unavailable		Dummy
Lat	Latitude of the unit	Numeric
Lon	Longitude of the unit	Numeric

4.1 Quantile Regression Forest

We use a quantile regression random forest to predict electricity generation at fossil fuel powered units. To understand random forests, let us start with a decision tree. Suppose we have an outcome y that we want to predict using K predictors $\mathbf{X} = (X_1, X_2, \ldots, X_K)^T$. In the case at hand, the functional relationship between a power unit's utilization rate y and predictors $\mathbf{X}$ is highly unlikely to be linear. The merit order, for example, establishes a clear threshold for whether a unit should produce at all. Decision trees such as the one displayed in Figure 4 are a useful tool in such cases. Initially, all observations of Y are placed in the root node. At subsequent decision nodes, observations are split so as to reduce the variance in one of the predictors X_k for $k \in \{1, ..., K\}$. For example, if all units with $X_k < 0$ have some similar values of y, and all units with $X_k \geq 0$ have other similar values of y, then a reasonable split is on whether $X_{i,k} < 0 \ \forall i$. This would, for instance, put all units with $X_k < 0$ into one internal node and all units with $X_k \geq 0$ into the other internal node. At every internal node, additional splits occur to further reduce variance. Splitting stops at the leaf nodes, the level of which is usually determined by some stopping criteria. Common such criteria are (i) that a further split would not achieve a minimum level of variance reduction, or (ii) that the number of remaining observations in a node would be below a predetermined minimum number.

Single decision trees generally have low bias because they can model complex patterns

Figure 4: Flow Chart of the Quantile Random Forest Model



very closely but may exhibit high variance as small changes in the data may lead to a very different tree. To reduce variance, researchers often apply bagging. This procedure picks a random subset of observations with replacement from the original training data set and constructs decision trees based on these random subsamples. The set of decision trees is then called a "forest", and predictions become the average across decision trees. Because each decision tree is based on a different random subsample of the training data, bagging will reduce variance without introducing bias. While bagging is effective for stabilizing high-variance predictions that are sensitive to small changes in the training data, high correlation between trees may still persist if a few features are strong predictors and dominate every tree. Random forests account for this by adding an extra layer of randomness by selecting a random subset of features at each split. This is commonly referred to as "feature bagging" because we restrict $\mathbf{X}$ to a random subsample at each split. A common choice is to use only a random third of all predictors at every split.

We use our data set to train the model and obtain our random forest with trees $t = 1, ..., T$. Every individual tree t is constructed such that, given the bagged subsample of the data and the random subset of predictors, the prediction of y^A given $\mathbf{X^A}$ has minimal uncertainty according to some variance metric.

Next, we apply the model on the data. This provides us with a predicted outcome $(\widehat{y})$ along the actual outcome (y) that we know. Consider observation $(y_i, \mathbf{X_i})$ from our data set. To obtain a model prediction $\widehat{y}_i$ based on $\mathbf{X_i}$, we drop $\mathbf{X_i}$ down each tree t to land in a terminal node n. Terminal node n_t of tree t contains a set of outcomes, $\{y^A_{n_t,t}\}$. Since we have T trees, we obtain T such sets. Their multiset union $\{y^A_{n_1,1}\} \uplus \{y^A_{n_2,2}\} \uplus ... \uplus \{y^A_{n_T,T}\} = S_i$ then contains all the outcomes from all the terminal nodes $n_1, ..., n_T$. In a way, one can think

of all the elements in S_i as "candidates" for the value of $\widehat{y}_i$ given $\mathbf{X_i}$.

The most straightforward estimate $\widehat{y}_i$ of y_i is the mean value of S_i, that is the arithmetic mean of all the values in the terminal nodes. However, as will be useful later when taking into account equilibrium constraints in the energy market, the QRF recovers the entire conditional distribution of possible predictions instead of just the mean value (Jarvis et al., 2022). Thereby, we can also calculate the predicted outcome $\widehat{y}_i(p)$ at a given percentile. For instance, for $p = 0.8$, we would order all the elements in S_i in ascending order and then pick the element that is larger than 80% of the other elements. In algebraic terms, that is

$$\widehat{y}_i(p) = \inf \left\{ y \in S_i : \frac{1}{|S_i|} \sum_{z \in S_i} I(z \leq y) \geq p \right\}$$

.

4.2 Percentile Choice

For every observation i, the QRF delivers a set S_i with all the candidate values for its predicted utilization rate $\widehat{y}_i$. The function $\widehat{y}_i(p)$ for $p \in [0, 1]$ then picks the value in S_i corresponding to the percentile p. For instance, if S_i contains 1,000 elements ranked in ascending order, $\{s_{i(1)}, s_{i(2)}, ..., s_{i(1000)}\} = S_i$, and we choose $p = 0.5$, we obtain as the predicted value of y for observation i, $\widehat{y}_i = s_{i(500)}$. Which value of p yields the most accurate prediction of y_i? Since we know y_i, we could choose an individual p_i $\forall i$ so that the resulting prediction $\widehat{y}_i(p_i)$ is as close to y_i as possible. Algebraically, the optimal percentile choice p_i^* is given by

$$p_i^* = \arg \min_{p \in [0,1]} |\widehat{y}_i(p) - y_i|$$

The choice of p_i^* also matters because the random forest does not incorporate the constraint that predicted utilization rates must be such that we remain in a market equilibrium. Because we leave electricity demand unaltered, predicted total electricity supply must equal actual supply. We incorporate this constraint indirectly by choosing the appropriate percentiles.

While solving for every p_i^* individually for every observation i in the data set is an approach that would satisfy the market equilibrium constraint, this is an infeasible strategy for a data set containing more than 600 units for which we track hourly utilization rates across six years. Instead, we do the following. First, we calculate predicted utilization rates $\widehat{y}_i(p)$ for $p \in \{0.1, 0.2, ..., 0.9\}$, between the 10th and 90th percentile in steps of 10, and predicted output $\widehat{Y}_i(p)$ at those same percentiles by multiplying the predicted utilization

rate $\widehat{y}_i(p)$ with the unit's installed electric capacity κ_i,

$$\widehat{Y}_i(p) = \widehat{y}_i(p) \cdot \kappa_i \quad \forall p \in \{0.1, 0.2, ..., 0.9\}$$

Next, we aggregate $\widehat{Y}_i(p)$ for all units i in bidding zone b, running on fuel f, and during month-of-sample m,

$$\widehat{Y}_{b,f,m}(p) = \sum_{i \in I_{b,f,m}} \widehat{Y}_i(p) \quad \forall p \in \{0.1, 0.2, ..., 0.9\}$$

to predict the p-th percentile of total generation by zone $\times$ fuel $\times$ month. We search for the percentile p that equates this measure to the actual level of total generated electricity $Y_{b,f,m}$ by imposing the market equilibrium condition on the zone $\times$ fuel $\times$ month level.

Because we calculate $\widehat{Y}_{b,f,m}(p)$ only for $p \in \{0.1, 0.2, ..., 0.9\}$, we almost surely have $\widehat{Y}_{b,f,m}(p) \neq Y_{b,f,m} \ \forall p \in \{0.1, 0.2, ..., 0.9\}$. To achieve equality, we linearly interpolate values of $\widehat{Y}_{b,f,m}$ between two adjacent percentiles. Algebraically, for every pair of adjacent percentiles $\left(p_{(z)}, p_{(z+1)}\right)$ for which $\widehat{Y}_{b,f,m}(p_{(z)}) < Y_{b,f,m}$ and $\widehat{Y}_{b,f,m}(p_{(z+1)}) > Y_{b,f,m}$, we look for the weight $\alpha_{b,f,m} \in (0, 1)$ for which

$$\alpha_{b,f,m} \cdot \widehat{Y}_{b,f,m}(p_{(z)}) \ + \ (1 - \alpha_{b,f,m}) \cdot \widehat{Y}_{b,f,m}(p_{(z+1)}) = Y_{b,f,m}$$

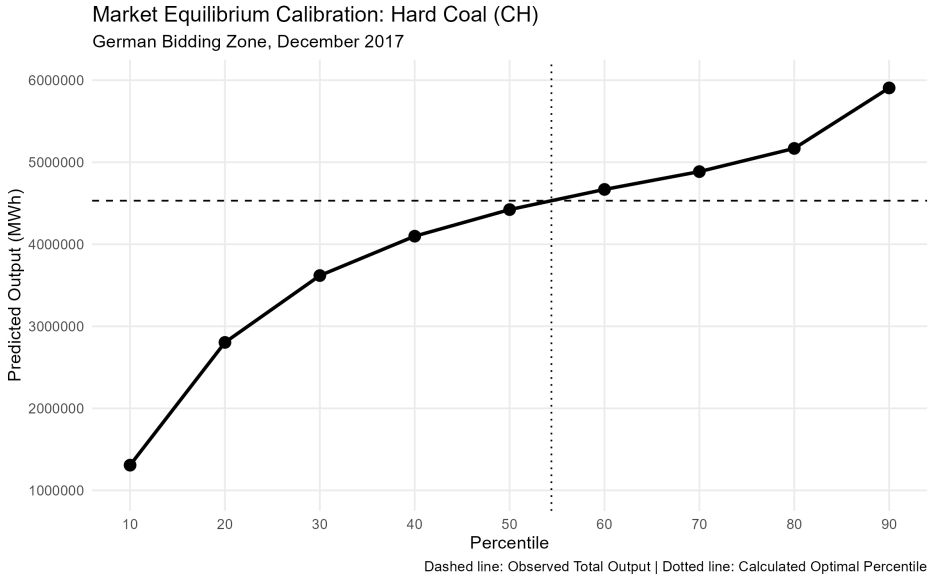
The according percentile that establishes the market equilibrium on the zone $\times$ fuel $\times$ month level is then given by $p^*_{b,f,m} = \alpha_{b,f,m} \cdot \widehat{Y}_{b,f,m}(p_{(z)}) \ + \ (1 - \alpha_{b,f,m}) \cdot \widehat{Y}_{b,f,m}(p_{(z+1)})$. Figure 5 depicts one example of how we pick the percentile that establishes the market equilibrium.

4.3 Predictive Quality of the Quantile Random Forest

Figure 6 provides an overview of the importance of predictors in the QRF model. The importance score is the sum of all variance reductions that are achieved at all the splits based on the respective variable. In general, variables with higher explanatory power and with more split points obtain a higher importance score. As Figure 6 shows, a unit's latitude and longitude are the most important predictors for a unit's utilization rate. These two predictors contribute to about 22.4% of all variance reduction in the predicted utilization rates.

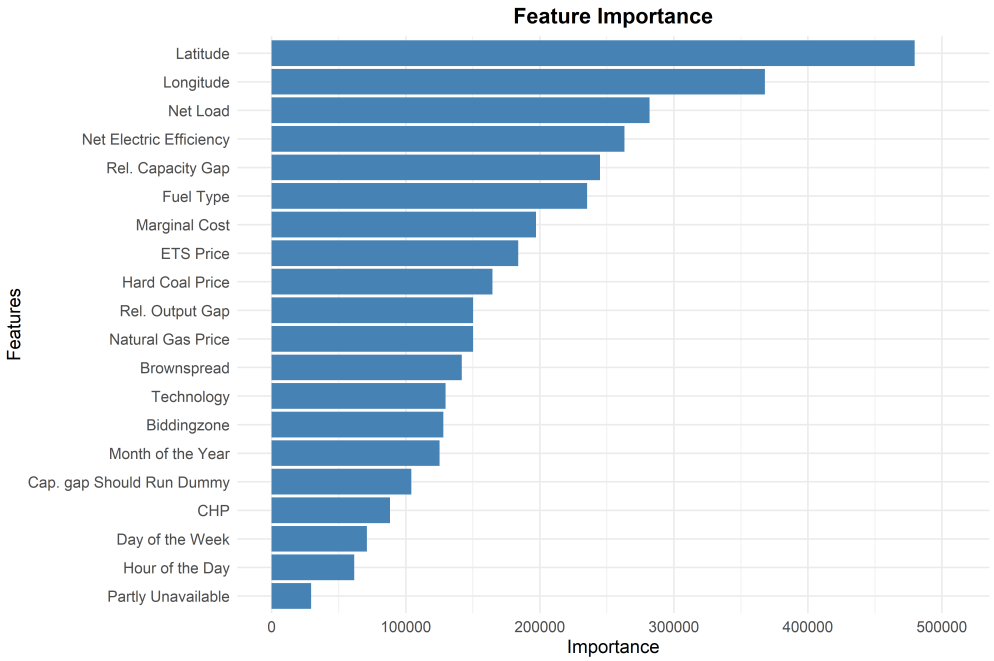
The next most important variables are net load and capacity gap. These variables directly relate to the merit order and energy market equilibrium. A high capacity gap denotes that the given unit is more likely to be active and generate relative to other units, whereas high net load denotes that every unit is more likely to run because overall electricity demand is

Figure 5: Predicted Output of Hard Coal Power Units in the German-Austrian-Luxembourgian Bidding Zone in December 2017



Notes: Hard-coal fired power units contained in the test data and located in the German-Austrian-Luxembourgian bidding zone jointly produced 4,530,922 MWh during December 2017, as indicated by the dashed horizontal line. The dots indicate deciles of predicted total output for this period, obtained from the quantile regression random forest model. Equilibrium requires that predicted generation by all units add up to total observed output. We use linear interpolation between deciles to satisfy this constraint. As can be seen in the graph, total output lies between the 5th and 6th decile. The 50th percentile under-predicts output (above 4,400,000 MWh) whereas the 60th percentile over-predicts it (short of 4,700,000 MWh). Algebraically, $\widehat{Y}_{b,f,m}(0.5) < Y_{b,f,m}$ and $\widehat{Y}_{b,f,m}(0.6) > Y_{b,f,m}$. Therefore, choosing $\alpha_{b,f,m} = 0.4407$ yields a weighted average of the 50th and 60th percentile predictions that brings us exactly to $Y_{b,f,m} = 4{,}530{,}922$ MWh. We also take this as the linearly interpolated prediction at the 54.407th percentile.

Figure 6: Relative Importance Score of Predictors



Notes: The figure shows the importance of the predictors in the One-Stage Random Forest. We use the "impurity" measure for importance scores, which measures the total reduction in variance (impurity) at all splits on the respective predictor across all trees in the random forest. The impurity measure is a unitless score and makes sense only relative to other predictors. For instance, a unit's latitude is the most important single predictor with an importance score of just below 500,000. Longitude, the second-most important single predictor, has a score of about 365,000.

Table 4: Goodness-Of-Fit Measures By Fuel Type

	All Units	Lignite	Hard Coal	Natural Gas	Border Points
	1. Weighted Percentile Prediction				
MSE	0.040	0.045	0.046	0.037	0.028
R^2	0.729	0.554	0.672	0.734	0.730
Corr	0.864	0.785	0.835	0.867	0.860
	2. Mean Prediction				
MSE	0.073	0.063	0.091	0.066	0.064
R^2	0.506	0.379	0.344	0.531	0.383
Corr	0.714	0.625	0.626	0.738	0.622

Notes: The table shows three metrics for quantifying prediction quality of the random forest model: 1. Mean Squared Error, 2. R-squared, and 3. the Correlation Coefficient. This table shows all three metrics for all units, and then for the four unit types (lignite, hard coal, natural gas, and border points). We compare the prediction at the weighted percentile to maintain a market equilibrium with the mean prediction. Numbers relate to one-fold cross validation, using 50% of hours-of-sample to train the model and assessing it using the other 50% (test data).

high.

Latitude and longitude, in contrast, are not connected to the merit order principle. A unit's geographic location within a bidding zone should be irrelevant in a frictionless market without local transmission constraints. However, the high importance of latitude and longitude indicates that important deviations from the hypothetical copperplate interferes with the merit order principle and significantly distorts generation away from what would happen in a frictionless market.

To assess the predictive quality of our model, we conduct out-of-sample validation. We randomly partition the 52,608 hours-of-sample (24 hours $\times$ 365 days $\times$ six years, plus two leap days in 2016 and 2020) fifty-fifty into equal-sized training and test sets. Following standard practice, we build the model on the training data, and then assess its predictive quality on the out-of-sample test data.

Table 4 shows the performance of the QRF algorithm to predict utilization rates in the test set. The upper panel presents goodness-of-fit measures using our weighted percentile predictions that keep aggregate generation constant. The MSE of predicted relative output is 0.040, with an R^2 of 0.729, and a correlation coefficient of 0.864. These metrics highlight the good predictive quality of our model, especially when considering that the utilization rate is continuous and bounded within the $[0, 1]$ interval with many zero and near-one observations. This in turn requires a great deal of non-linearity for the model to correctly predict utilization

Table 5: Goodness-Of-Fit Measures By Country (sorted by MSE, ascending)

	AT	SE	DE	BE	DK	IT	FR	NL	ES	CZ	PL
	Weighted Percentile Prediction										
MSE	0.016	0.017	0.023	0.026	0.029	0.037	0.038	0.043	0.044	0.047	0.067
R^2	0.879	0.831	0.852	0.830	0.758	0.715	0.761	0.710	0.627	0.649	0.441
Corr	0.940	0.915	0.926	0.915	0.878	0.857	0.880	0.857	0.810	0.826	0.720
	Mean Prediction										
MSE	0.061	0.051	0.060	0.066	0.087	0.073	0.066	0.075	0.088	0.073	0.076
R^2	0.531	0.492	0.618	0.574	0.272	0.439	0.585	0.498	0.249	0.451	0.363
Corr	0.773	0.710	0.794	0.761	0.554	0.672	0.778	0.712	0.517	0.683	0.615

Notes: The table shows three metrics for quantifying prediction quality of the model: 1. Mean Squared Error, 2. R-squared, and 3. the Correlation Coefficient, separately for units in a given country. We compare the prediction at the weighted percentile to maintain a market equilibrium with the mean prediction. Numbers relate to one-fold cross validation, using 50% of hours-of-sample to train the model and assessing it using the other 50% (test data).

rates.

The lower panel shows the goodness-of-fit measures if we simply used the model's mean predictions, disposing of the model's Quantile Regression part and disregarding equilibrium violations. Clearly, the weighted percentile predictions outperform the mean predictions: The MSE of the latter is almost twice as large, with a particularly pronounced difference for natural gas units and border points, for which the weighted percentile predictions work particularly well.

Comparing performance measures across types of units, the model works particularly well for natural gas units and border points, whereas it struggles a bit more with hard coal and lignite units. The latter have a substantially low R-squared of 0.554, which may reflect that lignite units are relatively sluggish and therefore tend to run either near full capacity or not at all. This would suppress R^2 because actual utilization rates vary relatively little with predictors. The relatively more flexible hard coal units may also be hard to predict, yet R^2 will be higher if the model still captures general generation patterns reasonably well.

We further document substantial differences in model performance across countries. Table 5 shows performance measures by country, by MSE in ascending order. One can see that the model mostly performs very well for small countries with high-capacity interconnections to other countries. Units in Austria, Sweden,[10] Belgium and Denmark have small MSEs of below 0.03.

[10]Sweden is considered "small" in this context because only its two southern-most bidding zones contain fossil fuel power units.

The model also works well for Germany, a country characterized by its decentralized geography, high degree of interconnectedness with other countries, and a very diverse energy mix. Therefore, price fluctuations plausibly induce large changes in generation patterns from which the model can learn.

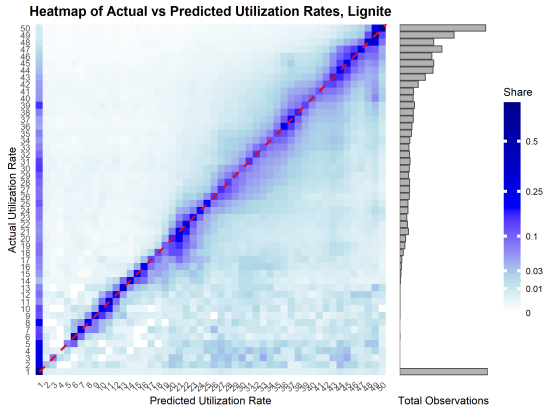
Conversely, the model struggles more in countries that rely on fewer fuel types. The Netherlands, albeit a small and interconnected country, relies heavily on natural gas to serve base load, in contrast to most other countries where natural gas is a peak demand fuels.

In Spain, the highly uneven population density coupled with insufficient transmission capacities plausibly causes significant friction in the Spanish energy market, weakening the predictive power of the merit-order principle. Lastly, Poland and Czechia depend on lignite and hard coal to a much higher degree than any other country in the sample. Accordingly, the marginal units in both Poland and Czechia are mostly coal units even under near-zero allowance prices, which is a phenomenon unique to these countries that the model might struggle to fully internalize.

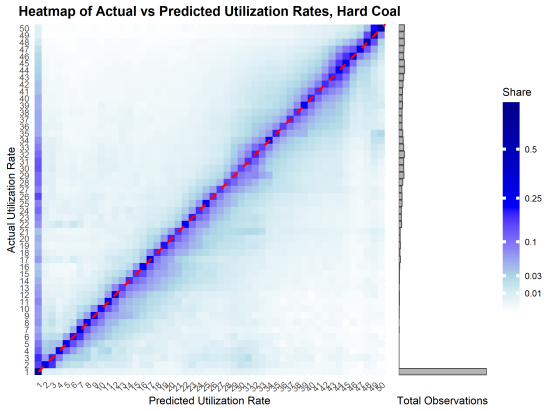
Next, we compare the predictions of units' utilization rates with its actual levels. Figure 7 presents heatmaps of actual (y-axis) vs. predicted (x-axis) rates in which we bin the outcome in bins of 0.02. The first bin therefore corresponds to utilization rates between zero and two percent. Every row shows the distribution of predicted utilization rates within the respective bin of actual rates, so that the graph should be read row-to-row. Figures 7a to 7d show the heatmap for lignite units (7a), hard coal units (7b), natural gas units (7c) and border inflows (7d). The heatmaps provide us with some important insights into the model's performance. First, the model predicts utilization rates with a high degree of accuracy as most observations lie on or very close to the 45 degree line. Second, there is a tendency to predict near-zero utilization rates among fossil fuel units even when actual rates are much higher. This happens along upper-medium rates of 50 to 80% for lignite and coal units, and along the entire range of utilization rates for natural gas units, whereas this pattern does not occur for border points. Beyond that, we rarely observe the opposite that predicted rates are far from zero while actual rates are near-zero.

Third, the model performs very well when actual utilization rates are near zero or near one. The model correctly predicts utilization rates of below 10 percent 76.4% of the time for lignite units, 84.1% for hard coal units, 91.4% for natural gas units, and 87.2% for border points. For actual utilization rates north of 90 percent, the model achieves a correct prediction 82.0% of the time for lignite units, 75.1% for hard coal units, 70.4% for natural gas units, and 66.9% for border points. Fourth, overall model precision is higher for natural gas units and border points than for lignite and hard coal units. Predictions are scattered more closely around the 45 degree line for gas units and border points, whereas they spread

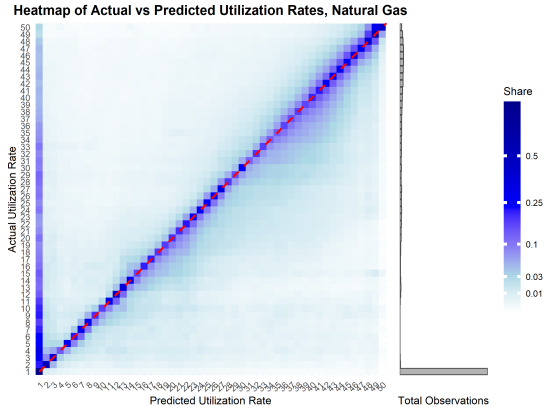
Figure 7: Actual vs. Predicted Utilization Rates, Weighted Percentile Predictions



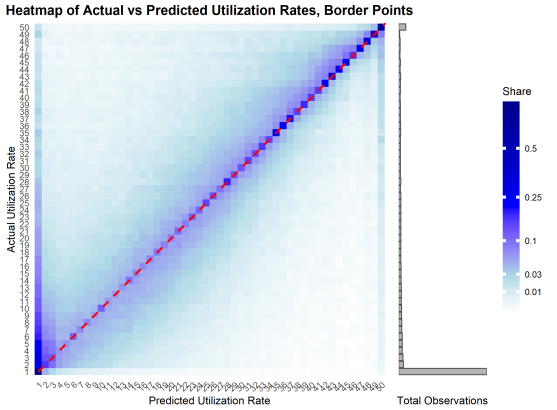
(a) Actual (y-axis) vs. Predicted (x-axis) Utilization Rates, One-Step RF, Lignite Units



(b) Actual (y-axis) vs. Predicted (x-axis) Utilization Rates, One-Step RF, Hard Coal Units



(c) Actual (y-axis) vs. Predicted (x-axis) Utilization Rates, One-Step RF, Natural Gas Units



(d) Actual (y-axis) vs. Predicted (x-axis) Utilization Rates, One-Step RF, Border In Points

Notes: The graphs show a heatmap of actual utilization rates (y-axis) and predicted utilization rates (x-axis) using the weighted percentile predictions. Utilization rates are split into fifty bins in steps of two percentage points. The heatmaps are such that one row adds up to one. In other words, we ask "what is the distribution of predicted utilization rates when actual utilization rates are in bin x". Darker colors indicate a higher density of predictions within the respective row. The bars to the right of the heatmap show the number of observations that contribute to every row. The graphs relate to one-fold cross validation, using 50% of hours-of-sample to train the model and assessing it using the other 50% (test data).

out a bit more for lignite and hard coal units. In addition to that, we have a notable quantity of cases in which predicted rates are substantially above actual rates for lignite units.

We also compare the histograms of actual and predicted utilization rates in Figure B.1 in the Appendix. The alignment between the actual and predicted histograms suggests that the model effectively recovers the underlying empirical distribution of utilization rates. Individual prediction errors appear to be non-systematic and tend to cancel out in the aggregate, ensuring that the predicted frequency of utilization rates remains consistent with observed market behavior.

4.4 Counterfactual Analysis

As described above, the set of percentiles $\mathbf{P}^* = \{p^*_{b,f,m}\}$ are chosen such that predicted electricity generation by all units equals aggregate electricity generation observed in the data. This subsection explains how we use the model to predict electricity generation in counterfactual scenarios that assume different prices of EU ETS permits, known as European Union Allowances (EUAs).

Across all counterfactuals, we keep net load unchanged in the counterfactual scenario. This is a simplifying assumption with two important implications. First, electricity demand has a zero price elasticity. That is, while fluctuations in the EUA price may be passed through to electricity prices (Fabra and Reguant, 2014), we assume that consumers do not adjust their demand in response to such electricity price shifts. Second, keeping net load unchanged also implicitly assumes that non-dispatchable power units would have generated the same amount of electricity in the counterfactual. This departs from other research that explicitly considers counterfactual changes to non-dispatchable generation, in particular from renewables.[11] We believe this is a tenable assumption because, unlike a structural shock to the entire grid capacity (such as a nuclear phaseout), the EU ETS changed electricity prices only gradually via a market-based scarcity signal that was too muted and volatile to change long-term investments in the power plant fleet on a broad scale.

A counterfactual change to the EUA price changes the marginal costs of power units and their ordering in the supply curve. Consider observation $i : (y_i, \mathbf{X_i})$ and denote by $\left(y_i^{cf}, \mathbf{X_i^{cf}}\right)$ its counterfactual equivalent where $\mathbf{X_i^{cf}}$ is observable. Our predictors that relate to the allowance price are now at their new, counterfactual values, while the other predictors remain at their original values. As before, we drop down $\mathbf{X_i^{cf}}$ every tree, obtain a candidate

[11] For instance, Jarvis et al. (2022) analyze the German exit from nuclear energy during the 2010s and assume that counterfactual generation from renewable sources would have been lower without the nuclear phaseout, because keeping nuclear energy in the fuel mix would have resulted in lower electricity prices, attenuating investments in renewable energies.

set S_i^{cf}, from which we obtain the prediction for y_i^{cf} as a function of the percentile p:

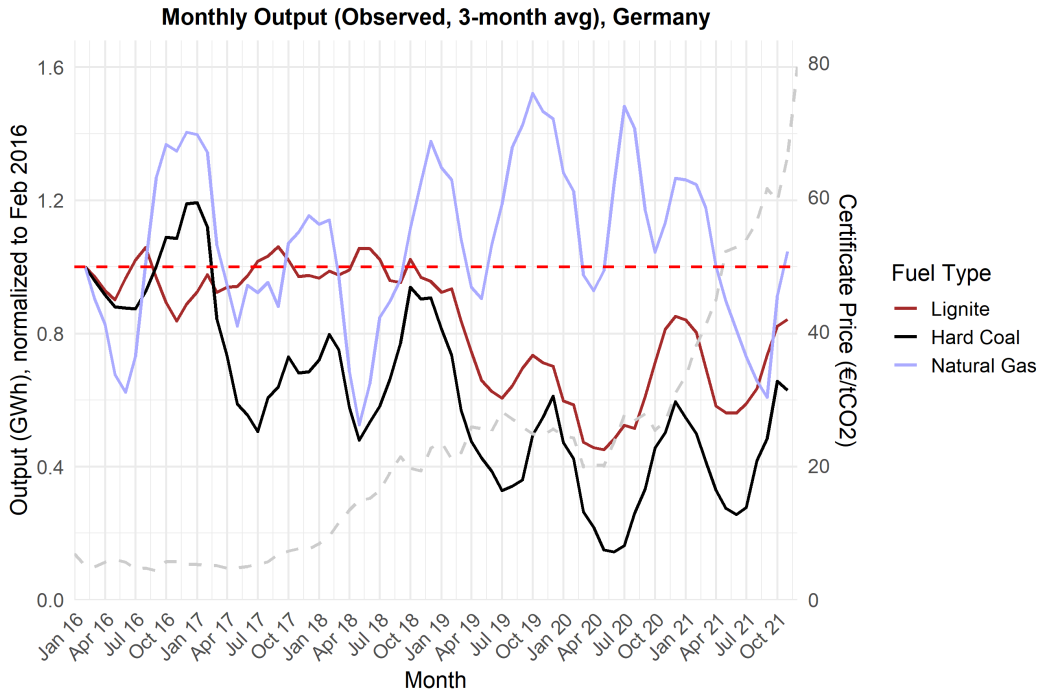
$$\widehat{y_i}^{cf}(p) = \inf \left\{ y \in S_i^{cf} : \frac{1}{|S_i^{cf}|} \sum_{z \in S_i^{cf}} I(z \leq y) \geq p \right\}$$

The natural question arises how we choose p in the counterfactual scenario. We cannot choose the counterfactual percentiles $\{p^{*\ cf}_{b,f,m}\}$ on the zone $\times$ fuel $\times$ month level as we did before under actual prices because (i) we cannot solve for $p^{*\ cf}$ by setting $\widehat{Y}^{cf}_{b,f,m}(p) = Y^{cf}_{b,f,m}$ as we do not observe $Y^{cf}_{b,f,m}$, and (ii) aggregate generation on the zone $\times$ fuel $\times$ month level changes in the counterfactual.

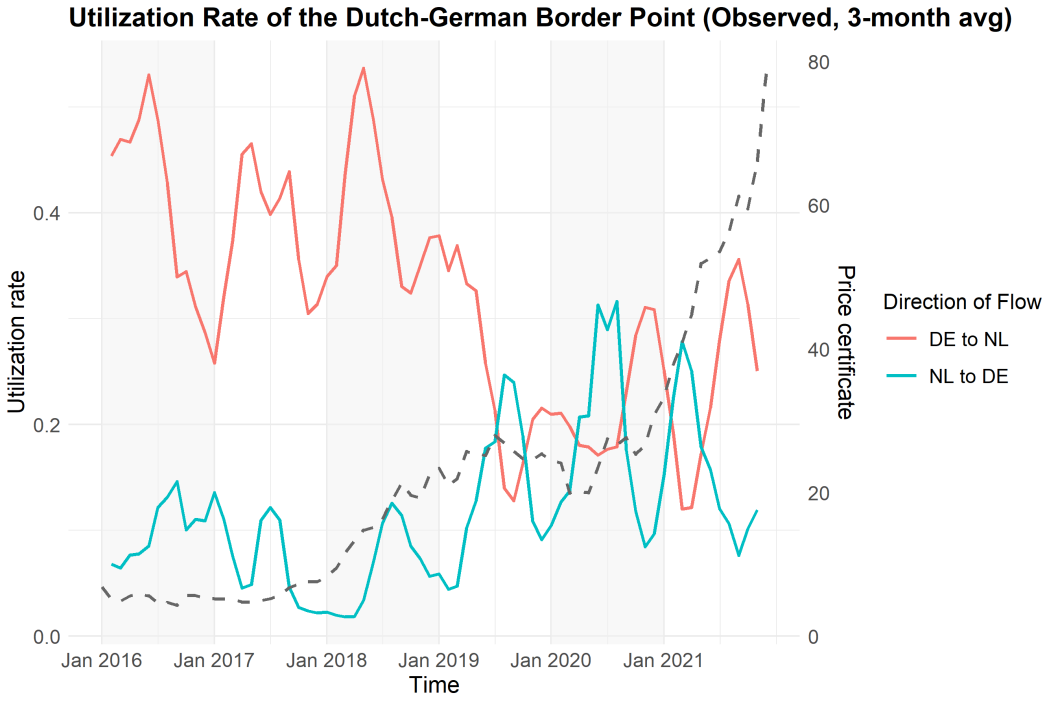
Solving this issue requires a notion of how the counterfactual market equilibrium differs from the observed one. There are two mechanisms at play that swing the European electricity market from one equilibrium to the other. First, there is fuel switching within bidding zones. When the allowance price rises, domestic lignite and coal units tend to become more expensive relative to domestic natural gas units so that electricity firms substitute natural gas for hard coal and ligniteFigure 8a illustrates this for the case of Germany. The figure plots the monthly output at fossil fuel power units by fuel type over time (as a 3-month moving average), normalized against the first observation in February 2016. One can see that generation at natural gas plants is as much as 50% higher in the later years when the allowance price is high, relative to the earlier years when the allowance price is low. The opposite holds true for lignite and coal units, which grow substantially more inactive in later years when allowance prices are higher.

Second, electricity generation relocates between bidding zones based on relative fuel costs and installed capacity. As allowance prices rise, the competitive advantage might shift from coal-heavy zone A to gas-heavy zone B. Once gas becomes the cheaper marginal source, zone B ceases importing and instead exports electricity to meet demand in zone A. We refer to this as 'between-zone relocation'. The Dutch-German border point constitutes a good example of this. With regard to their respective fuel mix, the Netherlands is a gas-heavy country whereas Germany is a lignite- and coal-heavy country. Figure 8b shows the utilization rate of the Dutch-German border point over time, with the red line indicating flows from Germany to the Netherlands and the blue line indicating the opposite. In the earlier years of the sample, when the allowance price was low and lignite and coal units presumably had a competitive advantage over gas, the gas-heavy Netherlands exported very little to coal-heavy Germany but imported a lot. As the EUA prise rose in the later years of the sample, and the competitive advantage presumably reverted, Dutch exports to Germany notably rose while imports from Germany declined.

Figure 8: Changing Generation Patterns And Carbon Pricing



(a) Fuel Switching Within Bidding Zone



(b) Switching Generation Across Bidding Zones

Notes: The figure illustrates the empirical correlation between carbon prices and fuel switching within and across bidding zones. Panel a shows 3-month moving averages of generation [GWh] by lignite, hard coal, and natural gas units in Germany. Panel b plots 3-month moving averages of the utilization rate of the Dutch-German border Point.

Our approach to constructing counterfactual equilibria on the European electricity market builds on these economic mechanisms. Because net load is measured separately per zone, holding net load constant boils down to keeping demand for fossil electricity constant in every bidding zone. Let residual demand (RD) denote the per-zone-demand for electricity that remains net of non-dispatchable generation. Because we further assume that the level of non-dispatchable generation is not affected by differing allowance prices, residual demand in every bidding zone remains constant.

Under actual allowance prices, we know that the combination of observed domestic fossil generation, electricity inflows and outflows is such that it satisfies residual demand in the respective bidding zone. Therefore, the counterfactual equilibrium also emerges from a combination of predicted domestic fossil generation, predicted electricity inflows, and predicted electricity outflows that are such that they add up to the same residual demand in every zone. Since our counterfactual predictions are determined by the percentiles in the conditional distribution of our quantile regression random forest, the counterfactual equilibrium is sufficiently characterized by a set of percentiles $\{p^{cf}_{b,f,m,}\}$ that yield predictions $\widehat{y}_i^{cf}\left(p^{cf}_{g(i)}\right)$[12] that abide by the constraint that every zone's implied predicted residual demand equals its observed one.

One further complication arises from the fact that residual demand does not necessarily equal demand for fossil electricity. Domestic residual demand is the demand that remains net of domestic non-fossil generation. However, domestic residual demand can also be satisfied by foreign non-fossil generation, for example when France imports solar energy from Spain. In order to treat residual demand as demand for fossil energy, we have to make the assumption that border flows are always flows of fossil-fired electricity. In other words, we rule out that renewable and nuclear energy is exported to other zones. Under this assumption, we can calculate the demand for fossil electricity in zone b as the sum of fossil generation in b plus imports of electricity from other zones minus exports of electricity to other zones. Furthermore, we can equate residual demand observed in the data with demand for fossil electricity and solve for the counterfactual equilibrium. This is because, by assumption, residual demand will always be served by fossil fuel units (either at home or through imports).[13]

Since we have 19 interconnected bidding zones, for any given month-of-sample m we have to choose percentiles $\{p^{cf}_{b,f,m}\}$ such that the model predictions at these percentiles are such

[12] $g(i)$ denotes the bidding zone, fuel type, and month corresponding to observation i.

[13] We assess the credibility of this assumption in detail in Appendix C.1. We find that, in most countries, violations of this assumption are very infrequent and not substantive when they occur. The exception to this are the Danish and Swedish bidding zones where hydroelectricity accounts for a substantial share of cross-border flows.

that all 19 implied residual demands equal their observed ones. In more concise notation, we look for $\{p^{cf}_{b,f,m}\}$ so that, for a given month m: [14]

$$\forall b: \quad RD_{b,m} \stackrel{!}{=} \widehat{RD}^{cf}_{b,m} = \widehat{\text{FG}}^{cf}_{b,m}(p^{cf}_{b,f,m}) \; + \; \widehat{\text{Imports}}^{cf}_{b,m}(p^{cf}_{b,f,m}) \; - \; \widehat{\text{Exports}}^{cf}_{b,m}(p^{cf}_{b'\neq b,f,m})$$

On the left-hand side of the equation, we have $RD^{cf}_{b,m}$, observed residual demand RD in zone b during month m. Our predicted counterfactual residual demand $\widehat{RD}^{cf}_{b,m}$ must be equal to that because we keep constant both gross electricity demand and non-dispatchable generation. Given our assumption on fossil border flows, residual demand equals fossil demand, and fossil demand in turn consists of (i) predicted counterfactual fossil generation $\widehat{\text{FG}}^{cf}_{b,m}$ in zone b during month m, (ii) predicted counterfactual electricity imports from neighboring zones to zone b during month m, minus (iii) predicted counterfactual electricity exports from zone b to neighboring zones during month m. This equality must hold for all zones b simultaneously.

Note that because we model all border points twice as border inflows to each of the respective two zones, the last term depends on other zones' percentiles. As an example, predicted residual demand in the Netherlands depends on our predictions for Belgian and German imports because Dutch exports are border inflows for Belgium and Germany, modeled as Belgian and German units with Belgian and German percentiles, respectively. Conceptually, one would have to solve for all percentiles simultaneously that equate predicted counterfactual and actual residual demand in all zones at once. In addition to the high dimensionality of the problem, we do not have a continuous and smooth function that maps percentiles p^{cf}_i to predictions $\widehat{y_i}^{cf}(\cdot)$ because we only know the value of $\widehat{y_i}^{cf}$ at $p = 0.1, 0.2, ..., 0.9$, linearly interpolating inbetween.

Instead of trying to solve directly for the counterfactual equilibrium, we approach the equilibrium through a two-step procedure. As described before, there are two mechanisms that describe the path from the observed equilibrium to a counterfactual equilibrium. First, there is fuel substitution, mainly gas replacing lignite and coal when the allowance price rises, and vice versa when it falls. Second, there is relocation of generation, for instance a coal-heavy region may generate less electricity domestically and import more from abroad when the allowance price rises.

In principle, these mechanisms are intertwined: When a Dutch gas plant replaces a German coal plant, fuel substitution and relocation happen simultaneously. We approximate the counterfactual equilibrium by treating these two mechanisms as separate. We first

[14] Technically, this should hold in every hour-of-sample, but because this is computationally infeasible, we resort to looking for a monthly equilibrium. We tolerate hour-by-hour deviations as long as they balance out over the course of one month.

estimate an "intermediate" solution in which we only consider within-zone fuel substitution and mechanically suppress between-zone relocation of generation. We do so by forcing monthly fossil generation to remain constant in every zone. From this "intermediate" solution onwards, we then focus on between-zone relocation of generation by adjusting our estimates so that they minimize the implied deviations from fossil demand across all zones. Thereby, we allow fossil generation to relocate from some zones to others while keeping total fossil generation across all zones constant. Because we explicitly model and predict border flows, this new equilibrium automatically satisfies cross-zonal transmission constraints.

To estimate the "intermediate" solution in which we only consider within-zone fuel substitution, we consider fossil generation in zone b during month m, which is the sum of lignite L, coal C, and gas G generation:

$$FG_{b,m} = L_{b,m} + C_{b,m} + G_{b,m}$$

We have chosen percentiles $\mathbf{P}^* = \{p^*_{b,f,m}\}$ such that $\widehat{L}_{b,m}(p^*_{b,L,m}) = L_{b,m}$, $\widehat{C}_{b,m}(p^*_{b,C,m}) = C_{b,m}$, and $\widehat{G}_{b,m}(p^*_{b,G,m}) = G_{b,m}$, and therefore, denoting $\mathbf{p^*_{b,m}} = \left(p_{b,L,m},^*, p^*_{b,C,m}, p^*_{b,G,m}\right)^{\mathrm{T}}$

$$\begin{aligned}
\widehat{FG}_{b,m}(\mathbf{p^*_{b,m}}) &= \widehat{L}_{b,m}(p^*_{b,L,m}) + \widehat{C}_{b,m}(p^*_{b,C,m}) + \widehat{G}_{b,m}(p^*_{b,G,m}) \\
&= L_{b,m} + C_{b,m} + G_{b,m} \\
&= FG_{b,m}
\end{aligned}$$

Now, we turn to the counterfactual scenario with a counterfactual allowance price and look for the "intermediate" solution that captures within-zone fuel substitution while suppressing between-zone relocation. Because the latter is equivalent to keeping total fossil generation within zone b constant, any choice of percentiles $\mathbf{p^{im}_{b,m}}$, where im denotes the "intermediate" solution, must be such that $\widehat{FG}_{b,m}(\mathbf{p^{im}_{b,m}}) \stackrel{!}{=} FG_{b,m}$

Of course, the choice of $\{p^{im}_{b,f,m}\}$ cannot and should not be random, but rather empirically informed. An informed starting value for $\{p^{im}_{b,f,m}\}$ (and in that sense also for $\{p^{cf}_{b,f,m}\}$, which will eventually establish the counterfactual equilibrium) is $\mathbf{P}^* = \{p^*_{b,f,m}\}$. Since $\mathbf{P}^*$ establishes estimates that constitute a market equilibrium at actual allowance prices, it might not be too far off the percentile values necessary to establish a counterfactual equilibrium. To find the "intermediate" solution based on $\mathbf{P}^*$, we do the following. Right at $\mathbf{P}^*$, we almost surely have $\widehat{FG}_{b,m}(\mathbf{p^*_{b,m}}) \neq FG_{b,m} \;\; \forall\{b,m\}$. To make this equation hold true, we look for a set of scalars $\{a_{b,m}\}$ for which

$$\widehat{FG}_{b,m}(a_{b,m} \cdot \mathbf{p^*_{b,m}}) = \widehat{FG}_{b,m}(\mathbf{p^{im}_{b,m}}) = \widehat{FG}^{im}_{b,m} \stackrel{!}{=} FG_{b,m} \;\; \forall\{b,m\}$$

In words, in a given zone b × month m pair, we scale all percentiles with a common factor $a_{b,m}$ until predicted fossil generation $\widehat{FG}^{im}_{b,m}$ equals observed one $FG_{b,m}$. It is important to note that $\widehat{FG}^{im}_{b,m}$ differs from $\widehat{FG}_{b,m}$ in two ways. First, $\widehat{FG}_{b,m}$ uses model predictions $\widehat{y}_i$ under actual allowance prices, whereas $\widehat{FG}^{im}_{b,m}$ is based on the model predictions $\widehat{y}^{cf}_i$ that use units' marginal cost structure under counterfactual allowance prices. The difference between $\widehat{y}_i$ and $\widehat{y}^{cf}_i$ captures the predicted fuel substitution (at any percentile). Second, $\widehat{FG}_{b,m}$ uses $\mathbf{p^*_{b,m}}$ whereas $\widehat{FG}^{im}_{b,m}$ uses $a_{b,m} \cdot \mathbf{p^*_{b,m}} = \mathbf{p^{im}_{b,m}}$. The factor $a_{b,m}$ corrects $\mathbf{p^*_{b,m}}$ so that aggregate predicted fossil generation $\widehat{FG}^{im}_{b,m}$ equals aggregate actual fossil generation $\widehat{FG}_{b,m}$ (on the zone × month level) as required by our definition of the "intermediate" solution.[15]

Table 6: Average Deviation of Fossil Demand Prediction in Intermediate Solution, €50 ETS price floor

Mapcode	**Avg. Dev. (TWh)**	**Mapcode**	**Avg. Dev. (TWh)**
DE_AT_LU	1.404	DK1	0.282
DE_LU	1.239	BE	0.267
FR	0.711	CZ	0.248
NL	0.537	IT-CNORTH	0.147
SE3	0.475	IT-CSOUTH	0.136
AT	0.421	IT-SOUTH	0.134
SE4	0.383	PL	0.107
ES	0.366	IT-Sardinia	0.035
DK2	0.333	IT-Sicily	0.001
IT-NORTH	0.312		

The set of percentiles $\{p^{im}_{b,f,m}\}$ that establishes the "intermediate" solution incorporate the within-zone fuel substitution. We use $\{p^{im}_{b,f,m}\}$ as the starting point to approximate the true counterfactual equilibrium by now allowing (and focusing on) between-zone relocation of generation. Table 6 shows how far the implied predicted residual demand of the intermediate solution is off from actual residual demand under a counterfactual €50 allowance price floor. The German-Austrian-Luxembourgian bidding zone (until September 2018) and German-Luxembourgian bidding zone (from October 2018) clearly stick out. Mechanically suppressing between-zone relocation of generation throws Germany relatively far off an equilibrium. This makes intuitive sense. Germany relies heavily on coal in its electricity generation, and is at the same time surrounded by gas-heavy neighbors such as the Netherlands, Belgium, Austria and partly Italy while Austria was part of the common bidding zone. Accordingly, in a high allowance price counterfactual, it is likely that Germany

[15] We also scale the percentiles of border points, $p^{im}_{b,BI,m} = a_{b,m} \cdot p^*_{b,BI,m}$ for consistency.

starts importing substantially more electricity generated by gas plants abroad while curtailing its own coal plants. The intermediate solution, however, mechanically blocks this channel and only allows for substitution between German coal and gas plants. In contrast, coal-heavy Poland is close to a true market equilibrium in the intermediate solution already. Since it only borders other coal-heavy zones, Germany and Czechia, there is little room for substantial cross-border relocation of electricity generation, and it is plausible that the bulk of the effect of a high allowance price materializes through fuel substitution within Poland.

To approximate the counterfactual equilibrium, we aim to close the deviations between predicted counterfactual residual demand and actual residual demand that are shown in Table 6. The intermediate solution implies a value for predicted counterfactual residual demand in zone b and month m equal to $\widehat{RD}^{im}_{b,m} = \widehat{RD}_{b,m}(\mathbf{P^{im}_m})$. $\mathbf{P^{im}_m}$ denotes the intermediate solution's percentiles across all bidding zones in month m. The model prediction for zone b's residual demand in any given month m depends on the entire set of percentiles of that month, $\mathbf{P^{im}_m}$, because it also depends on other zones' imports from b. As Table 6 shows, in general

$$\widehat{RD}_{b,m}(\mathbf{P^{im}_m}) \neq RD_{b,m} \quad \forall (b,m)$$

To approximate the counterfactual equilibrium in a given month m, we look for a set of scalars $\{w_{b,m}\}$ that we multiply with $\mathbf{P^{im}_m}$, the percentiles of the intermediate solution, to minimize the gap between predicted counterfactual residual demand and actual residual demand across zones. Algebraically, we aim to do the following to obtain the counterfactual equilibrium

$$\forall m: \ \{w^*_{b,m}\} = \underset{\{w_{b,m}\}}{\operatorname{argmin}} \sum_b \left(\widehat{\mathrm{RD}}_{b,m}(w_{b,m} \cdot \mathbf{p^{im}_{b,m}}) - \mathrm{RD}_{b,m}\right)^2$$

In every month-of-sample m, we look for the set of weights $\{w^*_{b,m}\}$ that minimizes the system-wide (across all zones) squared differences in predicted counterfactual and actual residual demand.

In practice, we cannot solve the problem as stated above because $\widehat{RD}_{b,m}(w \cdot \mathbf{p^{im}_{b,m}})$ is not a well-defined function of w. We would have to apply a numerical search for a global minimum across a 19-dimensional vector $\mathbf{w_m} = (w_{b_1,m}, ..., w_{b_{19},m})$ since we have 19 bidding zones, and do this 72 times for the 72 months that our sample spans. This is computationally impossible. Instead, we approximate the solution as follows, separately for every month m. We take $w_{b,m} = 1$ for every zone b as our starting value, given that $\mathbf{P^{im}_m}$ is the percentile matrix of our intermediate solution and therefore our best guess for the counterfactual equilibrium. Then, we alter only one value of $\mathbf{w_m}$, say $w_{b_1,m}$, and look for the value of $w_{b_1,m}$ that minimizes $\sum_b \left(\widehat{\mathrm{RD}}_{b,m}(w_{b,m} \cdot \mathbf{p^{im}_{b,m}}) - \mathrm{RD}_{b,m}\right)^2$, holding constant $w_{b_2,m}, ..., w_{b_{19},m}$. We continue doing

this for every value of $\mathbf{w_m}$, one at a time, and complete the whole cycle three times.

4.5 Identifying Assumptions

Our empirical approach and model predictions rely on several identifying assumptions which we summarize here. A more detailed discussion of potential threats to identification and how our model addresses them is relegated to Appendix C.2.

First, we take the fleet composition and the entry and exit of power units as exogenous due to their long-term nature. This avoids confounding short-run dispatch responses to carbon pricing with the longer-term effects of plant closures and openings. To the extent that increasing carbon prices may drive coal-fired power plants to close, we underestimate the effect of carbon pricing on co-pollution emissions.

Second, there are no time-varying confounding regulatory policies on the state level that would fundamentally change the hour-by-hour operating decisions of the existing power plant fleet. We argue that this is plausible during our analysis period. Structural policies such as national coal phase-out trajectories affect fleet composition through plant entry and exit, which we treat as exogenous as described above. Therefore, plant entry and exit only influence our predictions through the merit order, which the model handles directly. In other words, we rule out policies that would alter the hour-by-hour relationship between carbon prices and dispatch decisions for units already in the fleet.

Third, we follow previous work such as Grossi et al. (2017), Abrell et al. (2022) and Jarvis et al. (2022) and take fuel prices as exogenous because these are largely determined by fixed supply frameworks and global benchmarks. For instance, European natural gas prices largely follow long-term, oil-indexed contracts.

Fourth, in line with the previous literature we assume that wholesale markets are competitive and hence power plant operators had no market power. Specifically, we estimate supply curves based on marginal costs without mark-up. To support this assumption, we show in Appendix C.2 that (i) our predictions are not systematically biased at certain carbon price levels; (ii) on the unit-level, the model prediction error is largely uncorrelated with the allowance price; and (iii) the model's MSE and R^2 are stable year-on-year across the sample, spanning periods of low, medium, and high allowance prices.

Fifth, we abstract away from short-term re-dispatch and connect bidding in the day-ahead market to actual realized dispatch and generation. We also do not model market frictions such as within-zone transmission constraints. In Appendix C.2, we explain how our Random Forest, by design and thanks to including adequate additional unit-specific covariates, can account for and plausibly correct for these disturbances.

5 External Costs

The objective of this study is to quantify externalities that arise from fossil-fuel fired power generation in our sample. Those include global externalities from carbon emissions as well as mortality impacts of air pollution that can be attributed to the generating units in our sample. Are study focuses on the latter, more local negative externality.

We estimate external costs in several steps. The starting point is the predicted hourly electricity output of generating units for each of the scenarios described in the previous section. We compute CO_2 emissions based on the fuel input required to match predicted output. To estimate the associated co-emissions, we first construct measures of the amount of each co-pollutant emitted per ton of CO_2. These 'co-pollutant intensities' allow us to predict emissions of $PM_{2.5}$ precursors based on a unit's predicted CO_2 emissions. Ultimately, public health impacts of co-emissions depend on atmospheric dispersion, the chemistry of secondary air pollutant formation, and population exposure. To account for these processes, we use a custom-built chemical transport model (CTM) that accounts for these complex interactions and maps unit-level air pollutant emissions to premature mortality associated with $PM_{2.5}$ exposure which is predicted by the CTM.

Each step of this process is described in detail in the following section.

5.1 CO_2 Emissions

We use the net electric efficiency of the generating units in our sample to back out the energy input required for a unit's electricity output over a given time interval. We obtain the corresponding CO_2 emissions by multiplying this input requirement with the CO_2 emission factor for the unit's main fuel (lignite, coal, or natural gas). This exploits the direct proportionality between the amount of fuel burned and the amount of CO_2 emitted.

5.2 Air pollution emissions

We further use proportionality in the estimation of unit-level emissions of the main precursors to fine dust ($PM_{2.5}$) and ozone (O_3). Estimation is required because these emissions are not continuously measured in the European power sector.[16] However, self-reported data on annual pollution emissions are available from two official sources.

First, as mentioned in Section 2 above, generating units with a total rated thermal input of at least 50 MW are legally obliged to report their emissions of SO_2, NO_x, and particulate matter (PM_{10}) under the LCP and Industrial Emissions Directives. LCP reporting units

[16] In contrast, U.S. power plants are subject to a Continuous Emissions Monitoring System (CEMS).

(so-called installation parts) can typically be linked to one or more generating units in ENTSO-E. Second, data on annual releases of a broader set of air pollutants are available from the EPRTR. [17] EPRTR reporting units, so-called facilities, are more aggregate and may encompass anywhere from one to roughly a dozen installation parts. Only some of these facilities can be successfully matched to ENTSOE, generally at the power plant level, which requires assumptions to allocate facility-level pollution intensities from their facility across installation parts.

Both data sources can be used to compute annual co-pollutant intensities for each of the generating units in our sample. In 335 cases, we establish a one-to-one match to LCP installation parts. We compute the NO_x intensity by dividing reported emissions of NO_x by predicted CO_2 emissions in the baseline (likewise for other pollutants). In 191 cases, multiple generating units are linked to the same LCP installation. In those cases, we compute pollution intensities for aggregate generation and assign this measure to each generating unit. We apply the same procedure after matching one or several generating units to EPRTR facilities. Doing so allows us to compute co-pollutant intensities of generating units not contained in the LCP data as well as air pollutants not reported in LCP data, notably ammonia (NH_3) and black carbon.

Figure 9 displays the distribution of co-pollutant intensities among generating units in our sample, by fuel type. These are computed using facility-level data from EPRTR and averaged across all available years.[18]

5.3 Mapping pollution emissions to mortality impacts

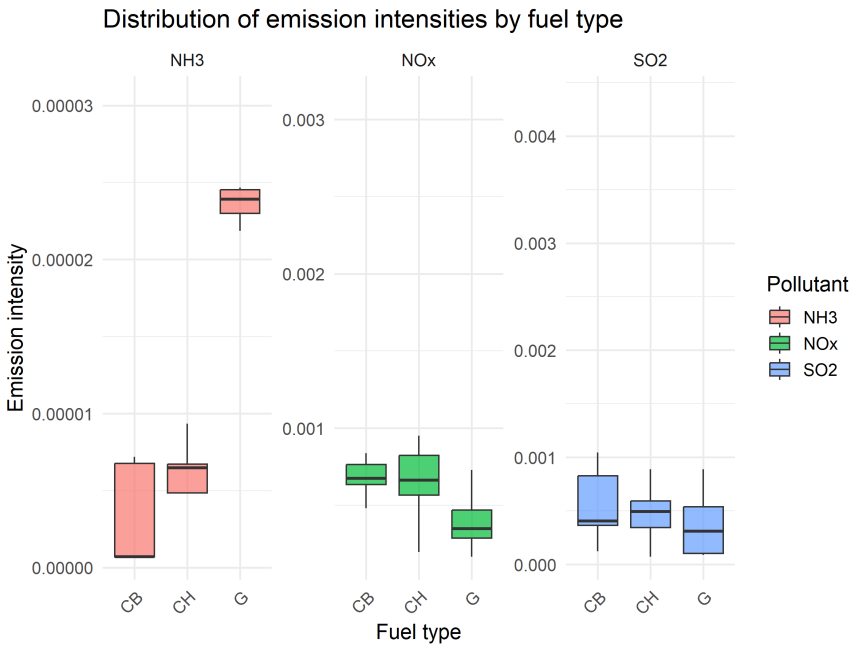
A well-known fact about air pollution is that the place of emission is not necessarily equal to the place of impact. Once pollutants are released from the smokestack, weather, topography, and pre-existing air pollution suspended in the local atmosphere influence their atmospheric dispersion and the formation of secondary pollutants such as ozone and $PM_{2.5}$. Therefore, mapping emissions at generating units in our sample to ambient pollution concentrations and population exposure requires us to use a chemical transport model (CTM), i.e. a computational tool capable of tracking the highly non-linear atmospheric transport and chemical reactions of air pollutants from the place of release to their final deposition.

We employ the nested GEOS-Chem Adjoint model developed by Gu et al. (2023). Based on primary pollution emissions of, among others, NO_x, SO_2, and NH_3, this CTM computes population exposure to both primary $PM_{2.5}$ (organic carbon, black carbon, and fine dust)

[17] We use version `eea_t_ied-eprtr_p_2007-2023_v14_r00`

[18] All calculations below are based on the facility-level data shown in Figure 9. We currently augment these with information from LCP data to improve coverage.

Figure 9: Co-Pollution Intensities

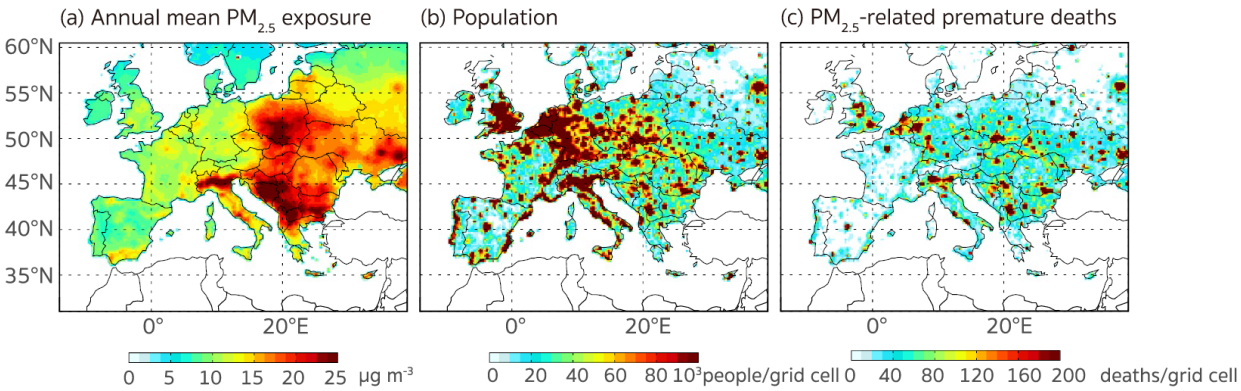


and secondary $PM_{2.5}$ (e.g., sulfate, nitrate, ammonium) on a $0.01° \times 0.01°$ nested grid for the European domain. The CTM is calibrated to 2015 weather and emissions data, which aligns well with the beginning of our sample period.[19] Health impacts in a given grid cell are computed as the number of $PM_{2.5}$-related premature deaths. Following the methodology proposed in the Global Burden of Disease Study (GBD) 2019 (Murray and *et al.*, 2020), this number is obtained by adjusting the age-stratified mortality within a cell for the $PM_{2.5}$ induced mortality among the "big six" causes of death (chronic obstructive pulmonary disease, ischemic heart disease, lung cancer, type-2 diabetes, and stroke). Figure 10 displays heat maps of annual mean concentrations, population density, and $PM_{2.5}$ related premature deaths in Europe at the 2015 baseline. We provide more detail on this calculation in Appendix A.

Importantly, the CTM's adjoint feature allows us to quantify the change in aggregate $PM_{2.5}$-related mortality induced by a power plant's change in precursor emissions. In an "adjoint run", the CTM computes partial derivatives of the total mortality count at each step of the computation. In a reverse application of the chain rule, these can be aggregated into marginal effects for each pollutant and grid cell, referred to as adjoint sensitivities. We assign the vector of sensitivities in a grid cell to all units located in the cell. We compute the unit's mortality impact by multiplying co-pollutant emissions with the appropriate sensitivity. This

[19]Since CTMs are computationally expensive, they are calibrated to a particular base year in which weather conditions are typical.

Figure 10: $PM_{2.5}$ Exposure, Population, And Health Burden In Europe (2015)



Source: Gu et al. (2023a)

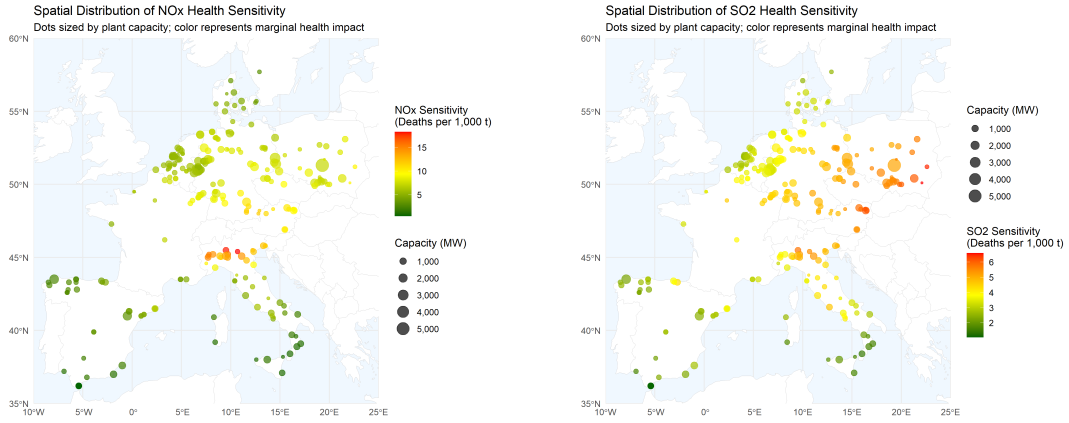
is useful for both source attribution (how many PM-related deaths can be attributed to a given power plant?) and counterfactual analysis (how does PM-mortality change with the carbon price?), and requires no additional computational costs.

Marginal impacts Figure 11 shows sensitivities for the power plant locations in our analysis sample. Depending on the place of release, an additional 1,000 metric tons of NO_x is associated with between 0.4 and 10.3 additional premature deaths in Europe (mean 2.9); the range for SO_2 is 0.0 to 1.7 per 1,000 metric tons (mean 0.4); the range for ammonia (not displayed on the map) is 3.9 to 189 premature deaths 1,000 metric tons emitted (69 on average across locations).

The sensitivities displayed in the maps depend on a multitude of factors, some of which are shared and others differ between pollutants. Common factors affecting how far a pollutant gets transported are weather and topography, but lifetime's differ across pollutants. Since NO_x is a relatively short-lived pollutant, the spatial correlation of sensitivities across regions is low compared to SO_2 which remains in the atmosphere longer and can thus travel farther. Since $PM_{2.5}$ forms via chemical reactions during atmospheric transport, the amount and composition of pre-existing air pollution also affects pollution levels. Last not least, differences in population density explain geographic variation of morbidity damages per ton emitted.

For illustration, consider the cluster of Northern Italian power plants that feature the highest NO_x-related sensitivities in Figure 11a. These plants are located in a very densely populated area with infamously little air exchange, both of which maximizes damage from short-lived NO_x. In contrast, the NO_x health sensitivity of Dutch plants is much lower despite being located in densely populated areas as well. This is because average wind speeds in the

Figure 11: Sensitivity Of $PM_{2.5}$-Related Premature Deaths To Precursor Emissions



(a) Sensitivity to NO_x emissions

(b) Sensitivity to SO_2 emissions

Notes: The maps displays how lethal emissions of NO_x and SO_2 are at a given power plant location are (in premature deaths per 1,000 tons emitted). Values are derived from adjoint sensitivities to $PM_{2.5}$ precursor emissions, computed with the GEOS-Chem model developed by Gu et al. (2023). Simply put, these measure For visual clarity, health sensitivities are aggregated to the plant level by taking the arithmetic mean of the unit-specific sensitivities, to avoid the overlap of individual unit markers at the same geographic location. Sensitivities are illustrated as additional premature deaths in Europe per 1,000 metric tonnes of the pollutant emitted at the respective plant location.

Netherlands are much higher than in Northern Italy, allowing for much quicker dispersion of NO_x.

Unlike NOx, SO_2 tends to stay in the atmosphere for several days, allowing it to be transported across state or national borders. Accordingly, the map in Figure 11b looks quite differently. Central Europe is characterized by a predominant southwesterly airflow over the course of the year so that the negative health impacts of a unit of emitted SO_2 tend to materialize in a corridor to the northeast of its source. This is a contributing factor to why power plants in Eastern and Southern Germany, Czechia and Austria feature high SO_2 sensitivities. Their emissions tend to travel northeastward across densely populated areas, and in particular into densely populated Southwestern Poland. This effect is compounded by notoriously bad air quality in this area which accelerates $PM_{2.5}$ formation. In contrast, power plants on the Spanish coast or in Southern Italy feature low sensitivities because the wind tend to transport their SO_2 emissions offshore.

5.4 Abatement cost

We compute the total abatement cost associated with a counterfactual scenario as the difference between generation costs of counterfactual output and actual electricity output across units and over time. Division by the corresponding difference in CO_2 emissions

provides a measure of the cost per ton of CO_2 avoided, or average abatement cost per ton,

$$AAC(s) = \frac{TC(s) - TC(0)}{E(s) - E(0)}. \tag{1}$$

To estimate the terms in this expression, we draw on the same information used to construct the empirical and counterfactual merit curves above. We compute aggregate CO_2 emissions as described in Section 6.2 above. To compute total generation cost in a given scenario s, we first compute unit-level generation cost as the product of marginal cost and output, and aggregate this across all fossil-fuel units and all hours-of-sample:

$$TC(s) = \sum_h \sum_i MC_{i,h} \cdot \hat{y}_{i,h}(s)\kappa_i. \tag{2}$$

Unit i's marginal cost MC in hour t to generate one MWh of electricity is determined by the sum of variable operating cost VOC and fuel cost required to generate one MWh:

$$MC_{i,t} = VOC_i + \frac{P_{i,t}^{\text{fuel}} + \tau_{i,t}}{\eta_i}. \tag{3}$$

We obtain fuel input cost by dividing the price ($P_{i,t}^{fuel}$) per MWh of the primary fuel –plus transportation costs ($\tau_{i,t}$) in the case of hard coal– by the unit's net electric efficiency (η_i).

Two remarks are in order. First, we do not include a fixed cost or capital cost component in our abatement cost calculation. In the day-ahead market, plant operators bid based on their avoidable costs (variable operating costs, fuel, and carbon) because their capital costs are already sunk. Importantly, these capital costs remain sunk costs in the counterfactual scenarios as well because we abstract from decommissioning. Even if a plant does not recoup its capital costs in a specific hour, it is still better off running as long as the market price exceeds its short-run marginal costs. Second, we do not include the price of EUA certificates because the social abatement costis given by incremental operating costs incurred by electricity firms due to carbon pricing. This operating cost is the actual change in fuel and variable operating expenditures incurred by the electricity system to achieve a reduction in CO_2. Abatement costs arise from distorting the order of power units in the supply curve, which, absent carbon pricing, is such that operating cost are minimal. The policy-induced distortion due to carbon pricing raises total operating costs in the market because some units with high operating costs yet low carbon intensity are pulled forward in the supply curve, whereas other units with low resource costs yet high carbon intensity are pushed back. In other words, the market bypasses technically "cheaper" generation in favor of generation with higher fuel costs but lower carbon intensity. The resulting abatement cost is therefore

the delta between the technically optimal (least-fuel-cost) dispatch and the policy-induced (low-carbon) dispatch.

6 Results

This section compares model predictions of electricity market outcomes under different assumptions about the carbon price. Our baseline scenario assumes a zero carbon price. The ETS scenario assumes that EUA prices follow their actually observed path. The price floor scenario assumes that EUA prices follow their actual path but never drop below €50.

6.1 Electricity generation

Table 7 summarizes aggregate output by source for different scenarios. In the six-year period from 2016 to 2021 that we observe, the fossil power units in our sample generate a total of 3,137.2 TWh of electricity. The top row shows our model predictions in the absence of carbon pricing. Hard coal has the largest share in electricity generation under a zero carbon price at 38.4%, followed by natural gas at 34.6%, and lignite at 27%. Border flows have a volume that is about a third of total fossil generation, highlighting how fossil power units frequently generate electricity for exports.

To assess how pricing carbon affects fuel switching and border flows, we compare this scenario to the observed generation under actual allowance prices (middle row) as well as to predicted generation under a counterfactual €50 allowance price floor (bottom row). Relative to the baseline of no carbon price, the EU ETS causes a substantial shift away from hard coal, which drops by almost 18%, and towards natural gas, which rises by more than 21%. In contrast, lignite-fired generation barely changes. This highlights that EUA prices fluctuating between €5 and €30 for most of the sample period were sufficiently high to shift some efficient natural gas units before hard coal units in the merit order, but not in front of the cheaper lignite units. Only as we further increase the allowance price to at least €50, as in the price floor scenario, do we observe a substantial reduction in lignite generation of 14% relative to the zero-carbon price scenario. The reduction in hard coal generation further increases to 32%, while the increase in natural gas generation further grows to more than 46%.

Turning to zone switching, the volume of cross-border flows also grows with the allowance price. Relative to the baseline, electricity imports are more than 10% higher under observed ETS prices, and would be more than 18% higher with a €50 allowance price floor. The positive correlation between the carbon price and the volume of cross-border flows is presumably driven by the asymmetric impact of carbon pricing on regional power-plant

Table 7: Generation [TWh] by type, aggregated 2016-2021

Type	Lignite	Hard Coal	Natural Gas	Total Fossil	Imports
Zero EUA price (€0)	845.7	1,204.9	1,086.6	3,137.2	1,055.6
Actual EUA price	829.8	988.3	1,319.2	3,137.2	1,163.8
% Change vs €0	-1.88%	-17.98%	+21.41%	0.00%	+10.25%
EUA price floor at €50	727.3	819.4	1,590.5	3,137.2	1,250.8
% Change vs €0	-14.00%	-32.00%	+46.37%	0.00%	+18.49%

Notes: Aggregate generation in terawatt-hours (TWh) by type of plant of all units in the analysis sample from 2016 to 2021. We compare three allowance prices. Row (i) contains predicted generation in a €0 allowance price counterfactual scenario. Row (ii) contains observed generation under actual allowance prices. Row (iii) contains predicted generation in a €50 allowance price floor counterfactual scenario.

fleets. By disproportionately increasing the marginal costs of carbon-intensive generation, the allowance price widens the cost heterogeneity between bidding zones, thereby expanding the economic scope for trade and incentivizing imports from regions with more efficient marginal units.

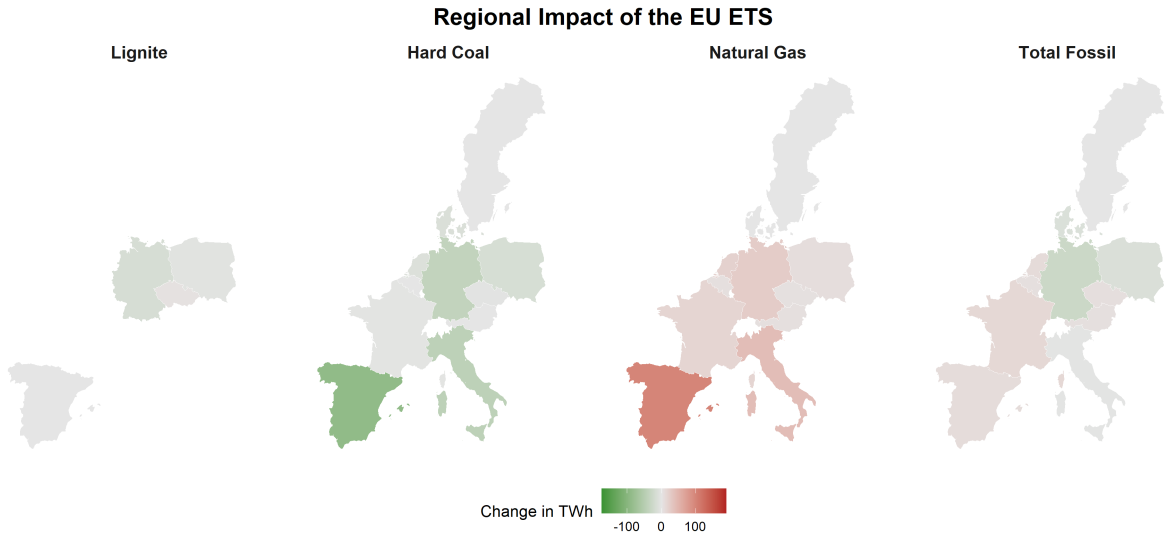
Figure 12 visualizes these policy-induced geographic shifts in electricity generation across countries in our sample on a map. Figure 12a shows the change in generation under actual ETS allowance prices relative to the €0 baseline. Notably, the largest fuel substitution occurs within Spain. Because Spain has substantial gas-fired capacity installed in excess of what is usually needed, changes in the carbon price induce very large shifts in which units generate. The share of Spanish coal units in total fossil generation falls from almost 70% (217 TWh) at baseline to just 39% (126 TWh) under the ETS.[20] At the same time, natural gas generation doubles from 99.6 TWh (31%) to 200 TWh (61%). The EUA price floor further increases the gas share by another 90 MWh to more than 87% of all fossil fuel generation, with coal generation droping by another 84 TWh to 12.6%.

In Italy, there is an almost one-to-one substitution away from coal towards gas. Relative to a zero carbon price, coal generation drops by 43 TWh while gas generation increases by 41 TWh, which is about 7% of total fossil generation, respectively. When further increasing the carbon price to €50, another 35.3 TWh switch from coal to gas, which is about 6.5% of the volume of total fossil generation.

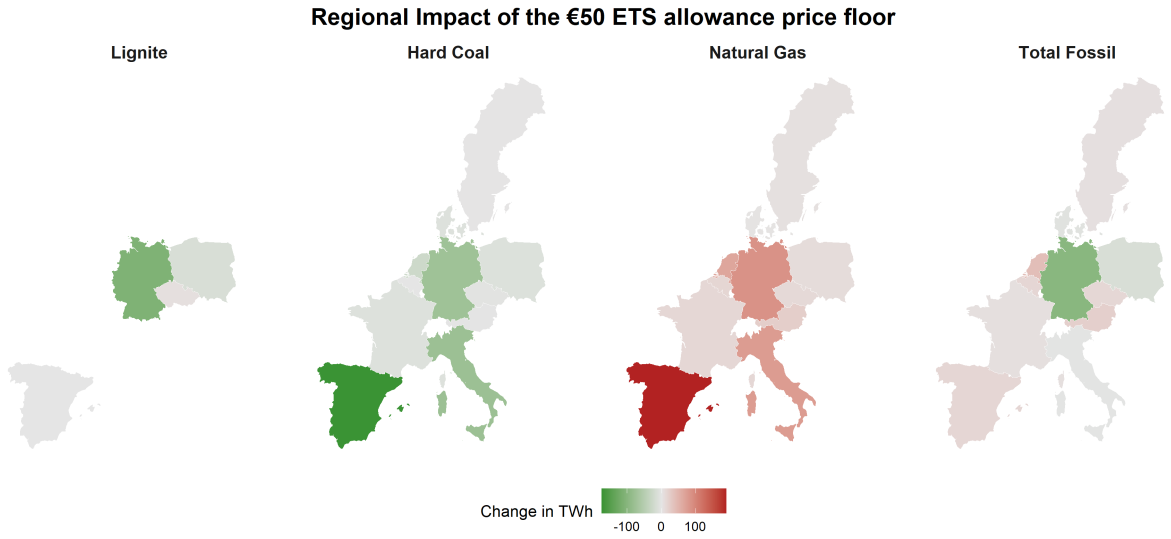
Germany notably increases electricity imports. Under the EU ETS, lignite and coal generation drops by 54 TWh, which is made up by 26 TWh higher gas generation and 31 TWh higher electricity imports. However, these effects are small relative to the about 1,000 TWh of fossil generation in Germany. The magnitudes become notably larger under a €50

[20]Refer to Appendix Table B.2 for the country×fuel generation quantities.

Figure 12: Generation changes in response to carbon pricing



(a) Total Change in Generation, Actual EU ETS allowance prices relative to €0. Table **??** contains the corresponding values.



(b) Total Change in Generation, €50 EU ETS allowance prices relative to €0. Table **??** contains the corresponding values.

price floor. Relative to the baseline, lignite and hard coal generation decrease by a total of 185 TWh, which is balanced by an additional 86 TWh of natural gas generation, 71 TWh of additional imports, and a reduction of electricity exports to its neighbors by 19 TWh. In general, one can see how generation is relocated away from coal-heavy Germany (-99 TWh) and towards its gas-heavy neighbors such as the Netherlands (+40 TWh), Austria (+22 TWh), and Belgium (+16 TWh).

6.2 CO_2 emissions

Table 8 summarizes the CO_2 emissions from the fossil fuel units in our sample. Absent carbon pricing, the power units in our six-year sample emit a total of 2,529.9 metric megatons of CO_2. Hard coal units contribute 45.3% to this quantity, lignite units 37.5%, and natural gas units 17.2%. Comparison to emissions under the actual EUA prices suggests that the EU ETS has achieved a reduction in CO_2 emissions of 137.8 Mt, or 5.45% relative to the zero carbon price scenario. This emission reduction is due almost exclusively to substitution away from hard coal (-18.32%) and towards natural gas (+20.53%). CO_2 emissions from lignite move very little at a reduction of only -1.81%. As mentioned in the previous section, the EU ETS mostly causes a relocation of lignite generation from Germany to Czechia and the according change in cross-border flows, with hardly any net abatement.

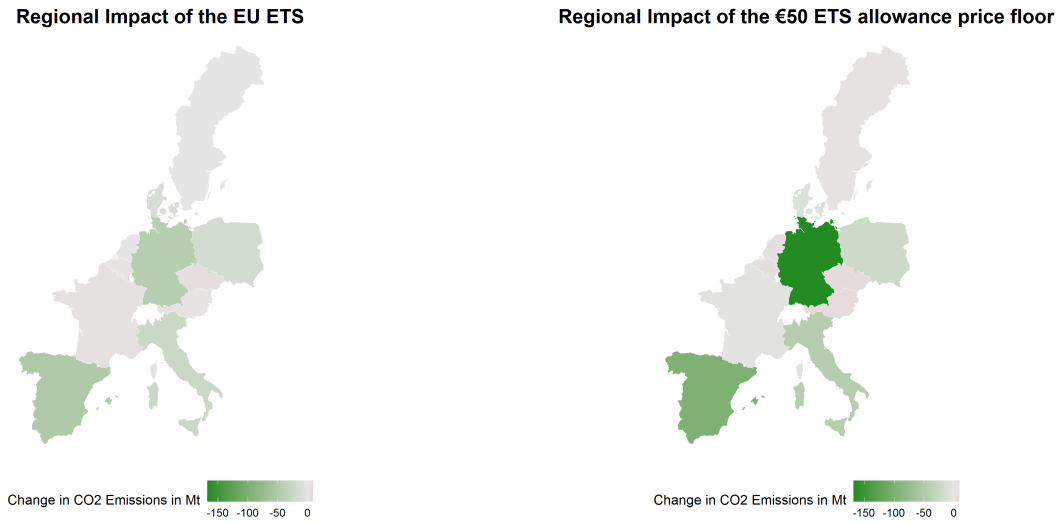
With a €50 EUA price floor, fuel switching from hard coal to natural gas drives further abatement, reducing total emissions by one eighth relative to a world with a zero carbon price. In contrast to actual ETS prices, the €50 price floor achieves a substantial reduction of emissions from lignite units at -14.58%. Lignite generation still moves from Germany to Czechia, yet far away from the nearly one-to-one relocation that occurs under actual prices.

Figure 13 shows which countries abate CO_2. Panel 13a displays a heat map of abatement relative to a zero carbon price baseline. The largest absolute abatement occurs in Spain with a reduction in emissions from 250 to 199 Mt (-20.5%).[21] Germany achieves the second-largest absolute abatement at 44 Mt, which only comes in at a 4.4% reduction relative to the 1,003 Mt of CO_2 emitted absent a carbon price. Italy, which also features substantial scope for domestic fuel substitution, abates 25 Mt of CO_2 (-8.9%). In contrast, the coal-intensive Polish power sector achieves relatively little abatement due to its small scope for domestic fuel substitution and its connection to the similarly coal-heavy countries Germany and Czechia. Polish emissions decrease by only -3.1% (18 Mt). Denmark achieves the largest relative abatement, reducing emissions by 36% (from 36 to 23 Mt).

CO_2 emissions increase in France (6.0%), Belgium (7.2%) and Austria (12.9%) because

[21] Refer to Table B.5 for the specific numbers.

Figure 13: Country-specific CO_2 abatement with carbon pricing, 2016-2021



(a) Actual EUA prices.

(b) Actual EUA prices with price floor

Notes: The maps show how carbon pricing scenarios reduces CO_2 emissions relative to the baseline of zero carbon prices. Panel (a) refers to actual EUA prices, panel (b) to actual EUA prices plus a price floor at €50. The corresponding values are reported in Appendix Table B.5.

their gas units become more active as they replace some coal-fired generation in neighboring zones. Czech emissions also increase by 4.8%, largely because their lignite units replace some of those in nearby Eastern Germany. Lastly, emissions in the Netherlands stay roughly constant due to two effects that cancel each other out: Domestic carbon abatement results from substitution away from coal to gas. Yet, at the same time, carbon emissions increase because of higher gas-fired generation, presumably to replace some German coal-fired generation. These opposing effects add up to a zero net change in emissions.

Figure 13b shows carbon abatement when imposing a €50 price floor on top of actual EUA prices. The absolute quantity of abated CO_2 increase from 138 Mt to 178 Mt; in relative terms, emissions are reduced by another 7.4%. As before, Spain achieves large emissions abatement due to its vast capacity in natural gas units that allow for large-scale fuel substitution. In total, Spain abates 47 Mt of CO_2, a reduction of 23.7%. Abatement in Italy also continues at a similar level as in Figure 13a, with 20 Mt of CO_2 abated, a reduction of 7.7%.

By far the largest abatement in absolute numbers occurs in Germany. Starting at 959 Mt of CO_2 under actual allowance prices, introducing a €50 price floor reduces German CO_2 emissions by 124 Mt (-12.9%). This huge reduction in CO_2 emissions results largely from the fact that German lignite plants are predicted to be significantly less active in the €50 price floor scenario, leading to large reductions in CO_2 emissions. Lastly, Poland, which abates

Table 8: CO_2 Emissions [Mt] by type, aggregated 2016-2021

Type	Lignite	Hard Coal	Natural Gas	Total
Zero EUA price (€0)	948.9	1,146.2	434.9	2,529.9
Actual EUA price	931.7	936.3	524.1	2,392.1
% Change vs €0	-1.81%	-18.32%	+20.53%	-5.45%
EUA price floor at €50	810.5	772.8	630.6	2,213.9
% Change vs €0	-14.58%	-32.57%	+45.00%	-12.49%

Notes: Aggregate CO_2 in metric megatons (Mt), 1 Mt = 1,000,000 tons, by type of plant of all units in the analysis sample from 2016 to 2021. We compare three allowance prices. Row (i) contains predicted generation in a €0 allowance price counterfactual scenario. Row (ii) contains observed generation under actual allowance prices. Row (iii) contains predicted generation in a €50 allowance price floor counterfactual scenario.

18 Mt (-3.1%) due to the EU ETS relative to a zero carbon price, only achieves a further reduction of emissions of 6 Mt (-1.0%) under the €50 price floor.

Juxtaposing Germany and Poland is particularly interesting given that both countries are coal-heavy, however, Germany relies heavily on lignite whereas Poland relies heavily on hard coal. At the same time, Poland achieves most of its abatement under actual ETS prices (-18 Mt/-3.1%). Further increasing the allowance price from its observed level to the €50 price floor achieves only little additional abatement (-6 Mt/-1.0%). In contrast, Germany achieves relatively little abatement under observed allowance prices (-44 Mt/-4.4%), but substantial abatement as allowance prices rise further to at least €50 (-124 Mt/-12.9%).

Overall, these are plausible results given the marginal cost structure of fossil power plants. At the fuel prices that prevailed in the late 2010s, and when carbon emissions go untaxed, lignite units generally have the lowest marginal costs, while natural gas units have the highest marginal costs, with hard coal units standing in the middle. Accordingly, as the carbon price start to rise, we first expect fuel substitution to occur between hard coal and natural gas, and only later between lignite and natural gas.[22] This rationalizes the finding that most of carbon abatement in hard-coal-heavy Poland already occurs as the carbon price rises from zero to its moderate real-world levels, whereas carbon abatement in lignite-heavy Germany only takes off as carbon prices rise further towards €50 per ton.

[22]Fuel substitution between lignite and hard coal also occurs, although at substantially smaller quantities. This is due to the fact that natural gas units are much less CO_2 intensive than both lignite and hard coal units, whereas hard coal only has a relatively small advantage over lignite in terms of CO_2 intensiveness.

Table 9: Predicted premature deaths in 2016–2021: Comparison to Zero Price Baseline

Type	Lignite	Hard Coal	Natural Gas	Others	Total
Zero EUA price	6,544	4,902	872	278	12,596
Observed (Actual EUA price)	6,382	4,006	1,036	226	11,650
Change vs. Zero Price [%]	−2.5	−18.3	+18.8	−18.7	−7.5
EUA price floor €50	5,543	3,357	1,266	193	10,359
Change vs. Zero Price [%]	−15.3	−31.5	+45.2	−30.6	−17.8

Notes: Predicted premature deaths from NOx, SOx, and NH_3 emissions (2016–2021). Percentages are calculated relative to the counterfactual Zero EUA price scenario.

6.3 Mortality

We report $PM_{2.5}$-related mortality for different scenarios in Table 9. At the given ETS allowance prices and emission levels, we estimate a total of about 11,650 premature deaths in our six year observation window due to air pollutant emissions from fossil fuel power units. In the absence of carbon pricing, almost 1,000 additional premature deaths would have occurred, an 8.1% increase relative to observed premature deaths. If the allowance price had continuously been at a minimum of €50 instead of its actual level, about 1,300 premature deaths, one in every nine deaths, would have been avoided.

These figures also highlight that mortality due to air pollutant emissions responds more strongly to carbon pricing than CO_2 emissions alone. Relative to actual ETS prices, CO_2 emissions would have been 5.8% higher under a €0 ETS price and 7.5% lower under a €50 price floor. At the same time, the changes in mortality are larger at an 8.1% increase and 11.1% decrease, respectively. This highlights the potential of emissions trading in the electricity sector to not only abate CO_2, but also significantly reduce air pollutant emissions.

6.4 Abatement Costs and Health Co-Benefits

With these estimates in hand, we analyze the efficiency implications of two policy experiments: (i) increasing the carbon price from zero to observed EUA prices and (ii) imposing a price floor to make sure EUA prices never fall below €50. Table 10 summarizes all outcomes needed for this.

Abatement Costs Abatement in the power sector is realized by fuel switching—within and between bidding zones—from lignite and coal units towards natural gas units, which tend to have higher marginal costs net of allowance costs. Thus, total generation costs during our six year observation period increase by €588 million with the actual EUA prices

Table 10: Abatement Costs and Health Co-benefits of Two EUA Price Scenarios

	(1) CO_2	(2) Total Costs	(3) Health Benefits	(4) Net Benefits	(5) Deaths
A. Aggregate Change	[Mt]	[m€]	[m€]	[m€]	
Actual EUA Price	-138	588	2,838	2,250	-946
EUA price floor € 50	-316	2,673	6,711	4,038	-2,237
Δ Increment	-178	2,086	3,873	1,787	-1,291
B. Per tonne CO_2 abated		[€]	[€]	[€]	
Actual EUA Price		4.3	20.6	16.3	
EUA price floor € 50		8.5	21.2	12.8	
Δ Increment		11.7	21.8	10.0	

Notes: The table reports the change in aggregate outcomes for the years 2016-2021 when going from zero carbon prices to actual EUA prices (row 1) plus an EUA price floor (row 2), as well as the increment between the two (row 3). Negative CO_2 emissions changes (in millions of tonnes) are interpreted as abatement, and total generation cost changes correspond to abatement costs (in millions of euros). Mortality reductions are expressed as negative changes to premature deaths due to PM2.5 exposure. To compute health benefits of mortality reductions, we multiply avoided deaths with a conservative value of statistical life (VSL) estimate of €3m.

and by €2.67 billion when the price floor is added. The associated abatement over six years amounts to 138 and 316 $MtCO_2$, respectively. While abatement costs per tonne average at €4.26 under observed EUA prices, imposing the €50 price floor leads to incremental abatement at a cost almost three times higher, averaging at €11.71 per tonne of CO_2. This shows that abatement becomes more expensive on a per-tonne basis. Moreover, this result is consistent with economic theory in that units that can contribute to abatement at lowest cost will do so first in a market-based mechanism such as the EU ETS where allowance trading is permitted. Raising EUA prices to €50 and beyond justifies more costly efforts on the part of firms to reduce their emissions.

To make such incremental changes explicit, we also reports incremental changes in outcomes that result when the EUA prices are combined with a price floor. Note that the price floor is only binding during the last weeks of 2021. Other than that, the EUA price was fluctuating between €5-15 in the first years and later between €20-30.

Health Co-Benefits Our model provides us with estimated mortality reductions expressed in terms of avoided premature deaths due to PM2.5 exposure. To compute health

benefits of mortality reductions, we multiply avoided deaths with a conservative value of statistical life (VSL) estimate of €3m.[23]

We estimate health co-benefits of carbon pricing to be as high as €6.7bn with a price floor and €2.8bn at observed EUA prices. Importantly, these are gross co-benefits that do not necessarily imply a welfare improvement. To see this, consider the hypothetical case that all external effects of fine dust pollution from power plants are internalized by Pigouvian regulations that target individual plants and pollutant species. This is highly unlikely to hold in practice, but it helps to build intuition for the insight that mortality reductions we measure need not enhance efficiency but could also reduce it if $PM_{2.5}$ is reduced to levels lower than the social optimum.

Since we have a measure of total generation costs, we are in a position to evaluate whether carbon pricing yields positive net co-benefits. We do so by subtracting abatement costs from gross co-benefits and find positive net co-benefits of €3.8 and €4.0 billion in the two scenarios. Net benefits might be higher because the VSL estimate we use is at the lower end; yet they might also be lower due to omitted elements of marginal costs that are related to air pollution control or due to omitted components of variable operating costs that do not cancel out when we take the difference between generation costs with and without carbon prices. To err on the side of caution, we maintain the low VSL estimate which allows us to conclude by a comfortable margin that the ETS yielded positive net co-benefits.

In panel B of Table 10, we examine costs and benefits per ton of carbon abated. As noted above, specific abatement costs increase with the carbon price. The specific co-benefit averages at €20.6 at actual EUA prices and increases only slightly to €21.8 when the price floor is introduced. We would expect the co-pollution benefit to rise with the carbon price if the units that can abate at lowest cost tended to be those with lower co-pollution damages. However, since co-pollution benefit are almost constant, the correlation between marginal abatement cost of carbon and co-benefits is not very high. As a result, net co-benefits per tonne abated fall from €16.3 under actual carbon prices to €10 for incremental abatement efforts undertaken only when the carbon price is €50 or higher.

Taken together, these results suggest that carbon pricing in the European power sector yielded substantial health co-benefits between 2016 and 2021. Those benefits can be attributed to short-run responses to carbon pricing such as fuel switching within and across bidding zones. While these short-run changes improved total welfare, our results indicate that marginal net co-benefits are diminishing with higher carbon prices. This is because

[23]The United States Environmental Protection Agency uses a VSL of $10.7 million in 2022 dollars (EPA, 2022), which is about €9 million. The European Commission's Directorate-General for Mobility and Transport (EU, 2024) applies a more conservative value of €4.5 million.

increases in marginal abatement costs in response to higher prices are not matched by similar increases in marginal health co-benefits.

Discussion Our results draw attention to a fundamental difference between carbon emissions and air pollutant emissions. The former increase linearly with fuel input. Twice the amount of coal or gas burned will double CO_2 emissions, and plant operators basically have no scope of action to affect this relationship.[24] Conversely, air pollutant emissions depend not only on the fuel burned, but also significantly on plant-specific emissions control technology, such as filters or exhaust fume recirculation. While air pollutant emissions are technically variable, the marginal costs to maintain emissions control technology, primarily reagents and auxiliary power, typically remain below €1 / MWh. This is negligible compared to the transformative impact of the EU ETS (up to €50 / MWh). Furthermore, as these technologies are 'end-of-pipe', they do not influence the carbon-driven merit order shift. This is because the costs of emission control technology are sunk and rather than a variable decision lever. This explains our finding that co-pollution benefits do not significantly vary over different magnitudes of carbon prices, in panel B of Table 10.

As was mentioned above, regulatory overlap of the EU ETS with air pollution policies set by local regulators at plant locations might interfere with our benefit calculation. If such local policies were chosen optimally, power plants would already operate in ways that internalize air pollution externalities. Accordingly, the ETS might be "over-correcting" at that specific location. However, this seems unlikely for two reasons. First, local policymakers are heavily constrained by harmonized European regulations that already target these power plants. Second, our estimates reveal that existing regulations remain sub-optimal as the marginal health benefits of further abatement significantly exceed the marginal costs. In this context, the carbon price acts as a corrective co-benefit, moving the system toward a social optimum that the existing mix of local and European regulation has failed to achieve.

7 Extensions

We intend to extend our analysis to shed light on two more relevant aspects.

Decomposition of co-benefits The sizable health co-benefits we find are the product of complex interactions between the market-based policy instrument and heterogeneities in

[24] Carbon Capture and Storage (CCS) is a technology that would alter the relationship between fuel burned and CO_2 eventually emitted into the atmosphere. However, CCS was not applied at any power plant in Europe during the late 2010s and early 2020s. Europe's first, rather experimental, post-combustion CCS plant opened in Italy in late 2024.

the regulated facilities with respect to (i) how generating units respond to CO_2 abatement incentives, (ii) their co-pollutant intensity, (iii) their geographic location and (iv) their impact on population exposure. A consequence of these multiple heterogeneities is that our quantitative estimates do not necessarily carry over to cap-and-trade programs in other parts of the world. More generally speaking, it cannot be taken for granted that co-pollution benefits that accrue to communities living downwind from plants that curb generation in response to policy outweigh co-pollution costs incurred in other communities that are downwind from plants that increase generation.

To better understand under which conditions we may expect to obtain similar co-benefits in other policy contexts, we decompose our estimates with respect to the fuel, region, and pollution intensity, population density.

Electricity prices For a more complete analysis of welfare impacts, it is important to also understand the effect of the EU ETS on electricity prices. Assuming complete pass-through of compliance costs to supply bids, we will estimate price effects in the day ahead market by examining counterfactual merit curves and find market clearing prices in each bidding zone.

8 Conclusion

In this paper, we develop an empirical framework for quantifying the ancillary health co-benefits of carbon pricing in the electricity sector, applied to the EU ETS over the period from 2016 to 2021. Combining a machine learning model of electricity dispatch, unit-level co-pollutant emission intensities, and a chemical transport model linking emissions to population exposure, we uncover the effects of carbon pricing on premature mortality due to changes in air pollutant emissions.

Our analysis of electricity generation reveals that carbon pricing induces substantial shifts in the fuel mix. Relative to a counterfactual world without carbon pricing, the EU ETS causes an 18% reduction in hard coal generation that is offset by a 21% increase in natural gas generation, translating into a 5.5% reduction in total CO_2 emissions. A hypothetical €50 price floor additionally sparks substitution away from lignite and increases total CO_2 abatement to 12.5% relative to a zero carbon price.

An important novel feature of our empirical framework is the modeling of cross-border electricity flows to predict how generation relocates across zones under different carbon prices. Germany offers the most striking illustration because it is a lignite- and coal-heavy country surrounded by several gas-heavy neighbors. As the carbon price rises, German domestic lignite and coal generation falls by a combined 185 TWh under the €50 price

floor, a reduction of 20%, which is offset approximately fifty-fifty by higher domestic natural gas output and by a surge in electricity imports from neighboring countries whose gas-fired plants become the marginal suppliers for German demand.

By incentivizing fuel switching from coal to natural gas, both within as well as across interconnected bidding zones, the EU ETS has redistributed air pollution in Europe. As a central result, we find that carbon pricing in addition to its primary objective of reducing CO_2 emissions, reduced PM-related premature deaths by about 7.5%, almost 1,000 deaths, relative to a counterfactual world without carbon pricing. Evaluated at a conservative VSL of €3 million, this translates into gross health co-benefits of about €20.6 / t CO_2 abated, well in excess of the €4.3 in abatement costs / t CO_2. In total, we calculate that the EU ETS provided net co-benefits of about €2.2 billion over our six year observation period. We take this as evidence that the decentralized responses to carbon pricing have reinforced the beneficial effects of existing air pollution regulations in the European power sector.

We further show that a carbon price floor of €50 per tonne of CO_2 would have delivered substantially larger co-benefits in absolute terms, avoiding an additional 1,291 premature deaths and adding an additional €1.8 billion in net benefits. This counterfactual experiment also reveals an important asymmetry: Marginal abatement costs rise as the carbon price increases and triggers more costly fuel switching, whereas marginal health co-benefits remain broadly stable. As a result, net co-benefits per tonne of CO_2 abated decline from about €16 under observed ETS prices to about €10 for the incremental abatement achieved under the €50 price floor, and will plausibly continue to diminish as carbon prices rise further in the future.

Our findings carry several implications for climate and air quality policy. First, they confirm that the EU ETS, in addition to reducing carbon emissions, has delivered local health dividends by internalizing air pollution externalities that pre-existing regulation had failed to fully address. The magnitude of these co-benefits suggests that evaluations of this policy instrument should also place significant weight on ancillary health benefits, particularly in coal-heavy electricity systems. Second, our results caution against assuming that these co-benefits will automatically scale with carbon prices in the long run. As marginal abatement costs rise and the remaining scope for low-cost fuel switching narrows, the realization of further health gains will increasingly depend on broader structural changes to the energy system that lie beyond the scope of our analysis of the short-run dispatch margin.

Several caveats deserve emphasis. Our analysis focuses on the intensive margin of electricity dispatch and abstracts from longer-run investment responses, which may have amplified the effects we document. Moreover, while we find that the health co-benefits of the EU ETS substantially exceed its abatement costs, this result rests on the assumption

that pre-existing air quality regulations are insufficient to close the gap between private and social costs, which we cannot verify directly. One can expand on our work by exploring the distributional dimension of our findings in more detail, examining how the health benefits from carbon pricing may differ across communities, and how these differences relate to socioeconomic variables.

References

Abrell, J., Kosch, M., and Rausch, S. (2022). How effective is carbon pricing?—a machine learning approach to policy evaluation. *Journal of Environmental Economics and Management*, 112:102589.

Agency, E. E. (2026). Emissions and energy use in large combustion plants in europe. Available online at Emissionsandenergyuseinlargecombustionplantsin Europe. Last accessed March 23, 2026.

Aunan, K., Berntsen, T., O'Connor, Persson, T. H., Vennemo, H., and Zhai, F. (2007). Benefits and costs to China of a climate policy. *Environment and Development Economics*, 12:471–497.

Borenstein, S. and Bushnell, J. B. (2022). Do two electricity pricing wrongs make a right? cost recovery, externalities, and efficiency. *American Economic Journal: Economic Policy*, 14(4):80–110.

Burtraw, D., Krupnick, A., Palmer, K., Paul, A., Toman, M., and Bloyd, C. (2003). Ancillary benefits of reduced air pollution in the US from moderate greenhouse gas mitigation policies in the electricity sector. *Journal of Environmental Economics and Management*, 45(3):650–673.

Butters, R. A., Dorsey, J., and Gowrisankaran, G. (2025). Soaking Up the Sun: Battery Investment, Renewable Energy, and Market Equilibrium. *Econometrica*, 93(3):891–927.

Carlson, C., Burtraw, D., Cropper, M., and Palmer, K. (2000). Sulfur dioxide control by electric utilities: What are the gains from trade? *Journal of Political Economy*, 108(6):1292–1326.

Center For International Earth Science Information Network-CIESIN (2018). Gridded population of the world, version 4 (gpwv4): Population count, revision 11 (version 4.11) [data set]. Technical report, Columbia University, Palisades, NY. NASA Socioeconomic Data and Applications Center (SEDAC).

Colmer, J., Martin, R., Muûls, M., and Wagner, U. (2024). Does Pricing Carbon Mitigate Climate Change? Firm-Level Evidence From the European Union Emissions Trading System. *Review of Economic Studies*, 92(3):1625–1660.

Cushing, L., Blaustein-Rejto, D., Wander, M., Pastor, M., Sadd, J., Zhu, A., and Morello-Frosch, R. (2018). Carbon trading, co-pollutants, and environmental equity: Evidence from california's cap-and-trade program (2011–2015). *PLOS Medicine*, 15(7):1–20.

Davis, D., Krupnick, A., and McGlynn, G. (2000). Ancillary benefits and costs of greenhouse gas mitigation - an overview. In OECD, editor, *Ancillary Benefits and Costs of Greenhouse Gas Mitigation*, pages 9–49, Paris.

Driscoll, C. T., Buonocore, J. J., Levy, J. I., Lambert, K. F., Burtraw, D., Reid, S. B., Fakhraei, H., and Schwartz, J. (2015). Us power plant carbon standards and clean air and health co-benefits. *Nature Climate Change*, 5:535–540.

Egerer, J. (2016). Open source electricity model for germany (elmod-de). Working Paper No. 83, German Institute for Economic Research.

Ekin, P. (1996). The secondary benefits of CO2 abatement: How much emission reduction do they justify? *Ecological Economics*, 16(1):13–24.

Ekins, P. (1996). How large a carbon tax is justified by the secondary benefits of CO2 abatement? *Resource and Energy Economics*, 18(2):161–187.

Ellerman, D., Marcantonini, C., and Zaklan, A. (2016). The european union emissions trading system: Ten years and counting. *Review of Environmental Economics and Policy*, 10(1):89–107.

Fabra, N. and Reguant, M. (2014). Pass-Through of Emissions Costs in Electricity Markets. *American Economic Review*, 104(9):2872–2899.

Fowlie, M., Holland, S. P., and Mansur, E. (2012). What do emissions markets deliver and to whom? Evidence from southern california's nox trading program. *American Economic Review*, 102(2):965–993.

Grainger, C. and Ruangmas, T. (2018). Who Wins from Emissions Trading? Evidence from California. *Environmental & Resource Economics*, 71(3):703–727.

Grossi, L., Heim, S., and Waterson, M. (2017). The impact of the German response to the Fukushima earthquake. *Energy Economics*, 66:450–465.

Gu, Y., Henze, D. K., Nawaz, M. O., Cao, H., and Wagner, U. J. (2023). Sources of pm2.5-associated health risks in europe and corresponding emission-induced changes during 2005-2015. *GeoHealth*, 7(3):e2022GH000767.

Henze, D. K., Hakami, A., and Seinfeld, J. H. (2007). Development of the adjoint of geos-chem. *Atmospheric Chemistry and Physics*, 7(9):2413–2433.

Hernandez-Cortes, D. and Meng, K. C. (2023). Do environmental markets cause environmental injustice? evidence from california's carbon market. *Journal of Public Economics*, 217:104786.

Hintermann, B. (2016). Pass-Through of CO2 Emission Costs to Hourly Electricity Prices in Germany. *Journal of the Association of Environmental and Resource Economists*, 3(4):857–891.

Hirth, L., Muehlenpfort, J., and Bulkeley, M. (2018). The entso-e transparency platform – a review of europe's most ambitious electricity data platform. *Applied Energy*, 225.

International Energy Agency (2026). Electricity 2026: Analysis and forecast to 2030. Technical report, International Energy Agency (IEA), Paris, France.

IPCC (2023). *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, Geneva, Switzerland.

Jarvis, S., Deschenes, O., and Jha, A. (2022). The private and external costs of Germany's nuclear phase-out. *Journal of the European Economic Association*, 20(3).

Lundy, S. . (2019). Generating unit annual capital and life extension costs analysis.

Martin, R., Muûls, M., and Wagner, U. J. (2016). The impact of the european union emissions trading scheme on regulated firms: What is the evidence after ten years? *Review of Environmental Economics and Policy*, 10(1):129–148.

Mauzerall, D. L., Sultan, B., Kim, N., and Bradford, D. F. (2005). NOx emissions from large point sources: variability in ozone production, resulting health damages and economic costs. *Atmospheric Environment*, 39(16):2851–2866.

Muller, N. Z. and Mendelsohn, R. (2009). Efficient Pollution Regulation: Getting the Prices Right. *American Economic Review*, 99(5):1714–1739.

Murray, C. and *et al.* (2020). Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The Lancet*, 396(10258):1223–1249.

Murray, C. J. L., Aravkin, A. Y., Zheng, P., Abbafati, C., Abbas, K. M., and Abbasi-Kangevari, M., e. (2020). Global burden of 87 risk factors in 204 countries and territories, 1990–2019: A systematic analysis for the global burden of disease study 2019. *The Lancet*, (10258):1223 – 1249.

Nawaz, M. O. and Henze, D. K. (2020). Premature deaths in Brazil associated with long-term exposure to PM2.5 from Amazon fires between 2016 and 2019. *GeoHealth*, 4(8):e2020GH000268.

Open Power System Data (2020). Data package conventional power plants. version 2020-10-01.

Pittel, K. and Rübbelke, D. T. G. (2008). Climate policy and ancillary benefits: A survey and integration into the modelling of international negotiations on climate change. *Ecological Economics*, 68(1-2):210–220.

Rypdal, K., Rive, N., Åström, S., Karvosenoja, N., Aunan, K., Bak, J. L., Kupiainen, K., and Kukkonen, J. (2007). Nordic air quality co-benefits from European post-2012 climate policies. *Energy Policy*, 35(12):6309–6322.

van Donkelaar, A., Hammer, M. S., Bindle, L., Brauer, M., Brook, J. R., Garay, M. J., Hsu, N. C., Kalashnikova, O. V., Kahn, R. A., Lee, C., Levy, R. C., Lyapustin, A., Sayer, A. M., and Martin, R. V. (2021). Monthly global estimates of fine particulate matter and their uncertainty. *Environmental Science & Technology*, 55(22):15287–15300.

van Vuuren, D. P., Cofala, J., Eerens, H. E., Oostenrijk, R., Heyes, C., Klimont, Z., den Elzen, M. G. J., and Amann, M. (2006). Exploring the ancillary benefits of the Kyoto Protocol for air pollution in Europe. *Energy Policy*, 34(4):444–460.

Zaklan, A. (2023). Coase and Cap-and-Trade: Evidence on the Independence Property from the European Carbon Market. *American Economic Journal: Economic Policy*, 15(2):526–58.

Appendix - For Online Publication Only

A Atmospheric Dispersion and Chemical Transport

We employ the chemical transport model (CTM) developed and applied by Gu et al. (2023) to quantify the sources of $PM_{2.5}$-related health risks in Europe. CTMs simulate the emission, chemical transformation, transport, and removal of atmospheric pollutants using meteorological inputs and emission inventories. Gu et al. (2023) adapt the global atmospheric chemistry model GEOS-Chem with its adjoint sensitivity model and high-resolution emission data to link emissions from specific locations and sectors in Europe to health impacts. In this way, the authors can attribute $PM_{2.5}$ exposure and associated premature deaths in Europe to specific emission sources, including countries, sectors, and precursor species.

The model is structured as follows

1. Forward GEOS-Chem Simulation The forward CTM simulates atmospheric processes driving $PM_{2.5}$ concentrations across Europe. Starting with emissions of precursor pollutants (NO_x, NH_3, SO_2, organic carbon, black carbon, etc.), the forward simulation comprises the advection and diffusion of pollutants by meteorology, the atmospheric chemical reactions forming secondary aerosols, as well as deposition processes that remove pollutants from the atmosphere.

The model is run on a nested European grid ($0.25^\circ \times 0.3125^\circ$ resolution) using the HTAPv3 emissions inventory to capture spatial variability in emissions and pollution. The raw simulation output consists of gridded $PM_{2.5}$ concentrations that, combined with population density, results in gridded population exposure to fine dust.

To reduce model biases, the framework integrates satellite-derived $PM_{2.5}$ concentrations which combine aerosol optical depth measurements with models to estimate surface $PM_{2.5}$ concentrations (van Donkelaar et al., 2021). This helps to correct systematic model biases across regions. A downscaling component is used to improve the model's capability to capture sub-grid variability, providing more spatial information than the model can resolve for, thus mapping surface $PM_{2.5}$ concentrations on a $0.01^\circ \times 0.01^\circ$ resolution (Nawaz and Henze, 2020). This step significantly improves the representation of fine spatial variability in population exposure when combined with fine resolution (1 km) gridded population data from the Center For International Earth Science Information Network-CIESIN (2018).

2. Health Impact Assessment The model converts simulated $PM_{2.5}$ exposure into premature mortality estimates using concentration–response relationships from

epidemiological studies. Specifically, Gu et al. (2023) combine population-weighted $PM_{2.5}$ exposure from the CTM with age-stratified exposure-response functions for Global Burden of Disease (GBD) mortality covering chronic obstructive pulmonary disorder (COPD), ischemic heart disease (IHD), lower respiratory illnesses (LRI), lung cancer (LC), type-II diabetes (T2) and stroke (Murray et al., 2020). Aggregate mortality impacts across Europe are summarized in the cost function

$$J_{PM_{2.5}} = \sum_{L} \sum_{A} \sum_{k \in D} \sum_{(I,J) \in k} (POP_{I,J,A} \times MOR_{I,J,A,L} \times AF_{I,J,A,L}) \tag{A.1}$$

where $AF_{I,J,A,L} = \frac{RR_{I,J,A,L}-1}{RR_{I,J,A,L}}$ denotes the attributable fraction, i.e., the excess mortality within a grid cell (I, J) and age bracket A that is due to disease $L \in$\{COPD, IHD, LRI, LC, T2D, stroke\}

3. Adjoint Sensitivity Model A central feature of the framework is the GEOS-Chem adjoint (Henze et al., 2007), which calculates the sensitivity of a cost function (here, $PM_{2.5}$-related premature deaths) to changes in emissions. Instead of running many forward simulations with perturbed emissions, the adjoint CTM offers a much more efficient way of computing sensitivities of the cost function (A.1) to all emissions simultaneously. The goal is to determine how emissions influence this quantity.

To explain this in a nutshell, consider the forward CTM simulation. The chemical transport model simulates pollutant emissions, transport, chemistry, and deposition:

$$\mathbf{x}_{t+1} = \mathcal{M}_t(\mathbf{x}_t, \mathbf{E}_t)$$

where $\mathbf{x}_t$ denotes pollution concentrations, $\mathbf{E}_t$) denotes emissions and $\mathcal{M}_t$ is the model operator describing transport and chemistry.

The object of interest is the gradient of the objective function with respect to emissions:

$$\frac{\partial J}{\partial E_i}$$

At the end of the forward simulation, the adjoint variables are initialized based on how the objective depends on the model state:

$$\boldsymbol{\lambda}_T = \frac{\partial J}{\partial \mathbf{x}_T}$$

The adjoint model propagates sensitivities backward in time through the CTM processes:

$$\boldsymbol{\lambda}_t = \left(\frac{\partial \mathcal{M}_t}{\partial \mathbf{x}_t} \right)^T \boldsymbol{\lambda}_{t+1}$$

This step accounts for how chemistry, transport, and deposition link earlier emissions to later concentrations. During the backward propagation, the sensitivity of the objective to emissions is accumulated:

$$\frac{\partial J}{\partial \mathbf{E}_t} = \left(\frac{\partial \mathcal{M}_t}{\partial \mathbf{E}_t} \right)^T \boldsymbol{\lambda}_{t+1} \tag{A.2}$$

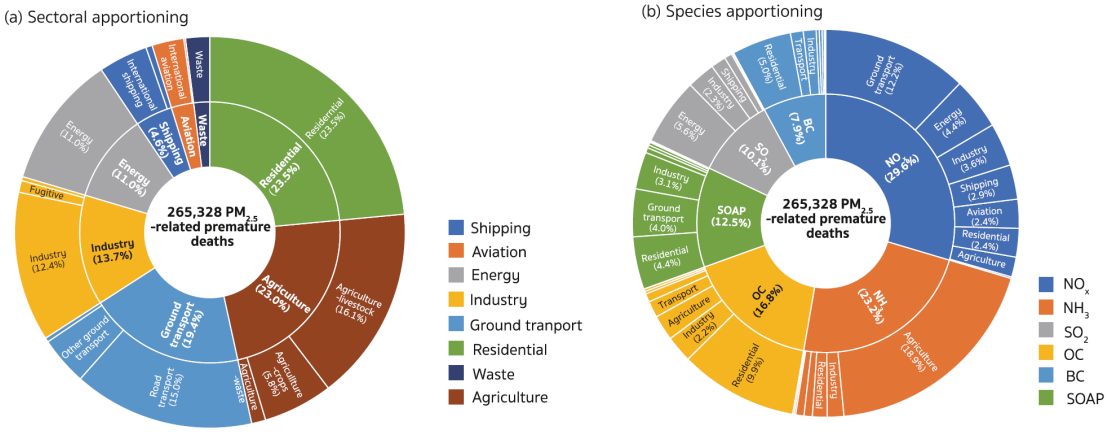
For small emission changes, the impact on the objective can be approximated using a linear expansion:

$$\Delta J \approx \sum_i \frac{\partial J}{\partial E_i} \Delta E_i$$

The formula allows researchers to attribute $PM_{2.5}$-related deaths to specific emission sources (countries, sectors, species). Doing this via the backward model run is computationally much more efficient (more than 11,500 times faster) than forward-modeling based calculations whereby emissions perturbations at different sources are simulated, one by one, in the forward simulation CTM.

4. Mortality and source attribution at baseline Gu et al. (2023) estimate about 449,813 $PM_{2.5}$-related premature deaths in Europe in 2015, relative to a total population of 598,970,000 in the domain.This amounts to a 46% reduction compared to 2005 when the EU ETS was launched. Roughly 59% of those premature deaths are attributable to domestic anthropogenic emissions, and 265,328 deaths (59%) are due to anthropogenic NO_x, NH_3, SO_2, OC, BC, SOAP emissions. Figure A.1 provides further detail on source attribution by sector and by $PM_{2.5}$ precursor pollutant. The energy sector accounts for 21,186 (11%) of $PM_{2.5}$-related premature deaths.

Figure A.1: Mortality due to anthropogenic $PM_{2.5}$ pollution in 2015



Source: Gu et al. (2023).

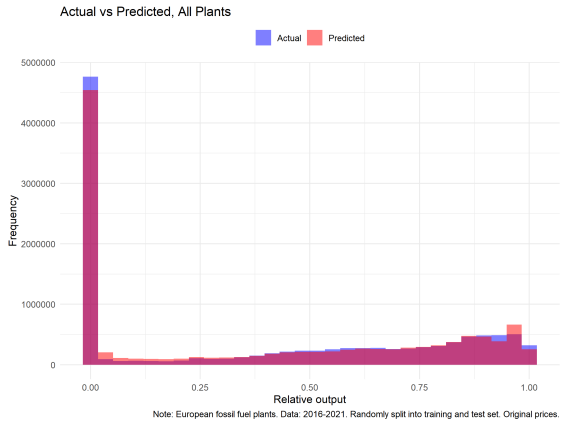
B Additional Tables and Figures

This section collects additional tables and figures that support the analysis in the main body of the paper.

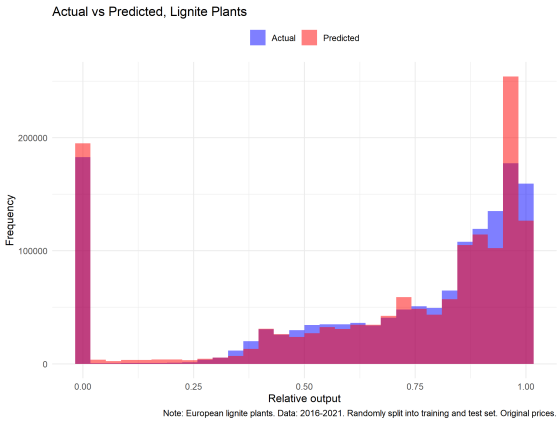
Table B.1: Percentiles per bidding zone and fuel type, averaged across month-of-sample

Bidding zone	**BI**	**CB**	**CH**	**G**
Austria	54.7		54.4	64.5
Belgium	61.8		48.7	54.2
Czechia	56.0	52.8	65.6	58.8
DE-AT-LU	62.8	42.3	53.7	62.3
DE-LU	62.9	51.3	57.3	60.5
Denmark 1	64.0		56.9	64.6
Denmark 2	54.2		47.9	53.4
Spain	47.7	59.1	64.0	76.4
France	64.6		54.9	53.5
IT C North	55.0			53.2
IT C South	46.8		57.8	67.7
IT North	46.8		63.4	62.9
IT Sardinia	68.4		60.0	
IT Sicily				67.8
IT South			76.4	61.9
Netherlands	54.6		45.5	65.7
Poland	59.5	54.5	68.2	49.2
Sweden 3	63.0		45.6	62.1
Sweden 4	49.4			47.4

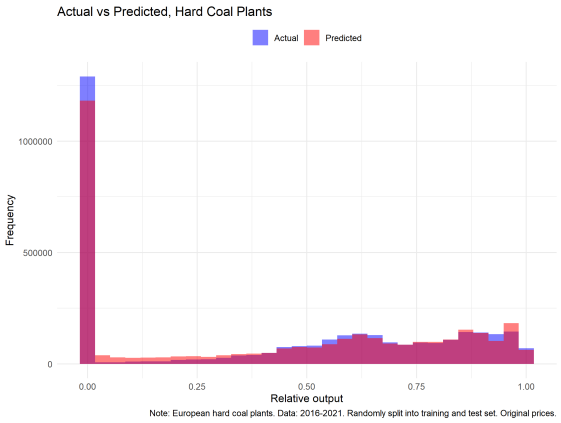
Notes: Averaged across months-of-sample, the table shows the (interpolated) percentile in every bidding zone that we use (on the zone-fuel-month level) to equate actual total output with predicted total output.



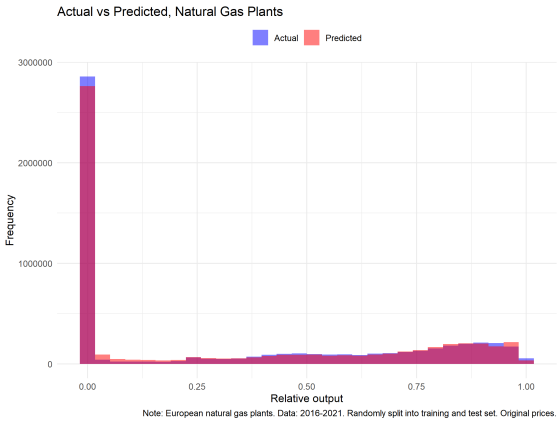
(a) Actual (y-axis) vs. Predicted (x-axis) Utilization Rates, Weighted Percentile Prediction, All Units



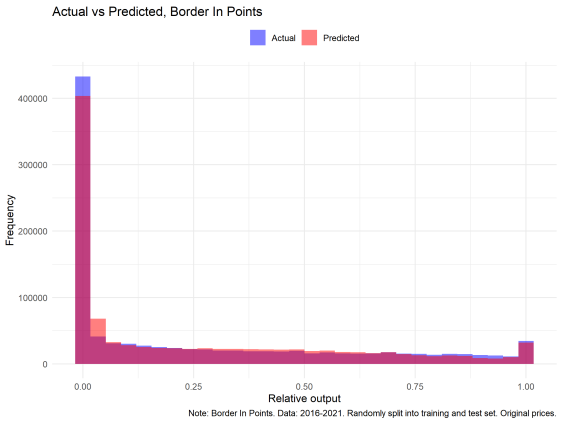
(b) Actual (y-axis) vs. Predicted (x-axis) Utilization Rates, Weighted Percentile Prediction, Lignite Units



(c) Actual (y-axis) vs. Predicted (x-axis) Utilization Rates, Weighted Percentile Prediction, Hard Coal Units



(d) Actual (y-axis) vs. Predicted (x-axis) Utilization Rates, Weighted Percentile Prediction, Natural Gas Units



(e) Actual (y-axis) vs. Predicted (x-axis) Utilization Rates, Weighted Percentile Prediction, Border In Points

Figure B.1: Actual vs. Predicted Relative Output, Weighted Percentile Prediction

Notes: The graphs show histograms of actual relative output (blue) and predicted relative output (red) using the weighted percentile predictions for all units (a) as well as for the four unit types lignite (b), hard coal (c), natural gas (d), and border in points (e).

Table B.2: Change in Electricity Generation and Flows (TWh) due to the EU ETS (Actual Prices)

Country	Lignite	Hard Coal	Natural Gas	Total Fossil	Imports	Exports
Germany	-16.2	-37.9	25.8	-28.3	31.7	-6.2
Poland	-4.3	-16.0	8.1	-12.2	14.7	-2.2
Denmark	—	-12.0	0.4	-11.6	0.7	8.0
Italy	—	-43.4	41.2	-2.2	11.1	0.5
Sweden	—	0.1	-0.4	-0.3	-11.3	5.8
Austria	—	-0.4	6.3	5.9	9.3	6.1
Belgium	—	0.0	6.2	6.2	5.0	16.6
Czechia	4.6	-3.3	5.5	6.8	2.1	4.5
Spain	0.0	-90.9	100.5	9.7	6.9	7.1
Netherlands	—	-10.2	20.5	10.3	4.7	12.7
France	—	-2.6	16.4	13.8	5.6	27.5

Notes: The numbers denote the changes in TWh when going from the €0 allowance price scenario to actual generation under actual allowance prices. Negative values in Imports/Exports indicate a decrease compared to the €0 scenario.

Table B.3: Change in Electricity Generation and Flows (TWh): €0 to €50 Scenario

Country	Lignite	Hard Coal	Natural Gas	Total Fossil	Imports	Exports
Germany	-109.7	-75.7	86.4	-99.1	71.6	-18.6
Poland	-14.4	-9.3	9.6	-14.1	12.9	-0.8
Denmark	—	-8.2	2.2	-6.0	39.9	31.9
Italy	—	-78.8	76.5	-2.4	8.5	1.6
Sweden	—	0.6	5.1	5.7	6.9	30.2
France	—	-8.5	14.3	5.8	14.3	32.6
Spain	-0.1	-174.9	190.5	15.4	0.7	16.9
Czechia	5.9	-3.6	12.2	14.4	2.8	8.7
Belgium	—	-0.1	15.6	15.5	5.4	22.7
Austria	—	-1.3	23.7	22.4	4.9	4.9
Netherlands	—	-25.6	65.9	40.2	-11.7	26.3

Notes: The numbers denote the changes in TWh when comparing a €0 carbon price scenario to a hypothetical €50 allowance price scenario.

Table B.4: Electricity Generation by Fuel and Scenario (TWh)

	Lignite			Hard Coal			Natural Gas		
Country	€0	Actual	≥€50	€0	Actual	≥€50	€0	Actual	≥€50
Germany	538.6	522.3	428.8	350.3	312.4	274.6	147.6	173.4	233.9
Italy	—	—	—	130.9	87.5	52.1	415.3	456.5	491.8
Poland	219.1	214.8	204.7	305.0	289.0	295.7	29.3	37.4	38.8
Netherlands	—	—	—	131.2	121.0	105.6	175.7	196.2	241.6
Spain	0.2	0.2	0.1	216.9	126.0	41.9	99.6	200.1	290.1
Czechia	87.8	92.4	93.7	7.8	4.6	4.2	6.9	12.4	19.1
Belgium	—	—	—	0.3	0.3	0.2	84.1	90.2	99.7
France	—	—	—	23.3	20.6	14.8	96.8	113.2	111.1
Austria	—	—	—	4.7	4.3	3.4	28.5	34.8	52.2
Denmark	—	—	—	34.0	21.9	25.8	2.4	2.8	4.6
Sweden	—	—	—	0.6	0.7	1.2	2.5	2.1	7.6

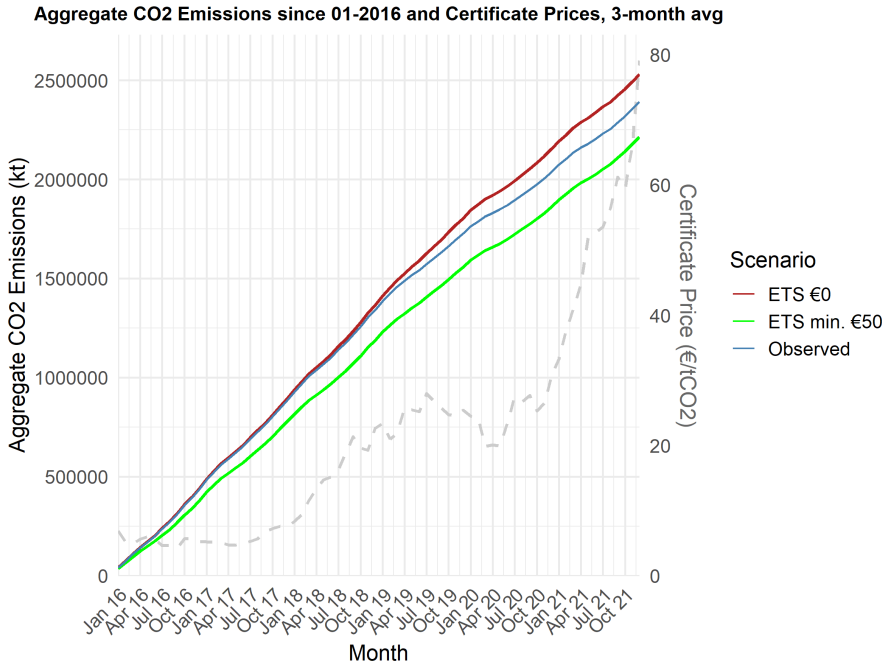
Notes: Numbers denote generated TWh of electricity by fuel across the countries in our sample, and across the three scenarios (€0 allowance price, actual allowance prices, €50 allowance price floor).

Table B.5: CO_2 Emissions by Scenario and Absolute Change (Mt)

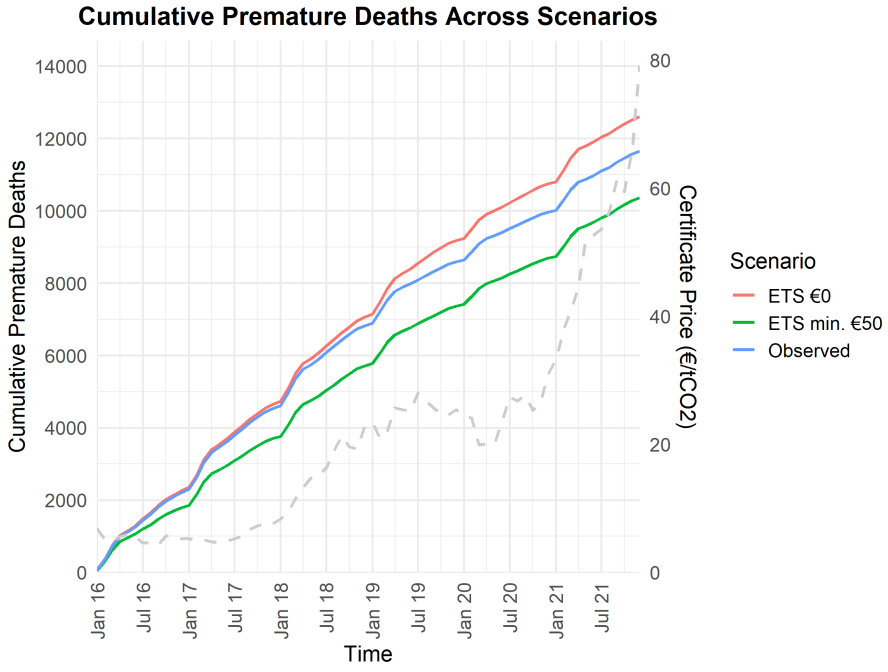
	Scenarios (Mt)			Change vs. €0	
Country	**€0**	**Actual**	**≥€50**	**ΔActual**	**Δ≥€50**
Germany	1003.5	959.3	835.7	-44.2	-167.8
Poland	560.4	542.8	537.0	-17.6	-23.4
Italy	282.7	257.4	237.6	-25.3	-45.1
Spain	249.7	198.6	151.6	-51.1	-98.1
Netherlands	181.3	181.3	185.8	0.0	4.5
Czechia	104.0	109.0	112.8	5.0	8.8
France	60.3	64.0	57.2	3.7	-3.1
Belgium	34.0	36.4	39.9	2.4	5.9
Denmark	35.5	22.7	26.8	-12.8	-8.7
Austria	16.7	18.8	24.7	2.1	8.0
Sweden	1.8	1.8	4.9	0.0	3.1
Total	**2529.9**	**2392.1**	**2193.4**	**-137.8**	**-336.5**

Notes: Δ columns represent the absolute change in Mt relative to the €0 baseline. Negative values represent emission reductions.

Figure B.2: Cumulative CO_2 emissions and premature deaths due to PM_2 across scenarios



(a) Cumulative CO_2 emissions



(b) Cumulative premature deaths due to PM_2

C Robustness

C.1 The share of fossil generation in border flows

We make the simplifying assumption that all border flows correspond to electricity generated by fossil units. To assess the credibility of this assumption we look at the activity of fossil units and the volume of electricity generated using fossil fuels relative to exported electricity. The rationale for this is the following. In the merit order, non-dispatchable units with (near-)zero marginal costs are always prioritized over fossil units. Within each bidding zone, the local market operator constructs a supply curve according to the merit order so that when domestic non-dispatchable generation cannot meet total demand, fossil units are activated to fill the gap. Once all the local market operators have established these hypothetical "pseudo-autarchic" equilibria, a secondary mechanism determines cross-border flows: if Zone A's hypothetical pre-trade clearing price is lower than Zone B's, Zone A exports to Zone B until prices equalize or transmission capacity is saturated. Crucially, if Zone A already required fossil generation to meet its own domestic demand, any subsequent exports to Zone B must be supplied by additional fossil units.

Table C.1: Share of Inactive Fossil Units by Bidding Zone

Zone	**Inactive (%)**		**Zone**	**Inactive (%)**
SE 4	90.15%		IT-CSOUTH	0.03%
DK 2	49.91%		BE	0.00%
SE 3	49.04%		CZ	0.00%
DK 1	10.94%		DE_AT_LU	0.00%
IT-Sicily	3.52%		DE_LU	0.00%
AT	2.57%		FR	0.00%
IT-Sardinia	1.97%		IT-NORTH	0.00%
IT-CNORTH	0.21%		IT-SOUTH	0.00%
ES	0.14%		NL	0.00%
			PL	0.00%

Table C.1 shows, for every bidding zone, the share of hours-of-sample during which every fossil unit in the zone is inactive. If all fossil units are inactive in a given hour, any electricity exports to neighboring zones must be non-fossil. One can see that this share is zero or near zero for most zones. A few zones are clear outliers though. The Southern Swedish zone SE 4 has zero fossil generation during most hours-of-sample because its only two natural gas units have been permanently inactive since December 2016. The Central-Southern Swedish zone SE 3 and the Eastern Danish zone DK 2 have zero fossil generation during close to 50% of hours-of-sample, respectively, because both zones feature only a handful of fossil units which

only operate sporadically.

Table C.2: Negative Fossil Export Share by Bidding Zone

Zone	Negative Share (%)		Zone	Negative Share (%)
SE 3	98.59%		IT-CSOUTH	17.14%
DK 1	87.47%		CZ	11.37%
SE 4	85.46%		DE_LU	9.31%
DK 2	82.54%		ES	4.90%
FR	74.12%		IT-NORTH	4.13%
IT-CNORTH	51.32%		DE_AT_LU	0.40%
AT	38.35%		NL	0.23%
IT-SOUTH	24.30%		PL	0.00%
BE	20.62%			
IT-Sardinia	20.44%			

Since we also observe electricity exports, we calculate the difference between domestic fossil generation and electricity exports to other zones. If the assumption holds, exports will never be larger than fossil generation, so that the difference should be positive. As Table C.2 shows, our assumption is violated quite frequently. The Danish and Swedish zones see export volumes exceeding fossil generation in the overwhelming majority of cases. The same holds for France, which generates huge volumes of nuclear energy. At the same time, there is relatively little violation of our assumption in Poland, the Netherlands, the German-Austrian-Luxembourgian / German-Luxembourgian zone, Northern Italy, Czechia. All of these zones are rely relatively heavily on fossil fuels and have a high degree of interconnectedness, so that they are a natural focal point of our analysis, and it is positive that our crucial assumption about fossil border flows seems to generally hold in these core zones. The same is true for Spain due to its low transmission capacities with France, and Spain is also of particular interest in the analysis due to its vast natural gas capacities that potentially allow for large-scale substitution from coal to gas under high counterfactual ETS allowance prices.

Finally, we address how significant possible violations of our assumption could be in practice. For instance, we may observe that fossil generation in zone A is 1,500 MWh in a given hour, while exports amount to 2,000 MWh. In this case, because of the merit-order principle, we can assume that 1,500 of these 2,000 MWh of exports are fossil and only the remaining 500 are non-fossil. Accordingly, if fossil generation is large relative to exports, this would lead to relatively minor violations of our assumption, and we can maintain that small changes in electricity exports in the counterfactual scenario will be entirely driven by changes in fossil exports. Table C.3 shows how large fossil generation is relative to exports when exports are larger than fossil generation. The closer this is to 100%, the more

Table C.3: Share of Fossil Generation in Exports by Bidding Zone

Zone	Fossil Share (%)		Zone	Fossil Share (%)
DE_AT_LU	89.52%		IT-CSOUTH	57.58%
NL	81.16%		IT-CNORTH	44.98%
CZ	79.44%		AT	37.06%
DE_LU	71.13%		FR	36.08%
IT-NORTH	69.90%		DK1	29.86%
IT-SOUTH	68.73%		DK2	11.67%
ES	67.85%		SE4	2.93%
IT-Sardinia	64.58%		SE3	2.54%
BE	59.15%		PL	N/A

plausible is the assumption that border flows correspond to fossil generation. As above, we see in central countries (Germany, the Netherlands, Czechia, Northern Italy) that (i) the assumption is violated in very few instances and (ii) violations are minor when they occur. In contrast, the assumption is not only frequently violated in Sweden and Denmark, but also, as the table shows, strongly violated.

C.2 Threats to Identification

C.2.1 Fleet Composition and Entry/Exit

In our counterfactual analysis, we maintain a constant fleet of fossil fuel power units based on the composition observed during the 2016–2021 period. Thereby, we focus on the short-term impact of carbon pricing on electricity generation, in line with other literature on the electricity sector (Abrell et al., 2022; Butters et al., 2025).

The exit of fossil fuel power units is heavily influenced by regulatory frameworks independent of the carbon price mechanism. For instance, several coal plants were decommissioned because of their noncompliance with stringent limits on pollutant emissions in the EU Industrial Emissions Directive. Likewise, national coal phase-out trajectories are often politically mandated and negotiated years in advance, tied to fixed decommissioning dates, and designed to ensure long-run grid stability. This renders the exit of coal units inherently insensitive to the secondary market signals provided by the EU ETS within our observation window.

Similarly, the entry of new generation units is constrained. Due to the same coal phase-out trajectories, there has been no scope to construct new lignite or hard coal power units even under a zero carbon price. At the same time, the build-up of new natural gas units is predominantly a policy-mandated transition to provide system flexibility as the share of

renewable energy rises, rather than a market-led expansion of the fossil fuel fleet. Beyond these policy constraints, the economic incentive for large-scale entry of new natural gas units appears limited by the significant latent capacity within the existing fleet. Our model predicts that even under a €50 allowance price floor, the average utilization rate for natural gas units across the sample only reaches 40%, up from 27% in the zero-price scenario. This indicates that the existing infrastructure remains far from a capacity threshold on average, suggesting that the system can accommodate the predicted fuel substitution. The only exception to this are Poland and Czechia in three and two years of our sample, respectively, where we observe predicted utilization rates of the natural gas fleet of above 70%, approaching the "practical maximum" (100% minus planned maintenance + forced outages + operating constraints). Still, local entry would only strengthen our results given that it would lead to broader fuel substitution and therefore larger effects. If anything this renders our estimates conservative.

Second, the decision to remain active in the energy market is primarily driven by short-run marginal costs. In the day-ahead wholesale market, operators place bids based on costs for fuel, carbon allowances, and variable operating expenditures since their capital and investment costs are already sunk. This implies that a power unit will continue to operate as long as the market clearing price exceeds its short-run marginal cost. An infamous example of this is the German hard coal unit Datteln IV. By the time it opened in 2020 with a nine-year delay, the German coal phase-out was already in progress, yet the new unit still opened because the operator (Uniper) argued that the sunk costs (approx. €1.5 billion) were too massive to write off, a decision that was clearly decoupled from the 2020 allowance price.

Third, the technical nature of existing fossil fuel infrastructure reinforces the stability of the fleet across different carbon price scenarios. The power plants themselves are high-value assets designed for multi-decade lifespans. By maintaining a constant fleet in our counterfactuals, we isolate the transformative impact of the EU ETS on short-term dispatch optimization and fuel substitution without introducing the speculative bias of long-term investment modeling. Lastly, at least in the high allowance price counterfactual, the model should be able to capture exits of lignite and coal units to an extent given that it can consistently predict zero generation at these units.

C.2.2 Confounding Regulatory Policies

In addition to the EU ETS, there are various national policies in place in the countries in our sample. As we argue, fortunately, our observation period from January 2016 to December 2021 is characterized by exceptional stability in national "top-up" policies that allow us to isolate the effect of the ETS allowance price on electricity generation in our data. The UK's Carbon Price Support (CPS), a national carbon price on top of the EU ETS, has

remained stable at £18 per tonne of CO2 throughout entire observation period, creating a time-invariant mark-up on marginal cost for UK power units. Sweden and Denmark have a national carbon tax, yet installations are generally exempt from this tax if they are already covered by the EU ETS.

C.2.3 Assumption of constant fuel prices

A potential concern in our counterfactual analysis is the assumption of constant fuel prices across different allowance price scenarios. One might argue that a significant increase in the EU-wide carbon price shifts demand sufficiently to put pressure on global fuel prices. However, we follow the approach of previous work such as Grossi et al. (2017), Abrell et al. (2022) and Jarvis et al. (2022) by treating fuel prices as exogenous to the electricity sector. This assumption is grounded in the structural reality that, during our 2016–2021 sample period, European fuel prices were largely determined by global benchmarks and long-term supply frameworks rather than local carbon-driven demand shifts.

During that time, European natural gas prices remained heavily influenced by global oil benchmarks through long-term, oil-indexed contracts. Since the oil market is the most liquid and globalized commodity market in the world, and the European fleet of oil-fired power plants is small, the EU ETS plausibly did not move the global oil equilibrium. Moreover, the supply side was characterized by significant structural rigidity; the physical nature of pipeline infrastructure and the prevalence of long-term, fixed contracts created a hold-up problem that insulated base fuel prices from fluctuations in the EU ETS.

C.2.4 Strategic Bidding and Market Power

A further potential threat to identification involves the assumption that power units always bid their true short-run marginal costs. We construct the supply curve and derive predictors for a unit's utilization rate from the supply curve based on our marginal cost estimates that reflect variable operating costs. In a perfectly competitive electricity market, this approach mirrors the true market dynamics since the merit-order principle dictates that operators bid their marginal costs of generation to ensure dispatch. However, in the presence of high market concentration, large utility firms may strategically submit bid above their true short-run marginal costs to exercise market power. Although this implies that we misidentify a unit's position in the merit order based on marginal costs only, the Random Forest can account for this by internalizing strategic markups by training on realized dispatch.

A deeper threat exists if these markups are systematically correlated with allowance price movements. It might be the case that firms would have adjusted their exercise of

market power in the counterfactual high allowance price scenario, for instance by increasing markups on gas units as coal becomes less competitive. In this case, our counterfactual predictions fail to capture this behavioral shift. However, given that our model focuses on the intensive margin of a highly integrated European grid, where cross-border competition often disciplines local market power, we treat these bidding strategies as a structural feature of the training data that remains reasonably stable across our counterfactual price ranges.

Likewise, the model is capable of "learning" the magnitude of systematic deviations where specific units consistently operate differently than their calculated marginal costs would suggest. By training on realized market behavior, the Random Forest endogenizes the dynamics of strategic markups. The identifying assumption is that the relationship between allowance prices and these markups is time-invariant. Under this assumption, the model leverages observed behavior from high-allowance-price periods to inform its predictions for counterfactual high-price scenarios during low-allowance-price periods. As long as the internalized markup dynamics during high-price periods are informative of how firms would have set markups in other times if the allowance price had been high as well, our counterfactual estimates provide a robust reflection of how the fossil fleet responds to carbon prices, rather than a stylized simulation based on the assumption of perfect competition.

We can draw on the relationship between the model's prediction error and the allowance price to assess whether our model systematically misses out strategic bidding. To consistently predict what power units would do if the carbon price was different (€0 or €50 price floor), we must assume that the model's prediction error is independent of our treatment, that is the carbon price itself and the derived metrics that we feed into the model (Abrell et al., 2022). If the model, for instance due to strategic bidding depending on the allowance price, systematically over-predicts output under low allowance prices and under-predicts under high allowance prices (or vice versa), our counterfactual results would be fundamentally biased.

We assess the assumption of the independence of the prediction errors with regard to the simulated carbon price in two ways. In Figure C.1a, we plot the prediction error (y-axis), that is the predicted utilization rate minus the actual utilization rate, against the ETS allowance price (x-axis). The dark red line shows how the average prediction error varies over different allowance prices, while the light red shaded area shows the standard deviation of the prediction error. Both metrics support the aforementioned assumption. The average prediction error fluctuates closely and in a seemingly random manner around zero across all values of the ETS allowance price. The predictive uncertainty also appears approximately constant: The standard deviation of the prediction error does not systematically grow or shrink with the allowance price.

In Figure C.1b, similar to how Abrell et al. (2022) assess this assumption in their analysis

of the UK carbon tax, we calculate the Pearson correlation coefficient between the prediction error and the ETS allowance price for every power unit in our sample, and then plot the histogram of these correlation coefficients. The graph shows that almost 82% of power units have correlation coefficients of no more than 0.1 in absolute size, which is a negligible amount of correlation. In other words, the carbon price is practically uncorrelated with the prediction error for the vast majority of the fossil fleet, so that the model's predictive precision does not change with the ETS allowance price. In addition to that, the symmetry of the histogram indicates that there is no global systematic bias in how the model handles carbon prices. Overall, the graphs in Figure C.1 show supportive evidence that the model's prediction error is not systematically related to the allowance price.

Beyond that, we also calculate the model's MSE and R^2 separately for every year. We can broadly partition the years of observations into periods of low prices (until early 2018), medium prices (mid-2018 to mid-2020), and high prices (since late 2020). If the predictive power of our model significantly changes between the years of observation, this would indicate that our model might fail to understand how power providers bid into the market and eventually provide electricity. As Table C.4 shows, this is not the case. Both the model's MSE as well as its R^2 are remarkably stable. The fact that neither metric changes substantially towards the end of the sample suggests that the model's predictive framework does not "break down" once the observed carbon prices approach unprecedented levels.

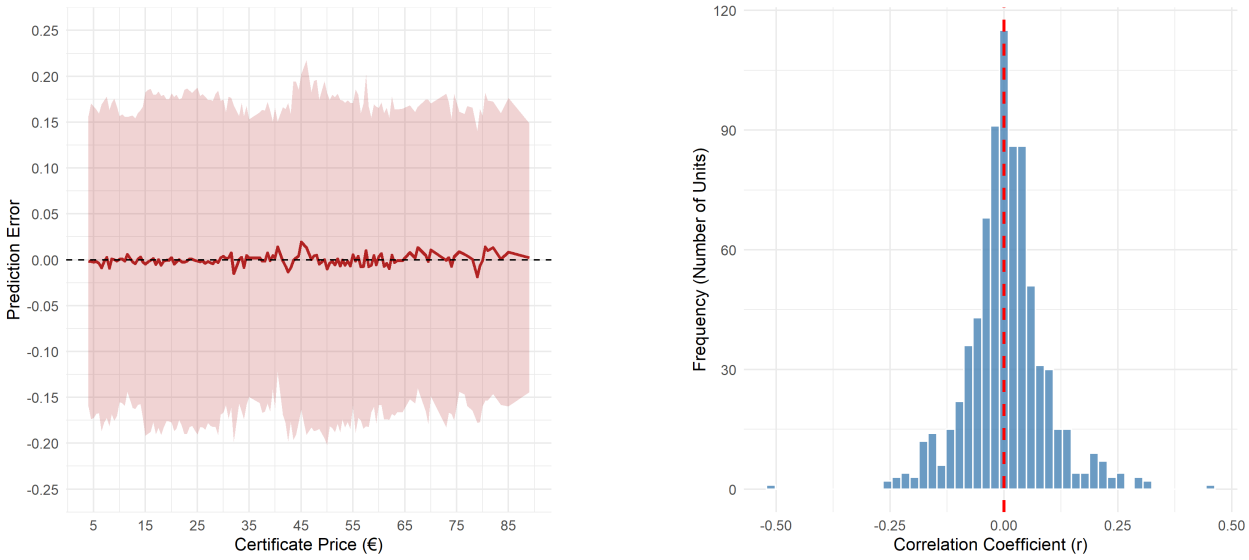
Table C.4: Random Forest Predictive Performance by Year

Year	**MSE**	R^2
2016	0.0268	0.8329
2017	0.0308	0.8098
2018	0.0315	0.8066
2019	0.0350	0.7756
2020	0.0303	0.7917
2021	0.0294	0.8067

Notes: The table reports the predictive performance of the Random Forest model by year. The R^2 and Mean Squared Error (MSE) are calculated based on actual vs. predicted unit-level utilization rates.

C.2.5 Merit Order vs. Actual Dispatch Order

Our empirical strategy feeds the Random Forest with information on the supply curve of fossil power units in the electricity market. Based on this information, the model aims to



(a) Average prediction error by ETS allowance price

(b) Correlation coefficient between prediction error and ETS allowance price, by power unit

Figure C.1: Correlation between prediction error & ETS allowance price

predict hourly utilization rates of power units. We derive a variety of metrics to describe a unit's relative position in the merit order, such as the capacity gap and the output gap as well as a unit's marginal cost and the zone's net load. In addition to that, time-constant unit-level information, such as their net electric efficiency, provides further information of a unit's relative position in the merit order.

The merit order establishes a market equilibrium on the day-ahead market. Unit operators submit supply bids specifying the volume of electricity they are prepared to generate at various price points for each hour of the following day. These bids are aggregated into a supply curve, and the market clearing mechanism subsequently allocates generation to units in ascending order of their bid prices until residual demand is met. However, the dispatch and generation volumes implied by the merit order is not the same as actual unit activity, which is the outcome that we observe in the data and what our model is supposed to predict.

Transmission constraints are an important contributing factor to why realized unit-level generation differs from generation as implied by the merit order. The day-ahead market implicitly assumes a copperplate, that is a frictionless electricity grid in which any unit can satisfy any demand that arises anywhere in the zone. This assumption is unrealistic, in particular in large bidding zones. When there is a local demand peak, it may be that only geographically close units can be dispatched to satisfy that local demand peak due to within-zone transmission constraints. Similarly, when a bottleneck is about to occur somewhere in

the grid, upstream power plants are instructed to reduce their output, while downstream plants must increase their output. This short-term redispatch occurs independently of the merit order and is therefore not captured by our merit-order-based predictors.

We do not observe within-zone transmission lines, and neither do we observe the distribution of demand within a zone. To approximate both, we provide the algorithm with information on a unit's geographical position (latitude and longitude). Transmission constraints are systematic physical properties of a bidding zone's grid that can be treated as approximately constant during our six-year observation window. By providing the Random Forest with information on a unit's geographic position, our aspiration is that the algorithm understands that units in a given area might run more often during peak hours than its marginal cost suggests because they are located downstream from a bottleneck. Similarly, we hope that local demand peaks appear with sufficient regularity such that the algorithm learns its generation patterns based on a unit's geographic position plus time dummies (hour-of-day, day-of-week, month-of-year) for when these local demand peaks tend to occur. In fact, as shown in Figure 6, latitude and longitude are the model's two most important predictors and substantially improve predictive power. These geographic features are especially critical for modeling dispatch in Spain, the bidding zone with the largest geographic area in our sample. Spanish electricity demand is concentrated in a few distant metropolitan centers, creating a system prone to internal transmission constraints.

Because the model predicts realized generation, the model learns the merit order, but also deviations from it. This predictive trade-off is a central feature of our empirical strategy. The alternative would be to model the "pure" day-ahead merit order using bid data. However, since we eventually convert the predicted generation pattern into pollutant emissions, such an approach would fail to capture the total air pollution burden that is determined in part by the physical constraints in the grid. To quantify realized emissions, we must train the model on actual dispatch data to include the "out-of-merit" generation required for grid stability and reliability. By utilizing observed generation, the algorithm is able to internalize the complex interactions between market signals and physical constraints that determine the actual utilization rates of fossil units.

Furthermore, training on ex-post dispatch data allows the model to account for intraday market activity and real-time balancing actions. Because we provide the Random Forest with contemporaneous fuel prices, ETS allowance prices, and actual realized net load, the algorithm can learn short-term dispatch corrections that a theoretical day-ahead model would necessarily miss. This ensures that our counterfactual predictions reflect not only shifts in the primary merit order, but also the secondary operational adjustments essential for maintaining a continuous equilibrium.