ENERGY INSTITUTE AT HAAS

**Energy Institute WP 304**

# The Private and External Costs of Germany's Nuclear Phase-Out

Stephen Jarvis, Olivier Deschenes, and Akshaya Jha

January 2020

# The Private and External Costs of Germany's Nuclear Phase-Out

Stephen Jarvis, Olivier Deschenes, and Akshaya Jha*

**Abstract**

Many countries have phased out nuclear electricity production in response to concerns about nuclear waste and the risk of nuclear accidents. This paper examines the impact of the shutdown of roughly half of the nuclear production capacity in Germany after the Fukushima accident in 2011. We use hourly data on power plant operations and a novel machine learning framework to estimate how plants would have operated differently if the phase-out had not occurred. We find that the lost nuclear electricity production due to the phase-out was replaced primarily by coal-fired production and net electricity imports. The social cost of this shift from nuclear to coal is approximately 12 billion dollars per year. Over 70% of this cost comes from the increased mortality risk associated with exposure to the local air pollution emitted when burning fossil fuels. Even the largest estimates of the reduction in the costs associated with nuclear accident risk and waste disposal due to the phase-out are far smaller than 12 billion dollars.

**JEL Codes:** Q4, Q5, C4

**Keywords:** Nuclear, Electricity, Fossil Fuels, Air Pollution, Machine Learning, Germany

*Jarvis: Energy & Resources Group and Energy Institute at Haas, University of California at Berkeley. Berkeley, CA 94720. Email: jarviss@berkeley.edu. Deschenes: Department of Economics, 2127 North Hall, University of California at Santa Barbara. Santa Barbara, CA 93106. Email: olivier@econ.ucsb.edu. Jha: H. John Heinz III College, Carnegie Mellon University. 4800 Forbes Ave, Pittsburgh, PA 15213. Email: akshayaj@andrew.cmu.edu. The authors thank seminar participants at the University of Wyoming, the University of Texas-Austin, and at the University of Regensburg. The authors also wish to acknowledge the Library at the University of California, Berkeley, which provided support for the completion of this research.

# 1 Introduction

The Fifth Intergovernmental Panel on Climate Change Assessment Report (IPCC 2013) and the 21st United Nations Climate Change Conference ("COP21") have both recommended that nuclear power should be a part of the global solution to climate change. This is because nuclear electricity generation produces minimal carbon emissions under normal operating conditions (Markandya and Wilkinson, 2007). In contrast, burning fossil fuels to produce electricity is known to emit both global pollutants that contribute to climate change and local pollutants that have negative consequences on human health (NRC and NAS (2010); Jaramillo and Muller (2016); Deschenes, Greenstone and Shapiro (2017); Holland et al. (2018)). Despite this, many countries have substantially decreased the share of their electricity production from nuclear sources. For example, Italy, Belgium, Spain, and Switzerland all have policies in place to phase-out nuclear power entirely. This is due in large part to concerns about long-term solutions for storing nuclear waste and public fears of catastrophic nuclear accidents. These fears intensified considerably following the incidents at Three Mile Island in 1979, Chernobyl in 1986, and Fukushima in 2011.

The decision to phase-out nuclear production in many countries seems to suggest that the expected costs of nuclear power exceed the benefits. Yet, there remains considerable uncertainty about some of these costs and benefits as there is a glaring lack of empirical studies quantifying the *full* range of economic and environmental impacts from large-scale nuclear sector closures.

This paper presents a first attempt at filling this important gap by documenting the impact of the phase-out of nuclear power in Germany on multiple market and environmental outcomes. In particular we focus on the shutdown of ten of the seventeen nuclear reactors in Germany that occurred between 2011 and 2017 following the Fukushima accident in Japan. This context affords us several advantages over previous research studying the impacts of nuclear power plants closures. First, and most importantly, Germany shut down over 8 GW of nuclear production capacity over a few months in 2011, representing close to a 5% reduction in total capacity. By 2017 this had increased to a total of 11 GW of closed nuclear production capacity. This is far larger than the reductions in capacity studied by previous research that focused on the shutdown of a small number of nuclear plants in the United States (Davis and Hausman (2016); Severnini (2017)).

Second, Germany plans to shut down all of its remaining nuclear reactors by 2022. Our study thus provides timely policy-relevant information on the consequences of Germany's nuclear phase-out moving forward. Third, studying electricity markets in the European context gives us the opportunity to examine how cross-border trade was impacted by a large shock to production in one country. Finally, Germany's nuclear phase-out was the direct result of political actions taken following extensive anti-nuclear campaigning in Germany as well as a sudden increase in the perceived risk of nuclear power following the Fukushima accident (Goebel et al., 2015). Importantly, the phase-out was not caused by changes in the economic or environmental conditions pertaining to nuclear production in Germany. This facilitates a causal interpretation of our analysis based on comparing the conditional averages of economic and environmental outcomes before versus after the nuclear phase-out.

This paper adds to the relatively small literature that explores the effects of the nuclear phase-out on the German electricity sector. For instance, both Traber and Kemfert (2012) and Knopf et al. (2014) used mixed economic-engineering models of the power sector to forecast changes to capacity investments, electricity prices and carbon emissions. More recently, Grossi, Heim and Waterson (2017) uses an event study framework to econometrically estimate the impact of the initial nuclear plant closures in 2011 on electricity prices over a three year window between 2009 and 2012. The broad consensus across this small existing literature is that nuclear power was replaced primarily by fossil fuel-fired production, resulting in higher electricity prices and more carbon emissions. However, by focusing on aggregate outcomes, the previous research ignores several important impact margins of the nuclear phase-out. Specifically, we show that much of the social cost of the switch from nuclear to fossil fuels is driven by changes in local air pollution concentration levels around individual power plants before versus after the phase-out.

This paper goes beyond the aggregate electricity sector by estimating the economic and environmental costs of the nuclear phase-out in Germany using rich plant-level data and ambient pollution monitor data. We contribute and expand on the existing literature in several important ways. First, our empirical analysis considers both the initial nuclear reactor closures in 2011 as well as the subsequent incremental shutdowns up until the end of 2017. Second, in addition to electricity prices and carbon emissions, we estimate the spatially disaggregated impacts of the phase-out on production costs, net electricity imports, and local air pollution. This is especially important because the increases in local air pollution as a consequence of shifting production from nuclear to coal represents over 70% of the overall costs of the nuclear phase-out.

To proceed, we develop a new machine learning framework to derive the appropriate counterfactual outcomes under a "no phase-out" scenario. Specifically, our machine learning approach predicts which power plants increased their output in response to the nuclear plant closures. In doing so, this paper contributes a new method that builds on Davis and Hausman (2016) in order to empirically assess how a change in electricity production or consumption at one location propagates throughout the electricity transmission network. This new methodology is useful in a number of different empirical contexts. For example, recent studies have explored how production at different fossil fuel-fired plants responds to changes in electricity consumption at a given location, whether it be plugging in an electric vehicle (Holland et al., 2018), installing a more energy efficient appliance, or siting new wind and solar resources (Callaway, Fowlie and McCormick (2018)). Finally, our paper also contributes to the small but growing literature in energy and environmental economics that integrates machine learning into causal inference techniques (Burlig et al. (2017); Cicala (2017)).

Our novel machine learning approach combines hourly data on observed power plant operations between 2010-2017 with a wide range of related information, including electricity demand, local weather conditions, electricity prices, fuel prices and various plant characteristics. Using these data, we first simply document that production from nuclear sources declined precipitously after March 2011. This lost nuclear production was replaced by electricity production from coal- and gas-fired sources in Germany as well as electricity imports from surrounding countries. We then more formally estimate the impact of the nuclear phase-out on market outcomes using our machine learning algorithm. This algorithm predicts the quantity of electricity produced by each power plant in Germany in each hour-of-sample under two scenarios: one with the nuclear phase-out and one without it. Consistent with the aforementioned descriptive trends, the results of this estimation procedure indicate that the lost nuclear electricity production due to the phase-out was replaced primarily by coal-fired production and net electricity imports.

Finally, we use our predicted changes in plant-level electricity production due to the nuclear shutdowns to calculate the costs of the shift away from nuclear power. We first show that the average operating cost per MWh of German electricity production increased as a consequence of the phase-out. This is unsurprising given that nuclear plants have lower marginal costs than fossil fuel-fired plants. In addition, we find that the switch from nuclear power to fossil fuel-fired production resulted in substantial increases in global and local air pollution emissions. Overall, we estimate that the social cost of the phase-out to German producers and consumers is $12 billion per year (2017 USD). The vast majority of these costs fall on consumers. Specifically, over 70% of the cost of the nuclear phase-out is due to the increased mortality risk from local air pollution exposure as a consequence of producing electricity by burning fossil fuels rather than utilizing nuclear sources.

The nuclear phase-out had benefits as well. In particular, shutting down nuclear plants reduces the risk of nuclear accidents and decreases the costs associated with storing nuclear waste (Dhaeseleer (2013); JECR (2019)). However, even the largest estimates of the benefits of the nuclear phase-out are far smaller than our estimated cost of $12 billion dollars a year. Moreover, consistent with previous work, we find that electricity prices in Germany are higher due to the phase-out. This increase in electricity prices results in increases in the profits earned by most electricity producers but imposes additional costs on German electricity consumers.[1]

Despite the substantial costs to German citizens, the nuclear phase-out still has widespread support. Specifically, more than 81% of German residents were in favor of the phase-out in a 2015 survey (Goebel et al., 2015). Existing evidence suggests that the average person greatly overestimates the expected costs of a nuclear accident, both in terms of likelihood and number of fatalities (Slovic, Fischhoff and Lichtenstein (1979); Slovic and Weber (2002); Slovic (2010)). In addition, the health costs associated with local air pollution exposure may simply be less salient than the risk of a nuclear accident, especially after the Fukishima accident in Japan. Regardless of the underlying causes, widespread anti-nuclear sentiment around the world has made it difficult to set policy pertaining to nuclear power based solely on a dispassionate benefit-cost analysis.

This paper proceeds as follows. The next section provides background on the German

[1]Neidell, Uchida and Veronesi (2019) similarly finds an increase in electricity prices due to the phase-out of nuclear power in Japan following the Fukishima accident. This phase-out-induced increase in prices resulted in a decrease in energy consumption, which in turn caused substantial increases in mortality during very cold temperatures.

electricity sector. Section 3 lists the data sources used for this analysis and presents descriptive trends in electricity prices, production by fuel type, costs, air pollution and other outcomes before versus after the nuclear phase-out. In Section 4, we estimate the impact of the phase-out on plant-level and market-level outcomes using a simple event study framework. We describe how our machine learning approach improves upon this event study approach in Section 5. Section 6 presents our estimates of the economic and environmental impacts of the phase-out. Finally, we discuss the policy implications of our findings in Section 7.

## 2 Background on Nuclear Power in Germany

The first nuclear power stations were constructed in Germany in the 1960s. Germany's nuclear production capacity expanded rapidly over the next three decades; the last nuclear reactor was commissioned in 1989. Despite no new reactors coming online in the 1990s and 2000s, roughly 25% of Germany's electricity production came from nuclear plants prior to 2011.

Nuclear power has long been controversial in Germany. There were protests as far back as the 1970s at a number of sites where nuclear facilities were either proposed or under construction. However, the Chernobyl disaster in Ukraine in 1986 created a focal point in the politics of nuclear power in Germany. Specifically, radioactive fallout affected much of the country and led to growing public concern. In 1998, the Schröder government took power through a coalition between the Social Democratic Party (SPD) and the Green Party. Over the next two years, the Schröder government banned the construction of new reactors and negotiated a policy of phasing-out nuclear power completely. This plan called for all nuclear reactors to be shut down by 2022.

The center-right Merkel government came to power in 2009. This government renegotiated the original phase-out policy by committing to extending the lifetimes of the newest reactors. This revised policy pushed back the shutdown of the last nuclear reactor into the 2030s. However, the specter of nuclear disaster rose again due to the Fukushima incident on March 11, 2011. In response, public opposition to nuclear intensified again, with an estimated 250,000 people taking to the streets nationwide to protest in the days and weeks following March 11, 2011. The resulting political pressure forced the Merkel government to declare a moratorium on planned extensions at existing nuclear power plants almost immediately after the Fukushima incident. In addition, eight older reactors were taken offline for testing.

By May of 2011, German policymakers decided to return to a version of the original plan: phase out all nuclear power by 2022. Specifically, of the seventeen reactors operating in 2011, the eight reactors already temporarily offline were closed immediately (8.4 GW of capacity), a ninth reactor was closed in 2015 (1.3 GW), a tenth in 2017 (1.3 GW), an eleventh in 2019 (1.4 GW), and the final six reactors (8.1 GW) will close in 2022. Our sample period ends in 2017. Consequently, our empirical analysis focuses on the closure of the nuclear reactors in 2011, 2015 and 2017, but not the subsequent closures in 2019 and 2022.

The phase-out of nuclear power is part of a wide-ranging transformation of Germany's energy sector known as the *Energiewende*. The primary goal of this policy is to reduce Germany's carbon emissions by at least 80% by 2050 relative to 1990 levels (BMWi, 2018). To achieve this, Germany has undertaken major investments in renewable electricity production, transmission grid infrastructure, and energy efficiency measures. The sweeping scope of the *Energiewende* policy highlights the importance of accounting for a host of potential time-varying confounders when assessing the impact of the nuclear phase-out. This motivates the development of our machine learning approach.

# 3 Data Description and Summary Statistics

This paper brings together the necessary data on the German power sector from a variety of different sources. First, we obtain data on hourly, unit-level electricity production for all power plants with production capacity greater than 100MW. These data are from the European Network of Transmission System Operators for Electricity (ENTSOE) and are only available from 2015-2017. We supplement these data with hourly total production by source (e.g. nuclear, coal, natural gas, oil, etc.) from the European Energy Exchange (EEX) from 2010-2017.

Germany's electricity transmission grid is owned by four different transmission system operators (TSOs) that are each responsible for a different geographical area on the grid: Amprion, TenneT, TransnetBW and 50Hertz. Each TSO reports hourly production from wind and solar sources for the period 2010-2017. The TSOs also provide data on the hourly level of electricity imports and exports in and out of Germany at border points, as well as the hourly total quantity of electricity demanded for their portion of the grid. These TSO data allow us to construct hourly net demand of electricity (total load minus renewable production), as well as hourly generation by source, and net imports at each grid border point.

We construct each plant's marginal cost over time using data on input fuel prices and carbon emission prices gathered from the following two main sources. First, Thomson Datastream provides data on daily natural gas prices in Germany and neighboring countries. The Intercontinental Exchange (ICE) lists monthly coal and oil prices as well as the monthly permit prices for carbon dioxide emissions set by the European Union Emissions Trading System (EUETS).

Our analysis of the environmental costs caused by burning fossil fuels to produce electricity also combines data from multiple sources. The European Environment Agency (EEA) reports annual carbon dioxide emissions for each plant that participates in the EUETS. The EEA also reports annual plant-level data on fuel inputs and local pollution emissions.[2] Station-level weather data comes from Germany's national meteorological service (DWD) and local pollution monitor data are from the German Environment Agency (UBA).

Finally, we compile other electricity sector data and power plant level characteristics from a variety of different sources (Open Power System Data (2018); BNetzA (2018); Egerer (2016)). Most notably, we utilize hourly, Germany-wide wholesale electricity prices from Thomson Datastream.

Taken together, our main estimation sample covers the period 2010-2017 and contains hourly data on wholesale electricity prices, hourly total and net electricity demand, hourly production by dispatchable sources, individual power plant characteristics (including marginal costs of production), and hourly plant-level generation (for the 2015-2017 only).

[Table 1 about here.]

Table 1 provides summary statistics for the electricity sector in 2010 (the first year in our sample) and 2017 (the last year in our sample). The top panel shows that, despite the

[2]These data are collected as part of monitoring for the EU Large Combustion Plant Directive.

closure of more than 10 GW of nuclear capacity between 2010 and 2017, total installed electricity generating capacity grew from 172.2 to 217.6 GW over this period. This is due primarily to rapid growth in renewable production capacity, from 52.1GW in 2010 to 112.5 GW in 2017 (see the bottom panel). Total electricity production increased by roughly 40 TWh between 2010 and 2017. Average wholesale electricity prices also declined precipitously from $70.70 in 2010 to $41.80 in 2017 (in 2017 constant USD). Finally, Germany is a net exporter of electricity throughout our sample period; annual net electricity exports increased from 3.5 TWh in 2010 to 33.5 TWh in 2017.

The middle panel of Table 1 reports summary statistics for the major types of power plants in Germany: nuclear, hard coal, lignite, natural gas, and oil. The extent of the nuclear phase-out in 2011 is immediately evident: production from nuclear sources roughly halved after 2011. At the same time, the number of coal-fired power plants (hard coal and lignite) also dropped due to the closure of older and smaller plants. However, production from coal plants remained roughly constant over our sample period; the small decline in hard coal generation was essentially offset by an increase in lignite generation. The marginal cost of production for both type of coal plants fell significantly during the 2010-2017 period, driven by a reduction in the price of coal. The 2010s were also a period of growth for the gas sector: 26 new plants were built and annual total natural-gas-fired production increased from 53.6 TWh to 72.3 TWh. Appendix Figure A.1 presents a more detailed breakdown of the quantity of electricity produced by different types of sources in Germany over 2010-2017.

[Figure 1 about here.]

Figure 1 shows the estimated marginal cost of each power plant in our sample operating in 2011. We assume that biomass, waste, hydroelectric, wind and solar resources have zero marginal operating cost. We also assume that nuclear plants have a marginal operating cost of approximately $10/MWh (in 2017 USD) based on prior research on Germany's power sector (Egerer, 2016). Finally, marginal costs for fossil fuel plants are calculated as the sum of fuel costs and an assumed amount of variable operating and maintenance costs that differs by fuel type.[3]

[3]Fuel costs are converted to dollars per MWh using the plant's thermal efficiency: how well the plant converts units of input heat to units of electricity output.

Figure 1 highlights that nuclear units uniformly have lower marginal costs than fossil-fuel-fired units. Nuclear power plants also emit virtually no carbon dioxide or local pollutants. We would thus expect that the shutdown of nuclear reactors will lead to increases in both production costs and pollution emissions. We test this hypothesis using a simple event study framework in the next section and our machine learning approach in Section 5.

# 4 Event Study Regressions

In response to the Fukishima nuclear accident, the German government suddenly and unexpectedly shut down eight nuclear reactors on March $15^{th}$ 2011. We can thus analyze the impact of these closures on market outcomes using the event study framework formulated in Davis and Hausman (2016) and more recently implemented by Grossi, Heim and Waterson (2017). Specifically, our event study framework estimates how total electricity production by each fuel type $i$ in each hour-of-sample $t$ responds to changes in electricity demand before versus after March 15th, 2011.

The independent variables of interest are equally-spaced bins of net electricity demand interacted with an indicator for observations after March $15^{th}$ 2011. As in the rest of this paper, "Net electricity demand" is defined to be electricity demand net of production from renewable sources. We consider net demand because production from renewable sources has near-zero marginal costs and is "non-dispatchable": wind and solar sources produce only when the wind is blowing or the sun is out. In order to implement the event-study, we restrict the sample to observations less than 12 months before or after March $15^{th}$ 2011 and estimate the following regression:

$$G_{i,t} = \sum_b (\alpha_{i,b} \cdot \mathbf{1}\{L_t \in B_b\}) + \sum_b (\beta_{i,b} \cdot \mathbf{1}\{L_t \in B_b\}\mathbf{1}\{t \geq 3/15/2011\}) + \gamma_m + \epsilon_{i,t} \quad (1)$$

where $G_{i,t}$ is the total quantity of electricity produced by fuel type $i$ in hour-of-sample $t$ in Germany. $L_t$ is net demand in hour $t$, and $\mathbf{1}\{L_t \in B_b\}$ is an indicator that takes on the value one if $L_t$ is in bin $B_b$ and is zero otherwise. Next, the indicator $\mathbf{1}\{t \geq 3/15/2011\}$ takes on the value one if the observation corresponds to an hour-of-sample on or after

March $15^{th}$ 2011 and is zero otherwise. Finally, we include month-of-year fixed effects (i.e.: $\gamma_m$) and cluster standard errors by week-of-sample.

Figure 2 plots the coefficient estimates of interest (i.e.: $\hat{\beta}_{i,b}$) along with their 95% confidence intervals. Panel (a) of this figure shows that average hourly electricity production from nuclear sources dropped by roughly 5 GWh across all levels of net demand. Panels (b)-(d) demonstrate that this lost nuclear production was offset in large part by increases in electricity production from fossil fuel fired sources. Specifically, production from lignite increased by roughly 1 GWh on average at low levels of net demand. Production from hard coal increased by 2-3 GWh on average across all levels of net demand. Finally, gas-fired electricity generation also increased by roughly 2 GWh on average, and by as much as 6 GWh for hours-of-sample with very high net demand.

[Figure 2 about here.]

While these results provide a simple examination of the data, the event study approach has several limitations in our context. First, hourly plant-level data on electricity production are not available prior to 2015. Consequently, the event study framework cannot be used to explore heterogeneity in how different plants respond to the nuclear phase-out beginning in 2011. This heterogeneity is especially important because the amount of local air pollution emitted per MWh of production can vary significantly across plants burning the same type of fuel. In addition, the monetary damage from local air pollution emissions is also tied directly to the number of people exposed to this pollution; the same level of pollution emissions from two different plants can have very different damages based on the number of people living near each of these plants.

Second, the event study framework relies on the assumption that changes in power plant operations around March 15, 2011 are caused by the nuclear reactor closures rather than changes in other factors that determine production behavior. To ensure that this assumption holds, we examine the impact of the phase-out in a fairly narrow window around the initial 2011 shutdowns. Focusing on this narrow window allows us to argue that firms could only respond to the nuclear shutdowns in the very short-run by adjusting output. However, subsequent nuclear plant shutdowns occurred incrementally and were pre-announced. As such, firms may have been able to take actions in anticipation of these later closures.

Finally, as discussed in Section 3, other important economic factors also changed over our 2010-2017 sample period independent from the nuclear phase-out in 2011. For example, coal and natural gas plants had similar marginal costs in 2011. However, coal prices decreased precipitously from 2011-2015 while natural gas prices increased over this period. Coal plants were thus increasingly more likely to produce in place of natural gas plants from 2011-2015 even absent any changes in nuclear power production. In addition, many older coal and gas plants were retired between 2010 and 2017, and a number of new fossil fuel-fired plants came online during this period as well. Summarizing, it is unlikely that market outcomes before versus after March 2011 were driven solely by the phase-out, especially as we look further in time after the 2011 shutdown decision.

# 5 Machine Learning Approach

## 5.1 Methodology

We use a machine learning approach to more credibly estimate the market and environmental impacts of the series of nuclear plant closures that occurred between 2011 and 2017. This machine learning approach has two advantages over the event study framework discussed in the previous section. First, hourly plant-level data on electricity production are not available prior to 2015; for this reason, we estimate the event study regressions using data on hourly aggregate electricity production by fuel type. As we noted earlier, plant-level heterogeneity is particularly important for estimating the damages from local air pollution exposure: different plants burning the same type of fuel may have very different emissions factors and number of people living nearby. The machine learning algorithm allows us to use hourly plant-level data from 2015-2017 to estimate plant-level heterogeneity in response to the nuclear phase-out over our entire 2010-2017 sample period.

Second, as discussed earlier, a variety of economic factors relevant for electricity production decisions changed over time independently from the nuclear phase-out. The event study framework affords us only limited ability to control for these factors. In contrast, the machine learning approach allows us to estimate the impact of the nuclear phase-out on plant-level economic and environmental outcomes controlling for a wide range of observed market factors.

Importantly, the goal of our machine learning framework is to best predict market outcomes for different values of the input variables. This differs from traditional econometric methods in two ways. First, we do not seek to identify the causal effect of one variable on another. Second, though we are able to provide bounds on our estimates, it has proven impossible to derive standard errors on the predictions from machine learning models absent randomization of treatment and control groups (Wager and Athey, 2018). Summarizing, our machine learning algorithm gives us substantially more accurate predictions of market outcomes than the event study approach at the cost of being unable to conduct traditional statistical inference on these predictions.

## 5.2 Data

We train our machine learning algorithm to predict power plant operations using a data set of roughly 4.5 million observations. The outcome of interest is the hourly quantity of electricity produced by each "dispatchable" plant in our sample. We subtract "non-dispatchable" renewable output from electricity demand because renewables have near-zero marginal cost and thus produce whenever nature permits (ex: the sun is out or the wind is blowing). Hourly data on plant-level electricity production are available for all EU member states since 2015 from ENTSOE.[4] We incorporate electricity imports and exports at each border interconnection between Germany and its neighboring countries into our framework by treating each border interconnection point as if it is a power plant. For example, consider the hourly net electricity imports from France to Germany. If France exports 100MWh of electricity to Germany, this border point would be "producing" 100MWh. Conversely, if France imports 100MWh of electricity from Germany, this border point would be "producing" -100MWh.

The dependent variables considered in our machine learning framework are the production levels from each power plant and border points in our sample. In all cases, we normalize the relevant dependent variable by dividing output by the maximum production capacity of each power plant or the maximum transfer capacity of the border point

[4]More specifically, the data are available for plants with capacity greater than 100 MW. This covers 100% of production from nuclear plants, 95% from lignite plants, 85% from hard coal plants, 50% from gas plants and 45% from oil plants. We treat the operating behavior of these plants as being representative of the remaining plants with capacity less than 100MW.

as applicable. Our algorithm focuses on dependent variables that are bounded between 0 and 1; we rescale the flows from border points from their original scale of -1/1 to 0/1 when applying the algorithm. We refer to this rescaled output as the operating rate for each power plant.

The independent variables include electricity demand, local weather, each plant's marginal cost, the availability of other power plants, and a wide range of power plant characteristics such as fuel type, efficiency, technology, and location. We estimate a predictive model that takes these independent variables as inputs and outputs a predicted operating rate for each power plant in each hour. Importantly, we have data on these independent variables from 2010-2017. This allows us to predict hourly, plant-level electricity production from 2010-2017 using our model despite only observing hourly plant-level production from 2015 onward.

We also build a predictive model for wholesale electricity prices. However, there is no cross-sectional variation in these prices; the hourly wholesale electricity price is the same throughout Germany. In this case, the independent variables for the time-series model of electricity prices include electricity demand, national average weather, and the marginal cost associated with the marginal unit (i.e.: the unit with the largest marginal cost that produces a positive quantity in that hour-of-sample).

## 5.3 Empirical Methods

We predict outcomes using a Random Forest regression algorithm (Breiman, 2001). In particular, we use the Quantile Regression Forest algorithm (Meinshausen, 2006). Random forests are especially well-suited for our empirical context for several reasons. First, each plant's production is based on a potentially complex combination of factors such as the marginal costs and availability of other plants, electricity demand at different locations, and transmission constraints. Consequently, the relationship between plant-level production and the independent variables listed above is likely to be highly non-linear and include multiple interactions. Random forest methods are well-suited to use variation in the data in order to find these interactions rather than pre-specifying how independent and dependent variables relate using polynomials or splines as in a more standard regression framework.[5]

In addition, the Random Forest algorithm ensures that the support of possible outcome predictions is bounded by the support of the outcome values in the training data-set. This prevents nonsensical predictions such as plants producing negative amounts of electricity or producing greater than their capacity. Finally, using the Quantile Regression Forests algorithm allows us to produce predictions for the full conditional distribution of the outcomes rather than just their expected value. This property both allows us to better understand the uncertainty in our analysis and to make corrections that ensure that our predicted outcomes meet certain physical constraints (e.g. that electricity supply equals electricity demand). More details can be found in Appendix B.

We use the Quantile Random Forest model to construct two data series. First, we predict hourly plant-level electricity production at each dispatchable plant (i.e. each fossil plant or border point) using the observed values of the independent variables over 2010-2017. This provides us with electricity production levels at each plant in the "factual" scenario with the nuclear phase-out. We note that the machine learning model is necessary for estimating plant-level production even in the factual scenario because there is no hourly plant-level production data prior to 2015.

Second, we use the model to estimate hourly production for the same set of dispatchable plants in the counterfactual scenario where there was no nuclear phase-out. Put another way, we predict plant-level production assuming that the nuclear reactors that were shut down in 2011, 2015, and 2017 would have remained operational until 2017. To do this, we first calculate the amount of electricity these nuclear plants would have produced in each hour-of-sample if they had remained online.[6] We subtract this counterfactual nuclear output from net electricity demand, thus reducing the production needed from the remaining dispatchable plants. In our primary specifications, we hold all of the other independent variables that do not depend on net demand fixed at their observed

[5]In their application for predicting housing values, Mullainathan and Spiess (2017) report that the Random Forest method results in the most accurate predictions, as measured by out-of-sample $R^2$, among the various methods evaluated (e.g., OLS, Regression Tree, LASSO, and Ensemble).

[6]We assume that the nuclear plants that were shut down would have operated at 80% of their capacity on average. We choose this relatively conservative 80% operating rate because the nuclear plants that were shut down tended to be older; newer nuclear plants often achieve operating rates of 90-95%. We adjust this counterfactual nuclear output based on observed fluctuations in monthly total nuclear production from 2012 to 2014 because there were no nuclear shutdowns during this period. This adjustment primarily reflects the fact that nuclear plants tend to go on maintenance during the summer months when demand is lowest.

values. A natural concern is that the phase-out led to changes in other independent variables such as the available production capacity from other plants. We discuss the various sensitivity analyses we implemented to address this concern in Section 5.5. Further details on the implementation of our machine learning algorithm can be found in Appendix Section B.

Finally, we calculate different market and environmental outcomes using the predicted hourly electricity production from each plant with versus without the nuclear phase-out. Though our exposition has focused on hourly plant-level production, we utilize a similar approach to assess the impact of the phase-out on wholesale electricity prices.

## 5.4 Model Validation

This subsection presents figures and tables comparing observed outcomes with the outcomes predicted by our machine learning algorithm.

[Figure 3 about here.]

Figure 3 reports daily average observed versus predicted wholesale prices in 2017 USD per MWh, as well as the difference between the two (i.e., the prediction error). It is evident that the machine learning model delivers very accurate predictions; the difference between observed versus predicted prices is nearly zero throughout the entire period. Nevertheless, the adjusted $R^2$ from the regression of observed average daily price on the predicted average daily price is 0.98.

Figure 4(a) compares observed hourly plant-level operating rates (i.e., percentage of capacity utilized) with the predictions from the machine learning model. Specifically the predicted electricity production (scaled on the y-axis) is plotted against the observed production (x-axis) so that observations on the 45 degree line indicate perfect prediction accuracy. Each pixel in the figure represents the predicted vs. actual operating rate in increments of 2% and darker areas correspond to a higher number of plant-hour (or plant-year) observations.

We check the out-of-sample cross-validated performance to avoid overfitting and give a fair assessment of how the model may perform when used to make predictions about our counterfactual no-phase-out scenario. The cross-validated out-of-sample $R^2$ is 0.61 and the mean squared error (MSE) is 0.061.[7]

However, even this small level of prediction error understates the relevant prediction accuracy of the machine learning model. Specifically, we will primarily use the predictions from our model to compare outcomes with versus without the phase-out at the plant-month and plant-year levels. We therefore also evaluate the predictive performance of the model at these levels of aggregation. Specifically, Figure 4(b) plots predicted versus observed annual average operating rates. As the figure shows, the performance is substantially improved, with most of the observations clustered close to the 45 degree line, and the areas of systematic error largely disappear. The cross-validated out-of-sample $R^2$ rises to 0.93 and the mean-squared error falls to 0.006.[8]

[Figure 4 about here.]

As an alternative metric to the cross-validated out-of-sample $R^2$ and MSE, we also evaluate accuracy of the machine learning predictions by testing whether variation in predicted hourly plant-level production is correlated with observed variation in ambient air pollution at nearby monitors. To this end, we use data from air pollution monitors in Germany spanning the entire 2010-2017 analysis period; we match each power plant to its three closest air pollution monitors.[9] Specifically, we construct a daily plant-level measures of air pollution concentrations as the inverse distance-weighted average of the readings from these three monitors. We then estimate panel regressions of daily average ambient pollution concentrations on daily total plant-level production. We include plant fixed effects, year fixed effects, and month-of-year fixed effects in order to control for seasonality in air pollution and electricity production, as well as, plant-specific emission intensities.

[Table 2 about here.]

Table 2 reports the results of this analysis. Each row reports the coefficient estimates,

[7] By comparison, a simple OLS regression with the same independent variables only achieves an out-of-sample $R^2$ of 0.37 and a mean-squared error of 0.091.

[8] A simple OLS regression with linear covariates is still clearly inferior with an out-of-sample $R^2$ of 0.63 and an MSE of 0.025.

[9] The average distance between power plants and the nearest air pollution monitors is 6.5 km, with a range of 0.25km to 31 km.

along with standard errors clustered by plant, from separate regressions for 5 air pollutants: $PM_{10}$, $PM_{2.5}$, $SO_2$, $CO$, and $NO_2$. For ease of interpretation, both the dependent variables and the plant-level production variables are standardized to have a mean of 0 and a standard deviation of 1. The columns correspond to different estimation samples. Column (1) is from models where the dependent variable is standardized observed daily plant-level production from 2015-2017. In column (2) the dependent variable is standardized predicted daily production over the same 2015-2017 period, and in column (3), the dependent variable is standardized predicted production over the full 2010-2017 period. The key comparison to assess the validity of the Random Forest prediction algorithm is between columns (1) and (2). The estimates in column (1) confirm that increases in observed daily production correspond to increases in pollution concentration levels for all pollutants except for $SO_2$. For example, a standard deviation increase in average daily production leads to a 0.13 standard deviation increase in average daily concentration of $PM_{10}$, or roughly a 1% increase in daily concentrations.

Column (2) replicates the analysis using the daily plant-level production predicted by the Random Forest model as the dependent variable. The resulting coefficient estimates are similar in magnitude and exhibit the same patterns and statistical significance as the specification in column (1) using observed production. Finally, column (3) reports estimates assessing the impact of predicted prediction on pollution levels over the entire 2010-17 sample period. These estimates documented in column (3) are similar to those in columns (1) and (2). Taken together, the analysis in Table 2 provides evidence that additional electricity production leads to higher concentrations of ambient air pollutants. More importantly, this table also provides evidence that our predicted plant-level production estimates are accurate even for the pre-2015 sample period where we do not have data on plant-level production.

## 5.5 Sensitivity Analyses

This subsection describes and motivates three sensitivity analyses we conduct that pertain to how we construct the counterfactual no-nuclear-phase-out scenario. The full results of these analyzes are described in Section 6.4. In our primary specifications, we assume that the effect of the phase-out flows solely through reductions in electricity production from the nuclear plants that were shut down. Put another way, we assume that the other observed factors in our model do not change as a result of the phase-out. This assumption makes sense for many of our predictors such as plant characteristics, temperature, and seasonality of demand. However, other factors may have changed as a consequence of the phase-out. For example, the phase-out may have led to an increase in retail electricity prices, which in turn might reduce aggregate electricity demand. As another example, over longer timescales, the phase-out may have accelerated investment in new replacement production capacity. We address these concerns by demonstrating how sensitive our results are to varying factors that may have changed as a result of the phase-out.

Our first sensitivity analysis focuses on how the nuclear phase-out impacts investment in fossil fuel-fired capacity. Prior studies have demonstrated that, if the phase-out had not occurred, the amount of fossil fuel-fired capacity necessary to ensure that demand is met even during peak hours in Germany would have been 4 GW lower by 2020 (Traber and Kemfert, 2012) and 8 GW lower by 2030 (Knopf et al., 2014). This reduction in capacity could be due either to fewer new fossil plants being built or older existing plants closing early. To capture this, we calculate how this 4GW (8GW) reduction by 2020 (2030) would impact fossil-fuel-fired capacity during our 2010-2017 sample period.[10] We then re-run the analysis for the counterfactual no-phase-out scenario removing the relevant fossil capacity from the system in each year.

Another sensitivity analysis accounts for the fact that the incentives to invest in renewable production may not have been as strong in the absence of the nuclear phase-out. To do this, we re-run our machine learning prediction model for the no-phase-out scenario assuming that renewable production would have been 30 TWh lower by 2017. We chose 30 TWh based on changes made to Germany's renewable energy targets in response to the phase-out decision. Specifically, in 2010, Germany planned on producing at least 30% of its electricity from renewables by 2020. However, this target was increased to 35% following the 2011 phase-out decision (Jacobs, 2012). The difference between these two targets requires a change in renewable production of roughly 30 TWh between 2010

[10]Getting to 4GW (8GW) less fossil capacity by 2020 (2030) can be achieved by assuming that fossil capacity falls by 0.4 GW per year from 2011 to 2030. For our 2010-2017 analysis period, we achieve this with the following modifications: Irsching opens in 2012 instead of 2011, Weisweiler (Blocks C & D) closes in 2011 instead of 2012, Boxberg opens in 2013 instead of 2012, KW Walsum opens in 2014 instead of 2013, GKM Mannheim (Blocks 3 & 4) closes in 2012 instead of 2015, Westfalen (Block E) opens in 2015 instead of 2014, Westfalen (Block C) closes in 2015 instead of 2016, Moorburg (Blocks A & B) opens in 2018 instead of 2015 and KW Voerde (Blocks A & B) closes in 2016 instead of 2017.

and 2017. Reducing renewable production by 30 TWh amounts to an 8% increase in net electricity demand by 2017 for the counterfactual case where the phase-out had not gone ahead. We argue that this 8% increase in net electricity demand is a relatively large response.[11] Consequently, this sensitivity analysis shows how our results change when considering an upper bound on the extent to which investment in renewables was driven by the phase-out.

Finally, one might be concerned that the phase-out increases wholesale electricity prices which in turn might decrease consumer demand. We argue that our second sensitivity analysis should assuage this concern. Specifically, as discussed above, an 8% increase in net demand due to the phase-out is an extremely large response; it is unlikely that consumer demand shifts by more than 8% due to the phase-out. In fact, it is plausible that changes in wholesale prices do not impact customer demand much at all. This is because the commercial and residential customers that make up around half of Germany's total demand are highly price-inelastic; wholesale electricity prices are only roughly a quarter of their overall retail price, with the remainder being network charges, renewable subsidy fees and taxes (BNetzA, 2018). Though larger industrial customers may be more price-elastic, changes in their electricity demand are extremely unlikely to result in changes in aggregate net demand that exceed 8%. Consequently, our third sensitivity analysis focused on changes in net demand due to changes in renewables also helps to address concerns that the phase-out impacted consumer demand through changes in wholesale prices.

# 6 Social Costs and Benefits of the Nuclear Phase-Out

This section presents the primary results on the full range of impacts of the nuclear phase-out. Specifically, we compare the market and environmental outcomes with versus without the nuclear phase-out using the predictions from our machine learning model.

[11] For example, previous work on the phase-out assumed that investments in renewables did not accelerate due to the nuclear plant closures (Traber and Kemfert (2012); Knopf et al. (2014)). Furthermore, the increases in wholesale electricity prices resulting from the phase-out were unlikely to impact the profitability of investment in renewable capacity. This is because all renewable capacity in Germany is remunerated through feed-in-tariffs that provide a guaranteed above-market price for the electricity produced.

## 6.1 Private Costs and Benefits of the Phase-Out

This subsection examines how the nuclear phase-out affected wholesale electricity prices, electricity production, revenues and operating costs. All currency units are converted from nominal Euros to constant 2017 USD.

[Figure 5 about here.]

Figure 5 presents our estimates of the impacts of the nuclear phase-out on electricity production and wholesale prices. First among these is Figure 5(a), which reports the monthly average difference in predicted production and net imports (in TWh) with minus without the phase-out policy. We report monthly average differences in fossil-fired electricity production (grey diamonds), net imports (red circles), and nuclear production (purple squares). The start of the nuclear phase-out in March 2011 is marked by the vertical black dashed line; the "with" minus "without" phase-out differences are zero before this point. By construction, we find a stark reduction in total nuclear production of 3-5 TWh per month. The cyclicality of this impact is due primarily to the fact that nuclear reactors typically schedule their maintenance and refuelling outages in the summer months.

The phase-out also caused a large increase in fossil-fuel-fired electricity production of 2-3 TWh per month and a smaller increase in net imports of electricity. Importantly, these increases are calculated taking into account the rise in renewable production over our sample period. Another notable result in Figure 5(a) is that the stark increase in fossil production starting in March 2011 persists over our entire sample period.

Figure 5(b) is constructed similarly and reports the impact of the nuclear phase-out on wholesale electricity prices in 2017 USD per MWh. The estimates clearly show that the phase-out resulted in an increase in wholesale prices, ranging from roughly from 0.5 to 8 dollars per MWh. Another key result in Figure 5(b) is that the increase in wholesale prices persists through the end of 2017, as was similarly noted for fossil fuel electricity production. Finally, the figure also shows that the phase-out may have exacerbated episodic increases in prices, such as the large price spike in January 2017 due to an unusual cold spell in Europe (European Commission, 2017).

[Table 3 about here.]

Column (1) in Table 3 complements the information in Figure 5 by reporting annual average predicted wholesale electricity price and electricity production in the scenario with the phase-out. Column (2) reports these predicted outcomes for the scenario without the phase-out. Column (3) reports the difference between the first two columns and Column (4) provides this estimated effect as a percentage by dividing column (3) by column (1). The estimates reveal that the phase-out caused inflation-adjusted wholesale electricity prices to increase by $1.80 per MWh on average, a 3.9% increase relative to the prices that would have prevailed if the phase-out had not occurred. Consistent with Figure 5(a), nuclear production fell by an average of 53.2 TWh per year during the phase-out period, corresponding to a 38% decline. The next rows decompose the previously documented increase in fossil production by source. The largest increases, both in absolute and percentage terms, are from hard coal and gas-fired production. Specifically, annual average production from hard coal increased by 28.5 TWh (32%) while gas-fired production increased by 8.3 TWh (26%). Finally, the phase-out caused net imports to increase by 10.2 TWh (37%) per year on average. In sum, the 2011 phase-out lead to large changes to Germany's electricity generation mix.

[Table 4 about here.]

Table 4 examines the impact of the nuclear phase-out on financial outcomes for power plants, once again organized by plant fuel type. We report predicted annual average revenues, operating costs, and operating profits. Revenues are calculated as the product of plant-level production and wholesale electricity prices; we thus ignore any additional revenues plants may receive, such as capacity payments, ancillary services payments, subsidies etc. Operating costs are the product of each plant's hourly production with its hourly marginal cost. Finally, operating profits are simply operating revenues minus operating costs. For net imports, we quantify revenues and costs as the net import of electricity multiplied by the wholesale price in the relevant neighboring country.[12] All of the entries in Table 4 are in billions of dollars (2017 USD) per year.

[12] Our analysis implicitly assumes that the phase-out caused no change to the electricity prices of neighboring countries. Fully modeling electricity markets for each of these interconnected countries would entail a prohibitive amount of additional data collection. This additional modeling would also be unlikely to dramatically alter the overall findings given the dominant role of domestic production in meeting Germany's electricity demand. Finally, since prices in interconnected electricity markets likely increased due to the phase-out, our net import cost estimates are likely to be a lower bound.

The nuclear phase-out had a large effect on the revenues and operating profits of the firms that owned the nuclear plants that were shut down. Specifically, annual average revenues across all nuclear plants declined by \$2.2 billion per year. Annual average operating profits earned by nuclear plants fell by \$1.6 billion (a 35% reduction). This decline is striking, especially given that it accounts for the increased revenues earned by the nuclear plants that remained open and were thus able to benefit from the increase in wholesale electricity prices.

The revenues previously earned by the shut-down nuclear plants were primarily redistributed to fossil plants, most notably hard coal and natural gas plants. This shift occurred at a less than one-for-one ratio since nuclear plants have a much lower operating costs per MWh than fossil plants. Despite this, annual average operating profits at fossil plants increased by roughly \$0.4 and \$0.3 billion due to the phase-out at lignite and coal plants respectively. This corresponds to sizable increases of 17% and 64%.

The redistribution of profits amongst electricity producers has interesting implications for the political economy surrounding the phase-out policy. In particular, the four large firms that owned nuclear plants in Germany clearly opposed the policy both privately and publicly. However, there are two important factors that may have tempered their opposition. First, these firms would have been allowed to operate their nuclear plants into the 2030s only if they paid a nuclear fuel tax. This nuclear fuel tax would have taxed away a large portion of the inframarginal rents that these nuclear plants earn. Second, the four firms that owned nuclear plants also had large fossil plant portfolios both in Germany and across Europe. As we have seen, these fossil plants earned larger profits due to the nuclear phase-out, which likely cushioned the reduction in profits earned by the four firms as a result of the nuclear closures.

## 6.2 External Costs and Benefits of the Nuclear Phase-Out

This subsection presents two separate analyses of environmental costs associated with the phase-out-induced increase in fossil-fuel-fired production documented in the previous subsection. Specifically, burning fossil fuels emits both global pollutants such as carbon dioxide that contribute to climate change and local pollutants that adversely impact the health of exposed populations.

### 6.2.1 Estimating External Damages Using Reported Emissions Rates

First, we estimate the change in carbon emissions due to the phase-out. To proceed, we calculate the change in the amount of fuel burned by each plant associated with the phase-out impact on each plant's hourly production and using each plant's thermal efficiency (i.e.: how well the plant translates input heat energy to output electricity). We then use the carbon intensity of different fuels documented in industry reports to convert changes in fuel burned to changes in plant-level $CO_2$ emissions.[13]

We also estimate the change in local pollution emissions due to phase-out-induced changes in plant production levels. Similar to the approach for $CO_2$ emissions, we translate changes in fuel use into changes in emissions using plant-level emissions rates for each local pollutant from the EU Large Combustion Plant Directive (LCPD). The LCPD database provides annual plant-level data on fuel inputs and emissions of sulfur dioxide ($SO_2$), nitrogen oxides ($NO_x$) and particulate matter (PM). The LCPD data covers the vast majority of large fossil plants in Germany.[14] We assign the small number of plants not in the LCPD database an emissions factor based on the average emissions factor of plants with the same fuel type.

We next monetize the damages caused by $CO_2$ and local air pollution emissions. For $CO_2$, we monetize damages assuming a social cost of carbon of \$50/t$CO_2$. To assess the health damages from increases in local air pollution, we rely on two studies that estimate the health impacts of local pollution in Europe (EEA, 2014; Jones et al., 2018). In particular, Jones et al. (2018) provide estimates of the annual health damages from the local air pollution emitted by roughly four hundred of the largest coal-fired power plants in Europe. We use these data to convert our predicted increases in plant-level kilotons of $SO_2$, $NO_x$ and PM emissions into monetized health damages.[15]

[13] The carbon intensities we use are 93.6 tCO2/TJ for hard coal, 55.9 tCO2/TJ for gas and 74.0 tCO2/TJ for oil. We consider three different intensities for lignite depending on the mining region that the plant sources its coal from. These are 113.3 tCO2/TJ (Rhineland), 111.2 tCO2/TJ (Lusatian) and 102.8 tCO2/TJ (Central).

[14] Specifically, the data covers 99% of lignite capacity, 98% of hard coal capacity, 90% of gas capacity and 91% of oil capacity.

[15] Specifically, we assume that increased emissions at a given fossil-fuel-fired plant in Germany would have the same health damages as if they were emitted at the nearest location for which we have health damages estimates. The mean distance between each of the power plants in our data set and closest of the 400 locations with damage estimates is 29km. The median is 14km. Jones et al. (2018) provides estimates for roughly 10% of the plants in our data-set, noting that these plants are among the 400 largest coal plants in Europe.

[Table 5 about here.]

Table 5 presents the results of this analysis. Specifically, this table reports the fuel-specific annual emissions for $CO_2$ (in Megatonnes, Mt) as well as the emissions of three local pollutants: $SO_2$, $NO_x$, and PM (in kilotonnes, kt). Lignite and hard coal are by far the two largest polluters, contributing more than 90% of emissions. Lignite and hard coal also contribute the most in terms of monetary damages from emissions, which are reported in billions of USD per year.

In aggregate, the phase-out led to an increase in $CO_2$ emissions of 36.3 Mt per year. This corresponds to a 13% increase relative to the scenario without the nuclear phase-out. This increase in $CO_2$ emissions was primarily attributable to an increase in emissions from hard coal plants of 25.8 Mt, with lignite and gas making up the remainder. Valuing carbon emissions at a social cost of carbon of \$50/tCO2, the phase-out results in estimated climate change damages of \$1.8 billion.

The phase-out also led to a roughly 12% increase in the total emissions of each the three local air pollutants we consider ($SO_2$, $NO_x$, and PM). Again, this increase is due primarily to increased emissions from hard coal plants. The bottom panel of Table 5 reports annual average mortality damages summed across all three local air pollutants. From 2010-2017, local pollution emissions from fossil plants were responsible for around \$65 billion in mortality costs each year. \$8.7 billion of this annual mortality cost can be attributed to the nuclear phase-out, representing a 15% increase in damages relative to the scenario without the nuclear phase-out.[16] Put another way, the phase-out resulted in more than 1,100 additional deaths per year from increased concentrations of $SO_2$, $NO_x$, and PM. The increase in production from hard coal plants is again the key driver here, making up roughly 80% of the increase in mortality impacts.

#### 6.2.2 Estimating External Damages Using Ambient Air Pollution Monitors

As an alternative to calculating damages using fuel inputs and reported emissions, we also compute damages using the estimated relationship between plant-level production and recorded air pollution at nearby monitoring stations. We have already shown that

[16]We use a Value of Statistical Life of \$7.9 million for Germany taken from Viscusi and Masterman (2017).

increased fossil fuel-fired production results in higher concentrations of $PM_{10}$, $PM_{2.5}$, CO and $NO_2$ in Table 2. In this subsubsection, we estimate a daily monitor-level regression of ambient air pollution on daily plant-level predicted production for the sample period 2010-2017:

$$P^{PO}_{i,d,m,y} = \alpha + \beta_s Y^{PO}_{i,s,d,m,y} + \mu_i + \delta_m + \delta_y + u_{i,d,m,y} \tag{2}$$

where $P^{PO}_{i,d,m,y}$ is recorded air pollution concentrations near plant $i$, on day $d$, in month $m$, and year $y$.[17] The "PO" superscript denotes that pollution is measured in the factual scenario with the nuclear phase-out.

$Y^{PO}_{i,s,d,m,y}$ represents daily electricity production at plant $i$ , powered by fuel $s$, in the factual phase-out scenario. The coefficient of interest, $\beta$, is estimated separately for each fuel type $s$, to account for the differing pollution intensities of lignite, hard coal, natural gas and oil plants. We include plant fixed effects ($\mu_i$) to control for plant-specific factors that are correlated with local air pollution conditional on production, such as the presence of pollution abatement technologies. We also include month-of-year fixed effects ($\delta_m$) and year-of-sample fixed effects ($\delta_y$) to control for trends and seasonality in air pollution and electricity production.

The regression coefficients $\beta_s$ capture how a one MWh increase in production from fuel type $s$ impacts local air pollution concentration levels. To estimate the change in local air pollution attributable to the phase-out (through its effect on electricity production at plant $i$), we multiply each coefficient estimate $\hat{\beta}$ by the phase-out driven change in production at plant $i$, burning fuel $s$. Formally, we calculate $\Delta POLL_{i,d,m,y}$ $= \hat{\beta}_s \times (Y^{PO}_{i,s,d,m,y} - Y^{NPO}_{i,s,d,m,y})$ which is the estimated increase in pollution levels due to phase-out-induced increases in fossil-fuel-fired production at plant $i$ (using fuel $s$).

We calculate the increase in premature mortality due to this increase in air pollution concentrations using dose-response estimates from the ESCAPE project (Lancet 2014).[18] Specifically, the ESCAPE project reports that mortality rate when $PM_{2.5}$ exposure is X + 5 micrograms per cubic meter divided by the mortality rate when $PM_{2.5}$ exposure is X micrograms per cubic meter is 1.07. The corresponding hazard ratio for a 10 micrograms

[17] As before, we calculate pollution at each plant as the inverse distance-weighted average of the measurements at the three pollution monitors closest to this plant.

[18] The European Study of Cohorts for Air Pollution Effects (ESCAPE) is one of the few studies on the health impact of air pollution exposure in Europe. It is based on 22 European cohort studies with a total study population of more than 350,000 participants.

per cubic meter increase in $NO_2$ is 1.01.[19]

Based on these hazard ratios, we can calculate the increase in mortality caused by the additional air pollution due to the phase-out using the following formula:

$$VSL \times POP \times MR \times \left(1 - \frac{1}{\exp(\rho_j \Delta POLL_j)}\right) \quad (3)$$

for j=$PM_{2.5}$ or $NO_2$. The value of statistical life (VSL) used to monetize the premature mortality due to phase-out-induced increases in air pollution is \$7.9 million as in the previous subsubsection (Viscusi and Masterman, 2017). POP and MR are the population and mortality rate in the exposure group. The parameter $\rho_j$ corresponds to the hazard ratios described above and $\Delta POLL_j$ is the change in ambient air pollution caused by the phase-out for air pollutant $j$. Finally, we assume that only the population residing within 20 km of the fossil power plants is exposed to the additional air pollution due to the phase-out (approximately 7.5% of the total population of Germany). This population measure is calculated using satellite-based data from the Socioeconomic Data and Applications Center (SEDAC) at NASA.

[Table 6 about here.]

The estimates of monetized mortality damages are reported in Table 6. Specifically, we present the annual average impact of the phase-out on pollution concentrations, premature mortality and the monetized damages from this premature mortality. A few key results emerge. First, there is again clear evidence that the phase-out resulted in significant increases in local pollution that in turn led to large and costly increases in premature mortality. Second, the changes in $PM_{2.5}$ and $PM_{10}$ concentration levels due to the phase-out were responsible for much larger health impacts than the change in $NO_2$ air pollution (about 10 times more). Finally, the primary drivers of excess mortality are the hard coal and lignite power plants. The estimates in column (3) suggest that the additional production from burning hard coal due to the phase-out led to \$3 billion in annual mortality damages. Phase-out-induced increases in production from lignite led to \$1.2 billion in annual mortality damages. Overall, this analysis points to the phase-out causing annual premature mortality damages of 4.3 billion USD per year.

[19] In order to calculate the hazard ratio for $PM_{2.5}$, we convert $PM_{10}$ to $PM_{2.5}$ by making the simple assumption that $PM_{10} = 0.5PM_{2.5}$. There are no dose-response functions for CO and $SO_2$ in the ESCAPE project.

Taken together, the results in Tables 5 and 6 paint a remarkably consistent picture of the monetized mortality damages attributable to the nuclear phase-out. That being said, our preferred estimate is the 8.7 billion USD per year in damages calculated based on reported emissions (Table 5). This is because the analysis using reported emissions considers a more complete set of pollutants and implicitly draws on a more sophisticated analysis of pollution transport and exposure. The results presented in Table 6 based on our estimated relationships between pollution concentrations and electricity production (Table 6) serves as a valuable complementary validation exercise, especially given it was derived using an entirely distinct approach. Lastly, we want to emphasize that the air pollution costs of the phase-out are economically sizable, amounting to a roughly 10-15% increase in damages from premature mortality due to air pollution emissions from Germany's power sector.

#### 6.2.3 Estimating Risks from Nuclear Accidents and Waste Storage

Nuclear power plants emit virtually no global or local air pollution. However, nuclear energy does come with catastrophic accident risk and requires storing the waste that results from nuclear production, which has important costs as well. For instance, JECR (2019) estimates that the cost of the Fukushima accident over the next forty years is between 35-80 trillion yen ($330-750 billion). Most of this cost will not be incurred by the firm that owned the Fukishima nuclear power plant; the costs of the Fukishima accident are largely borne by Japanese society as a whole.

More generally, estimates from the literature suggest that the external costs of nuclear power due to waste storage and accident risk fall between €1-4 per MWh (Dhaeseleer, 2013). This wide range is due to differing estimates of accident probabilities and severity, as well as varying assumptions on discount rates. If we value the external costs of nuclear power at $3 per MWh, the expected benefits from the nuclear phase-out are $0.2 billion per year. Even if we value the external costs of nuclear at $30 per MWh, a value far higher than the magnitudes typically found in the literature, the expected benefits from the nuclear phase-out are still relatively moderate at $2 billion per year. This is markedly smaller than our estimate of the external costs associated with the nuclear phase-out.

## 6.3 Total Costs and Benefits of the Nuclear Phase-Out

This subsection bring the analysis together by summarizing the full range of private and external costs and benefits of the nuclear phase-out. The private costs of the phase-out consist of changes in the operating costs of the power plants in our sample as well as any net costs from changes to imports and exports. The external costs of the phase-out include the monetized climate change damages from carbon emissions, the damages from mortality, and morbidity caused by the air pollution attributable to the change in electricity production mix. Finally, the benefits of the phase-out consist of reductions in the costs associated with nuclear waste and accident risks.

[Table 7 about here.]

Table 7 reports the aggregate cost and benefits of the phase-out. The phase-out resulted in replacing low cost nuclear production with higher cost sources such as fossil fuels and net imports; this increases average operating costs in Germany by $1.6 billion per year. Whilst not trivial, these private costs are small relative to the external costs associated with the phase-out. Specifically, burning fossil fuels to produce electricity rather than using nuclear plants emits global pollutants such as $CO_2$ as well as local pollutants such as $PM_{2.5}$, $SO_2$ and $NO_2$.

The climate damages from phase-out-induced increases in $CO_2$ emissions alone amount to $1.8 billion per year. However, the largest impact of the phase-out by far has been the external costs from local air pollution emissions. Specifically, increased exposure to local air pollution results in an additional 1,100 excess deaths due to poorer air quality. We estimate the monetized mortality impacts to be $8.7 billion per year when using reported emissions, with a further $0.2 billion per year in morbidity costs. The average reduction in the external costs from nuclear waste and accident risks are small by comparison at $0.2 billion per year. Overall we estimate the annual ongoing costs of the nuclear phase-out as approximately $12.2 billion per year.

## 6.4 How does the Phase-Out Impact Investment?

Keppler (2012) argues that extending the lifetime of the nuclear reactors in Germany would have required investments of roughly €500 million per reactor, or €8.5 billion in total (roughly \$10 billion). These investment costs are avoided due to the nuclear phase-out. However, Knopf et al. (2014) argues that the phase-out led to 8GW of additional fossil-fuel-fired capacity being required by 2030. If we assume coal-fired capacity has capital costs of \$3500/kW while gas-fired capacity has capital costs of \$1000/kW, the total additional investment costs in fossil-fuel-fired capacity as a result of the nuclear phase-out range from \$8-\$28 billion. Subtracting the avoided investment costs in nuclear from this range, the net investment costs of the phase-out are between -\$2 billion to \$18 billion. That being said, our central estimate of the *annual* net increase in intensive margin costs as a consequence of the nuclear phase-out is roughly \$12 billion. Therefore, no reasonable comparison of the investment costs with versus without the phase-out can overturn the conclusion that the phase-out fails a simple benefit-cost test by a large margin.

However, one could argue that the nuclear phase-out accelerated investment in renewable sources. Increased investment in renewables drives down the production costs and air pollution damages associated with shifting away from nuclear. To explore this argument, we estimated a scenario where the phase-out incentivized a steady increase in investment in renewables. We set the level of this annual investment in renewables such that Germany produces its target of an additional 30 TWh per year of renewable generation by 2020. In this "renewables" scenario, fossil-fuel-fired power plants are tasked with producing roughly 5 TWh less electricity each year. Consequently, the increase in annual average private operating costs due to the phase-out is \$1.4 billion instead of \$1.6 billion in the baseline analysis. The phase-out-induced increase in climate damages costs is \$1.3 billion (versus \$1.8 billion in the baseline analysis) while the phase-out-induced increase in air pollution damages is \$7.6 billion (versus \$8.7 billion in the baseline analysis). Combined, allowing for a sizable increase in renewable production as a consequence of the phase-out decreases the total net costs of the phase-out by only \$1.8 billion per year (from \$12.2 billion in our baseline to \$10.4 billion with increased renewables).

## 6.5 Robustness Checks

Two externally estimated parameters play a key role in our estimates: (a) the Value of Statistical Life (VSL) used to monetize the additional morality due to phase-out-induced local air pollution, and (b) the external costs of nuclear waste and accident risks. Our central estimate of the cost of the nuclear phase-out is based on a VSL of $7.9 million from Viscusi and Masterman (2017). We believe this Germany-specific VSL to be most reliable/up-to-date. Nevertheless, the Organization for Economic Cooperation and Development (OECD) estimates that the Germany-specific VSL is approximately $3 million, which is one of the lowest VSL estimates for Germany we've seen in the literature.[20] Even using this extremely low VSL of $3 million, we find that the air pollution costs of the phase-out are $3.3 billion (versus $8.7 billion in the baseline analysis). This significantly more conservative assumption on VSL reduces the total cost of the phase-out to $6.4 billion per year (versus $12.2 billion in the baseline analysis).

Similarly, we can value the external costs associated with nuclear waste and nuclear accident risk at $30 per MWh. This is roughly 10 times larger than the external costs of nuclear power estimated in previous studies (Dhaeseleer, 2013). This extremely conservative (i.e.: high) estimate increases the benefits of the phase-out from $0.2 billion per year to $2 billion per year. However, replacing both the VSL and external costs of nuclear power with extremely conservative estimates is still not sufficient to overturn the conclusion that the nuclear phase-out resoundingly fails a simple benefit-cost test.

# 7 Conclusions and Policy Discussion

Following the Fukashima disaster in 2011, German authorities made the unprecedented decision to: (1) immediately shut down almost half of the country's nuclear power plants and (2) shut down all of the remaining nuclear power plants by 2022. We quantify the full extent of the economic and environmental costs of this decision. Our analysis indicates that the phase-out of nuclear power comes with an annual cost to Germany of roughly $12 billion per year. Over 70% of this cost is due to the 1,100 excess deaths per year resulting from the local air pollution emitted by the coal-fired power plants operating in place of the shutdown nuclear plants. Our estimated costs of the nuclear phase-out far exceed the right-tail estimates of the benefits from the phase-out due to reductions in nuclear accident risk and waste disposal costs.

Moreover, we find that the phase-out resulted in substantial increases in the electricity prices paid by consumers. One might thus expect German citizens to strongly oppose the

[20] Viscusi and Masterman (2017) discusses the shortcomings of the OECD estimates of VSL.

phase-out policy both because of the air pollution costs and increases in electricity prices imposed upon them as a result of the policy. On the contrary, the nuclear phase-out still has widespread support, with more than 81% in favor of it in a 2015 survey (Goebel et al., 2015). This support cannot be chalked up to a lack of concern regarding climate change. Indeed, German citizens widely support the transition to renewables as part of the Energiewende program even though the costs of this transition were €26 billion in 2017 alone. German citizens are also highly aware of the costs associated with the transition to renewables, with charges for renewable subsidies now making up about a quarter of the electricity price paid by residential households.

This raises the question: what drives the global shift away from nuclear power despite the substantial economic and environmental costs associated with this policy? We discuss two potential mechanisms. First, the nuclear phase-out may be the result of rational decision-making by risk averse agents. Specifically, we compare the social costs of the phase-out against the *expected* benefits of this policy. However, nuclear accident risk imposes uncertainty on citizens and the costs associated with nuclear waste disposal are also arguably relatively uncertain. It is thus possible that a sufficiently risk-averse policymaker could phase-out nuclear to avoid the tail risks associated with nuclear accidents and waste disposal, even though the air pollution costs associated with the phase-out are higher in expectation.

To get a sense of the level of risk aversion required to justify the phase-out, we calculate the probability of a major nuclear accident that would result in the expected benefits from the phase-out being equal to the costs. For this back-of-the-envelope calculate, assume that, absent the phase-out, nuclear plants would have been shut down in the same order but by 2032 instead of 2022. This gives 2032-2011 = 21 years over which the phase-out would reduce nuclear production. Our estimated cost of the phase-out is $12 billion per year; this implies a cumulative cost of the phase-out of $250 billion over 2011-2032. The upper bound estimates of the cost of the Fukushima accident are roughly $750 billion (JECR, 2019). Assume for simplicity that there can either be no accidents or there can be one Fukushima magnitude accident during this 21 year window. The probability of this Fukushima-scale accident occurring would have to be $0.33 \approx \frac{\$250 \text{ billion}}{\$750 \text{ billion}}$ in order for the expected benefits of the phase-out to be equal to the costs of the phase-out. This is far greater than even the most conservative estimates of the probability of an accident of this magnitude occurring in Germany.[21] This in turn suggests that policymakers would have to exhibit an extremely high level of risk aversion in order to rationalize the phase-out based on risk aversion alone.

That being said, citizens may also be anti-nuclear because the risks associated with nuclear power are more salient than the air pollution costs associated with fossil-fuel-fired production. Specifically, the literature on the harmful effects of air pollution is becoming more definitive by the day. However, there is still relatively limited public understanding of the scale of the adverse health consequences of local air pollution exposure. This might be because it is difficult to attribute any single death entirely to pollution exposure from a single power plant. Instead, pollution concentration levels are the result of a wide range of different emitters and air pollution slightly but persistently increases the mortality risk of large exposed populations. Similarly, the costs of climate change will primarily be born by future generations, and linking a future climate event to the carbon emissions from a power plant smokestack is even less straightforward. In contrast, a nuclear accident is a highly visible, yet low probability, event that can be clearly linked back to the offending nuclear reactor. This may lead both policymakers and the public to over-estimate the ex-ante probability that nuclear accidents will occur as well as costs of these accidents (Slovic, Fischhoff and Lichtenstein (1979); Slovic (2010)).

Regardless of the underlying causes, it is clear that the German citizenry cares deeply about climate change yet is distinctly anti-nuclear. Policymakers around the world thus face a difficult trade-off. On the one hand, many climate change experts have argued that nuclear power is a necessary part of the shift away from carbon-intensive fossil fuels. Moreover, many voters are willing to incur substantial costs to reduce the risk of climate change. However, many of these same voters are also unwilling to support nuclear power due to fears surrounding nuclear accidents and nuclear waste disposal. Facing this political pressure, countries around the world are shifting away from nuclear production despite the substantial increases in operating costs and air pollution costs associated with this policy. This highlights that it is essential for policymakers and academics to convey the relative costs of climate change and air pollution versus nuclear accident risk and waste disposal to the voting public.

[21]For instance, Wheatley, Sovacool and Sornette (2017) estimates that there is a 50% chance that a Fukushima event (or larger) occurs every 60-150 years across the entire global fleet of nuclear reactors. Germany had less than 4% of the world's nuclear reactors in 2011. Moreover, nuclear reactors in Germany almost certainly come with less accident risk than other parts of the world.

# References


**Beelen, Rob, Ole Raaschou-Nielsen, Massimo Stafoggia, Zorana Jovanovic Andersen, Gudrun Weinmayr, Barbara Hoffmann, Kathrin Wolf, Evangelia Samoli, Paul Fischer, Mark Nieuwenhuijsen, et al.** 2014. "Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project." *The Lancet*, 383(9919): 785–795.

**BMWi.** 2018. "Sixth Energy Transition Monitoring Report: The Energy of the Future." Federal Ministry of Economic Affairs and Energy (BMWi) Report.

**BNetzA.** 2018. "Monitoring Reports."

**Breiman, Leo.** 2001. "Random Forests." *Machine Learning*, 45(1): 5–32.

**Burlig, Fiona, Christopher Knittel, David Rapson, Mar Reguant, and Catherine Wolfram.** 2017. "Machine Learning from Schools about Energy Efficiency." National Bureau of Economic Research Working Paper 23908.

**Callaway, Duncan S, Meredith Fowlie, and Gavin McCormick.** 2018. "Location, location, location: The variable value of renewable energy and demand-side efficiency resources." *Journal of the Association of Environmental and Resource Economists*, 5(1): 39–75.

**Cicala, Steve.** 2017. "Imperfect Markets versus Imperfect Regulation in U.S. Electricity Generation." National Bureau of Economic Research Working Paper 23053.

**Davis, Lucas, and Catherine Hausman.** 2016. "Market Impacts of a Nuclear Power Plant Closure." *American Economic Journal: Applied Economics*, 8(2): 92–122.

**Deschenes, Olivier, Michael Greenstone, and Joseph S. Shapiro.** 2017. "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program." *American Economic Review*, 107(10): 2958–89.

**Dhaeseleer, William.** 2013. "Synthesis on the Economics of Nuclear Energy." DG Energy Report.

**EEA.** 2014. "Costs of air pollution from European industrial facilities 20082012." European Environment Agency EEA Technical Report 20/2014.

**Egerer, Jonas.** 2016. "Open Source Electricity Model for Germany (ELMOD-DE)." DIW Berlin, German Institute for Economic Research Data Documentation 83.

**European Commission.** 2017. "Quarterly Report on European Electricity Markets."

**Goebel, Jan, Christian Krekel, Tim Tiefenbach, and Nicolas R Ziebarth.** 2015. "How natural disasters can affect environmental concerns, risk aversion, and even politics: evidence from Fukushima and three European countries." *Journal of Population Economics*, 28(4): 1137–1180.

**Grossi, Luigi, Sven Heim, and Michael Waterson.** 2017. "The impact of the German response to the Fukushima earthquake." *Energy Economics*, 66: 450 – 465.

**Holland, Stephen P, Erin T Mansur, Nicholas Muller, and Andrew J Yates.** 2018. "Decompositions and Policy Consequences of an Extraordinary Decline in Air Pollution from Electricity Generation." National Bureau of Economic Research Working Paper 25339.

**Jacobs, David.** 2012. "The German Energiewende History, Targets, Policies and Challenges." *Renewable Energy Law and Policy Review*, 3(4): 223–233.

**Jaramillo, Paulina, and Nicholas Muller.** 2016. "Air pollution emissions and damages from energy production in the U.S.: 20022011." *Energy Policy*, 90(C): 202–211.

**JECR.** 2019. "Follow up Report of Public Financial Burden of the Fukushima Nuclear Accident." Japan Center for Economic Research Report.

**Jones, Dave, Charles Moore, Will Richard, Rosa Gierens, Lauri Myllvirta, Sala Primc, Greg McNevin, Kathrin Gutmann, Anton Lazarus, Christian Schaible, and Joanna Flisowka.** 2018. "Last Gasp: the coal companies making Europe sick." Europe Beyond Coal.

**Keppler, Jan Horst.** 2012. "The economic costs of the nuclear phase-out in Germany."

**Knopf, Brigitte, Michael Pahle, Hendrik Kondziella, Fabian Joas, Ottmar Edenhofer, and Thomas Bruckner.** 2014. "Germany's Nuclear Phase-out: Sensitivities and Impacts on Electricity Prices and CO2 Emissions." *Economics of Energy & Environmental Policy*, 0(Number 1).

**Markandya, Anil, and Paul Wilkinson.** 2007. "Electricity generation and health." *The lancet*, 370(9591): 979–990.

**Meinshausen, Nicolai.** 2006. "Quantile Regression Forests." *J. Mach. Learn. Res.*, 7: 983–999.

**Mullainathan, Sendhil, and Jann Spiess.** 2017. "Machine learning: an applied econometric approach." *Journal of Economic Perspectives*, 31(2): 87–106.

**Neidell, Matthew J, Shinsuke Uchida, and Marcella Veronesi.** 2019. "Be Cautious with the Precautionary Principle: Evidence from Fukushima Daiichi Nuclear Accident." National Bureau of Economic Research.

**NRC and NAS.** 2010. "Hidden Costs of Energy: Unpriced Consequences of Energy Production and Use." National Research Council (US). Committee on Health, Environmental, and Other External Costs and Benefits of Energy Production and Consumption. National Academies Press.

**Open Power System Data.** 2018. "Data Package Conventional power plants."

**Severnini, Edson.** 2017. "Impacts of nuclear plant shutdown on coal-fired power generation and infant health in the Tennessee Valley in the 1980s." *Nature Energy*, 2(4): 17051.

**Slovic, Paul.** 2010. "The psychology of risks." *Sade e Sociedade*, 19(4): 737–747.

**Slovic, Paul, and Elke U Weber.** 2002. "Perception of risk posed by extreme events." *Regulation of Toxic Substances and Hazardous Waste (2nd edition)(Applegate, Gabba, Laitos, and Sachs, Editors), Foundation Press, Forthcoming.*

**Slovic, Paul, Baruch Fischhoff, and Sarah. Lichtenstein.** 1979. "Rating the risks." *Environment*, 21(3): 14–39.

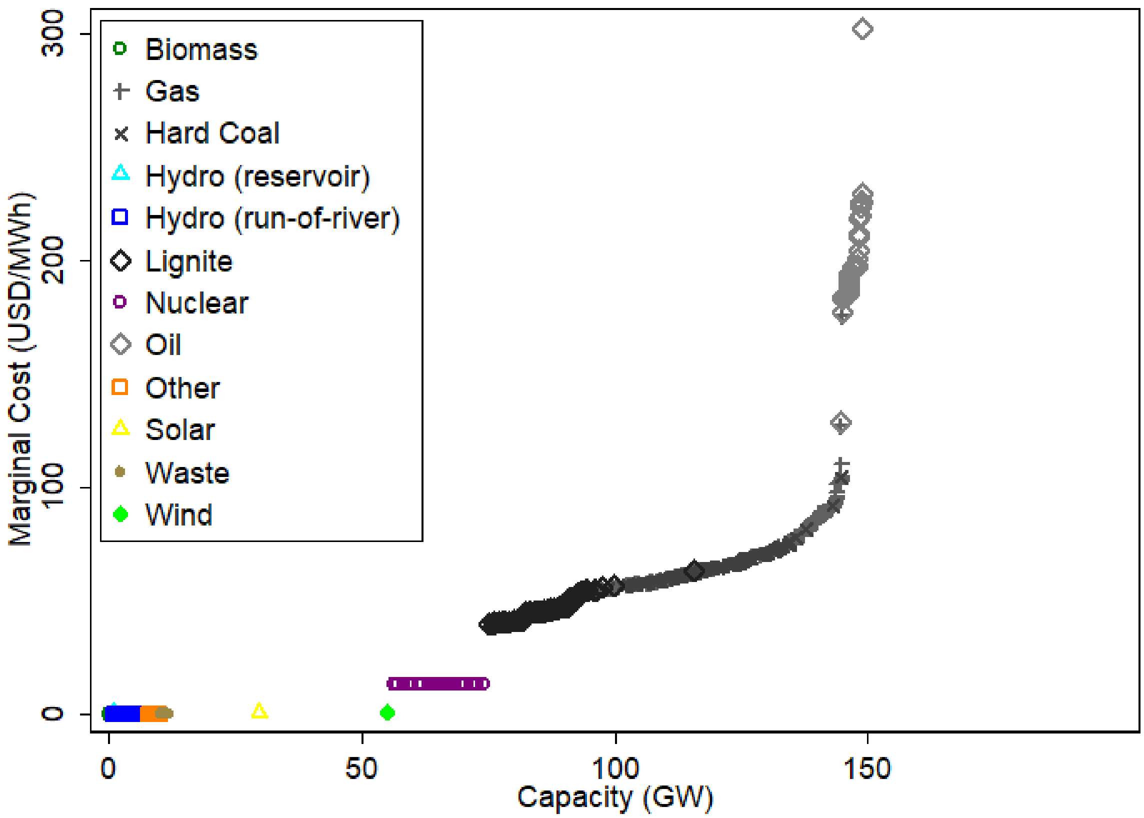
**Traber, Thure, and Claudia Kemfert.** 2012. "German Nuclear Phase-out Policy: Effects on European Electricity Wholesale Prices, Emission Prices, Conventional Power Plant Investments and Eletricity Trade." DIW Berlin, German Institute for Economic Research Discussion Papers of DIW Berlin 1219.

**Viscusi, W. Kip, and Clayton J. Masterman.** 2017. "Income Elasticities and Global Values of a Statistical Life." *Journal of Benefit-Cost Analysis*, 8(2): 226250.

**Wager, Stefan, and Susan Athey.** 2018. "Estimation and inference of heterogeneous treatment effects using random forests." *Journal of the American Statistical Association*, 113(523): 1228–1242.

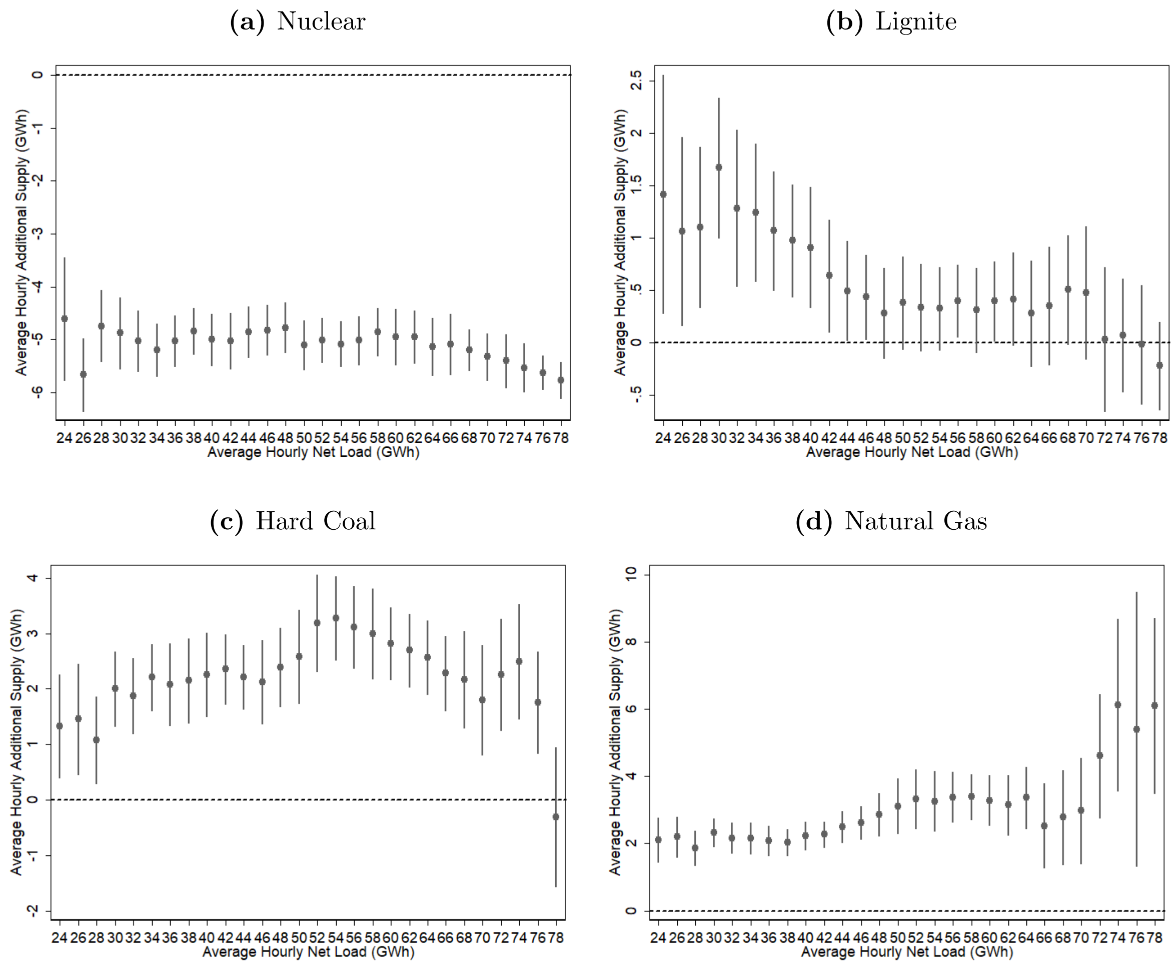
**Wheatley, Spencer, Benjamin Sovacool, and Didier Sornette.** 2017. "Of Disasters and Dragon Kings: A Statistical Analysis of Nuclear Power Incidents and Accidents." *Risk Analysis*, 37(1): 99–115.

**Figure 1:** Marginal Cost Curve in 2011



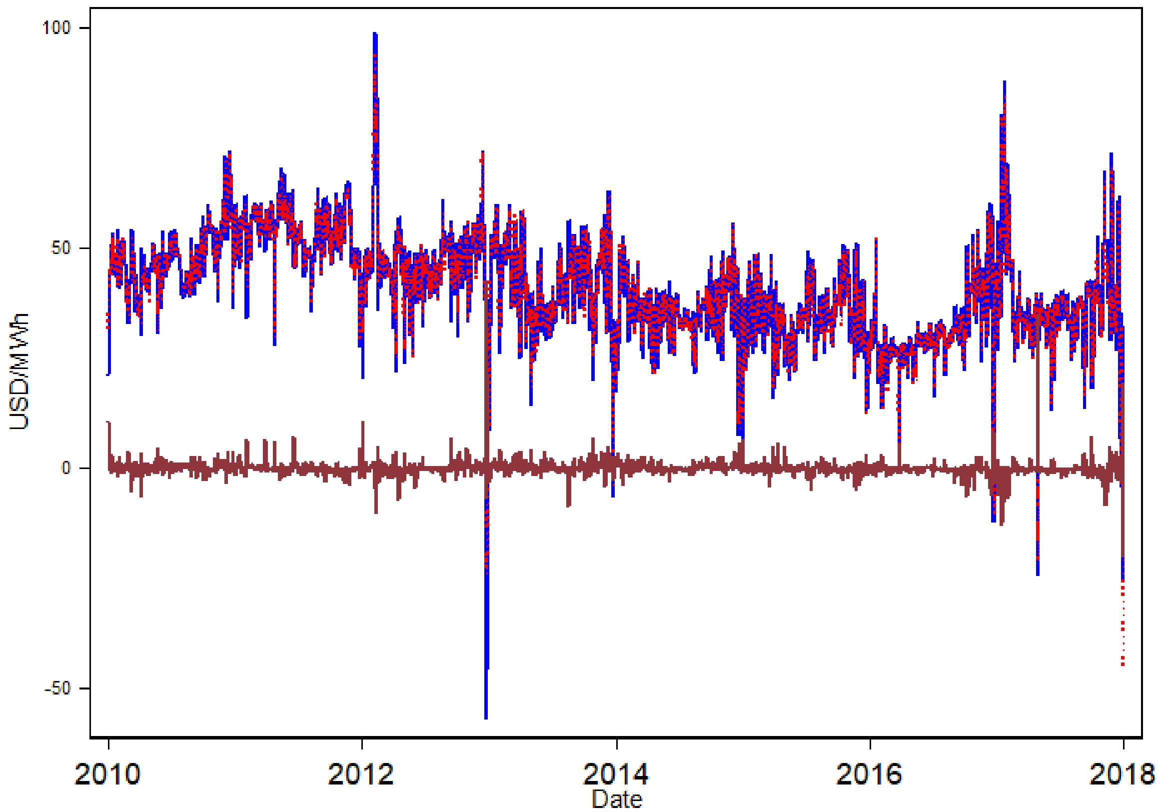

**Notes:** This figure plots estimated marginal costs for power plants in Germany in 2011. Specifically, plants are ordered in terms of marginal cost to create an aggregate supply curve. For a given marginal cost $c$ (plotted on the y-axis), the x-axis provides the sum of the production capacity (in GW) over all plants with marginal cost less than or equal to $c$. Marginal costs are in 2017 U.S. dollars. For coal, gas and oil plants, marginal costs are calculated as the sum of fuel costs and an assumed variable operating and maintenance cost that differs by fuel type. Fuel costs are converted to dollars per MWh using the plant's thermal efficiency: how well the plant converts units of input heat to units of electricity output. For this figure, we consider the fuel costs on February 1st, 2011. Nuclear plants are assigned a marginal cost of $10 per MWh as in Egerer (2016). Hydro, wind and solar have zero marginal costs. For simplicity, the small amount of remaining sources are also assigned a marginal cost of zero (i.e. biomass, waste and other). For ease of presentation, this figure does not show how electricity imports and exports factor into the aggregate supply curve; importantly, we account for imports and exports in our analysis.

**Figure 2:** Event-Study Estimates: Effect of the 2011 Nuclear Closures on Production



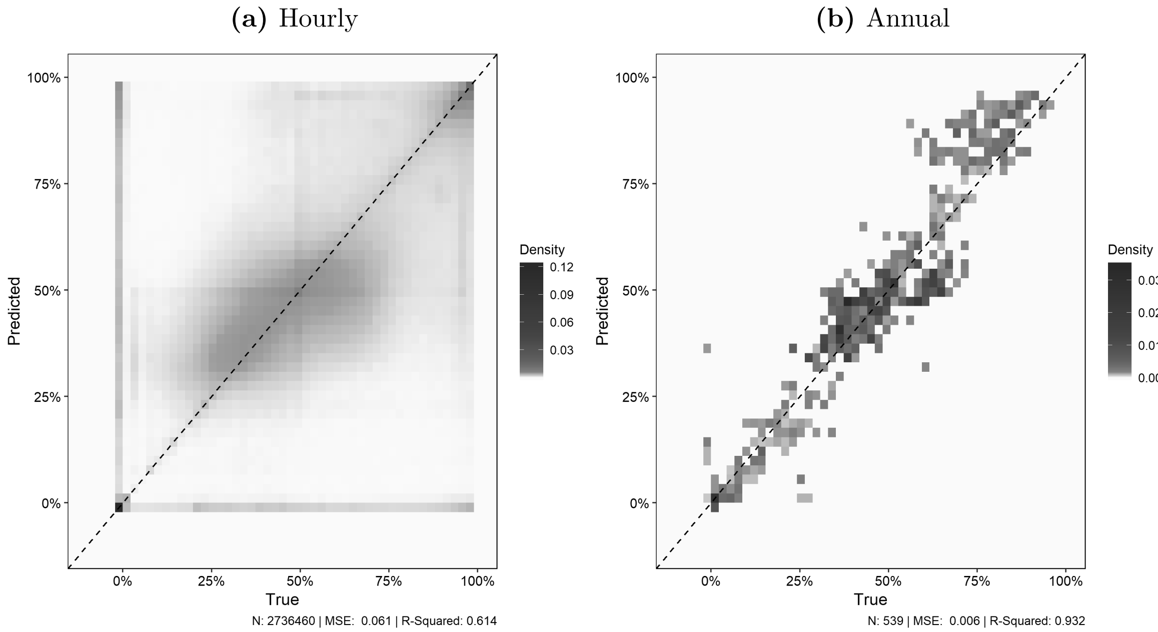

**Notes:** This figure plots the results from an event study analysis of the effects of the nuclear phase-out in Germany in 2011. The estimates correspond to changes in electricity production by source after relative to before March 15, 2011. Panel (a) presents the estimates for nuclear production, separately for each of 28 equally sized bins of net demand (i.e.: electricity demand minus production from renewables). Panels (b)-(d) present the corresponding estimates for production from lignite, hard coal, and natural gas respectively. The panels also include the point-wise 95% confidence interval around each of the estimated effects; the standard errors used to construct these confidence intervals are clustered by week-of-sample.

**Figure 3:** Machine Learning Model Performance: Wholesale Electricity Prices




**Notes:** This figure illustrates the accuracy of the aggregate predictions from the machine learning model presented in Section 5. The model predicts the wholesale electricity price in each hour-of-sample from 2010-2017. The figure depicts the observed wholesale electricity price (solid blue line), our model prediction (dotted red line) and the difference between these two (solid maroon line along the x-axis). Whilst the model predicts prices at the hourly level, the data in this figure have been averaged to a daily resolution for ease of presentation.

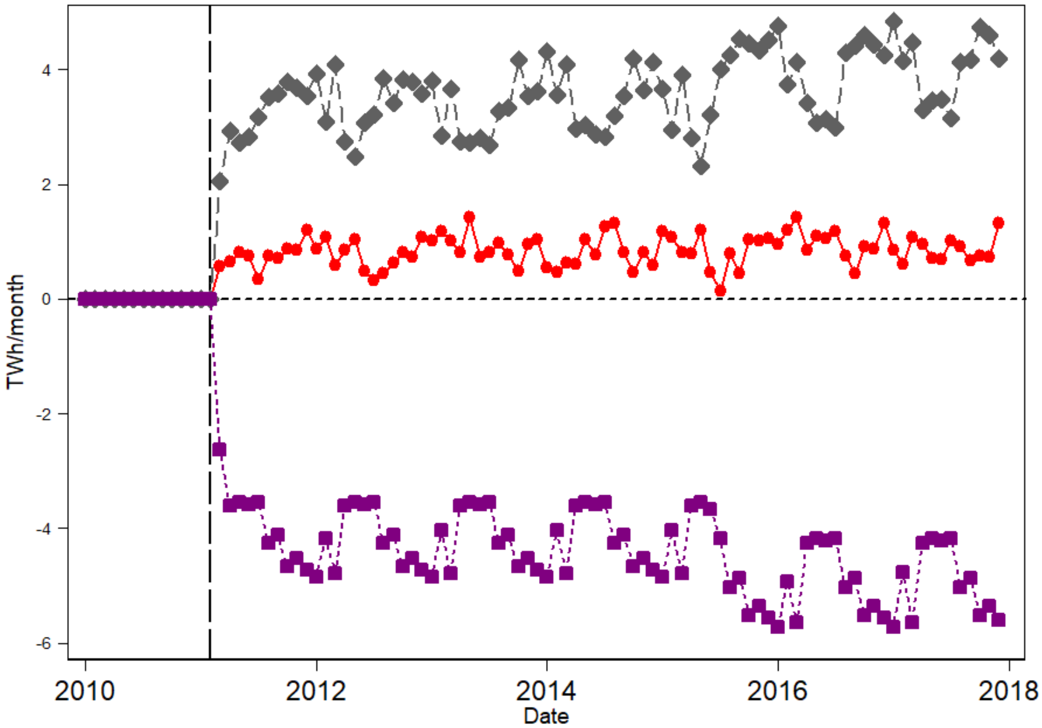
**Figure 4:** Machine Learning Model Performance: Plant-Level Electricity Production



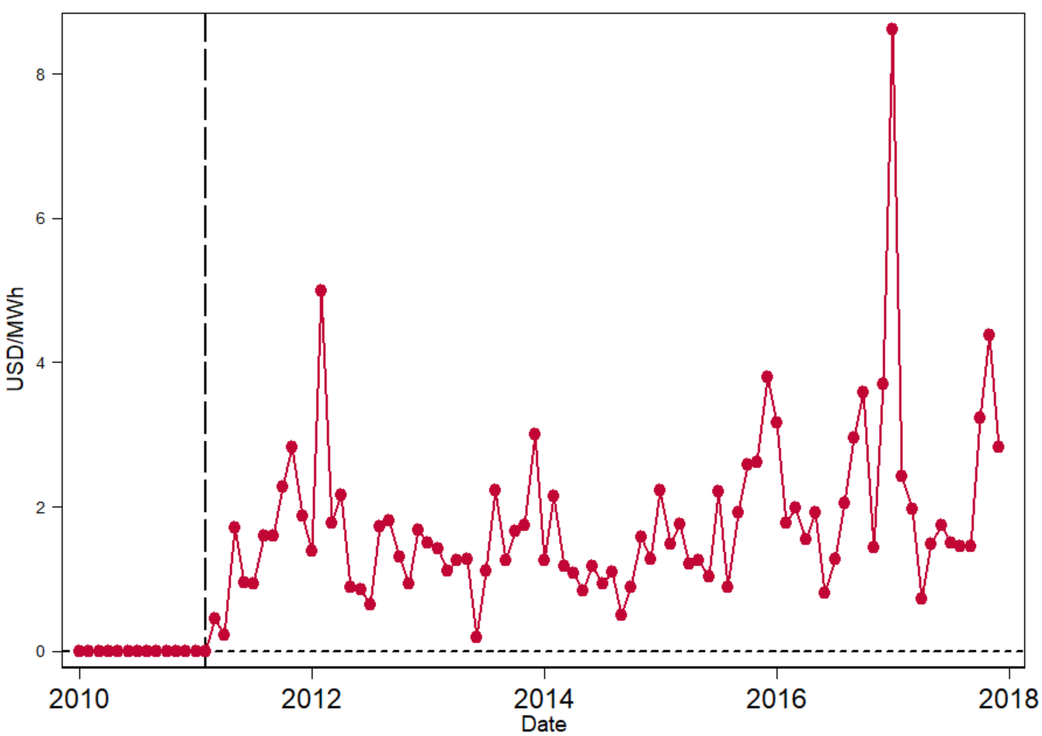

**Notes:** This figure illustrates the accuracy of the plant-level predictions from the machine learning model presented in Section 5. The model predicts the operating rate of each power plant in each hour, where a value of 0% means that a plant is offline and a value of 100% means that the plant is running at maximum capacity. Values on the 45 degree line indicate perfect accuracy, and we summarize this both visually and by computing measures of Mean Squared Error and R-Squared. We compute these metrics using out-of-sample cross-validation to avoid overfitting and give a fair assessment of how the model may perform when used to make predictions about our counterfactual scenario. We use five-fold cross-validation: we divide the 2015-2017 training dataset into five randomly generated subsets, or"folds". We then estimate our predictive model using four fifths of the data and check how this model performs in predicting outcomes for the remaining one fifth of the data. We repeat this for each of the five folds and then average the resulting measures of performance. Panel (a) shows prediction accuracy at an hourly timescale. Panel (b) shows prediction accuracy after taking annual averages of our hourly predictions. Darker areas indicate higher numbers of plant-hour (or plant-year) observations. Each pixel represents the predicted vs. actual operating rate in increments of 2%.

**Figure 5:** Impact of the Phase-Out on Electricity Production and Prices

**(a)** Electricity Production




**(b)** Wholesale Electricity Prices




**Notes:** This figure plots the monthly difference between the predictions from our machine learning algorithm with the phase-out minus without the phase-out. The start of the phase-out in March 2011 is marked by the vertical black dashed line. Panel (a) reports the estimates for all fossil-fuel-fired electricity production (grey diamonds), net imports (red circles), and nuclear production (purple squares). Panel (b) presents the change in wholesale electricity prices.

**Table 1** Summary Statistics

| | 2010 | 2017 |
|---|---|---|
| **All Electricity Sector** | | |
| Total Capacity (GW) | 172.4 | 217.6 |
| Electricity Production (TWh) | 551.4 | 591.2 |
| Wholesale Electricity Price (USD/MWh) | 70.68 | 41.81 |
| Net Electricity Imports (TWh/Year) | -3.5 | -33.5 |
| **By Source** | | |
| **Nuclear Plants** | | |
| Number of Plants | 16 | 7 |
| Average Capacity (MW/Plant) | 1,196.9 | 1,359.4 |
| Annual Electricity Production (TWh) | 134.7 | 70.5 |
| **Hard Coal Plants** | | |
| Number of Plants | 109 | 87 |
| Average Capacity (MW/Plant) | 236.5 | 288.0 |
| Annual Electricity Production (TWh) | 93.9 | 83.5 |
| Marginal Costs (USD/MWh) | 64.9 | 41.8 |
| **Lignite Plants** | | |
| Number of Plants | 74 | 61 |
| Average Capacity (MW/Plant) | 274.0 | 344.1 |
| Annual Electricity Production (TWh) | 130.9 | 137.9 |
| Marginal Costs (USD/MWh) | 54.2 | 28.9 |
| **Gas Plants** | | |
| Number of Plants | 242 | 268 |
| Average Capacity (MW/Plant) | 96.9 | 98.6 |
| Annual Electricity Production (TWh) | 53.6 | 72.3 |
| Marginal Costs (USD/MWh) | 77.6 | 41.8 |
| **Oil Plants** | | |
| Number of Plants | 53 | 50 |
| Average Capacity (MW/Plant) | 79.0 | 80.6 |
| Annual Electricity Production (TWh) | 1.9 | 3.8 |
| Marginal Costs (USD/MWh) | 197.5 | 125.8 |
| **Renewables (Hydro, Solar, and Wind)** | | |
| Total Capacity (GW) | 52.1 | 112.5 |
| Annual Electricity Production (TWh) | 60.6 | 157.1 |

**Notes:** This table reports summary statistics for Germany's electricity generation sector in 2010 versus 2017. Electricity prices and marginal costs are in constant 2017 USD. While not reported in Table 1, we nuclear plants have a marginal operating cost of approximately $10/MWh (in 2017 USD) based on prior research on Germany's power sector (Egerer, 2016).

**Table 2** Estimated Relationship Between Ambient Air Pollution and Electricity Production

| | (1) Actual, 2015-17 | (2) Predicted, 2015-17 | (3) Predicted, 2010-17 |
|---|---|---|---|
| $PM_{10}$ | 0.132*** | 0.140*** | 0.186*** |
| | (0.013) | (0.009) | (0.014) |
| $PM_{2.5}$ | 0.152*** | 0.138*** | 0.169*** |
| | (0.009) | (0.008) | (0.015) |
| $SO_2$ | -0.005 | -0.008 | 0.021 |
| | (0.013) | (0.013) | (0.012) |
| CO | 0.172*** | 0.152*** | 0.197*** |
| | (0.019) | (0.019) | (0.024) |
| $NO_2$ | 0.167*** | 0.184*** | 0.251*** |
| | (0.014) | (0.010) | (0.017) |
| Year and Month FEs | X | X | X |
| Plant FEs | X | X | X |

**Notes:** This table reports coefficient estimates from a panel regression of daily air pollution concentrations on daily plant-level electricity production. Both the dependent variable and the independent variable are standardized to have a mean of 0 and a standard deviation of 1. The regressions include plant fixed effects, month-of-year fixed effects, and year-of-sample fixed effects and the standard errors are clustered at the plant level. ***,**, and * denote statistical significance at the 0.1%, 1%, and 5% levels respectively.

**Table 3** Estimated Impact of the Nuclear Phase-Out on Wholesale Electricity Prices, Electricity Production by Source, and Net Imports

| | **Average with Phase-Out (1)** | **Average w/out Phase-Out (2)** | **Change (3)** | **Change (%) (4)** |
|---|---|---|---|---|
| **Production (TWh/Year)** | 574.4 | 574.2 | 0.2 | 0.0% |
| Nuclear | 86.2 | 139.4 | -53.2 | -38.2% |
| Lignite | 160.4 | 154.3 | 6.1 | 3.9% |
| Hard Coal | 118.3 | 89.8 | 28.5 | 31.7% |
| Gas | 39.8 | 31.6 | 8.3 | 26.2% |
| Oil | 11.1 | 10.7 | 0.4 | 3.7% |
| Net Electricity Imports | -17.2 | -27.4 | 10.2 | 37.1% |
| Renewables and Others | 175.8 | 175.8 | 0.0 | 0.0% |
| **Wholesale Prices ($/MWh)** | 47.3 | 45.5 | 1.8 | 3.9% |

**Notes:** This table reports annual average electricity production by type and wholesale electricity prices with versus without the nuclear phase-out, as estimated using our machine learning algorithm. These annual averages are calculated using data spanning from immediately after the phase-out in March 2011 to the end of 2017. In our baseline specification, the "renewables and others" production category experiences no change by construction. This is relaxed in one of our sensitivity analyses (see Section 6.4).

**Table 4** Estimated Impact of the Nuclear Phase-Out on Revenues, Operating Costs, and Operating Profits

| | Average with Phase-Out (1) | Average w/out Phase-Out (2) | Change (3) | Change (%) (4) |
|---|---|---|---|---|
| **Revenues** (Billion $/Year) | 19.3 | 18.6 | 0.7 | 3.9% |
| Nuclear | 4.1 | 6.4 | -2.2 | -35.0% |
| Lignite | 7.6 | 7.1 | 0.6 | 8.0% |
| Hard Coal | 5.8 | 4.3 | 1.5 | 34.4% |
| Gas | 2.0 | 1.5 | 0.5 | 30.9% |
| Oil | 0.5 | 0.5 | 0.0 | 7.0% |
| Net Electricity Imports | -0.7 | -1.1 | 0.4 | 36.6% |
| Renewables and Others | – | – | 0.0 | 0.0% |
| **Costs** (Billion $/Year) | 14.2 | 12.6 | 1.6 | 12.7% |
| Nuclear | 1.0 | 1.7 | -0.6 | -37.9% |
| Lignite | 5.1 | 4.9 | 0.2 | 4.0% |
| Hard Coal | 4.9 | 3.7 | 1.1 | 30.1% |
| Gas | 2.3 | 1.9 | 0.4 | 23.2% |
| Oil | 1.9 | 1.8 | 0.0 | 2.5% |
| Net Electricity Imports | -0.9 | -1.4 | 0.4 | 31.4% |
| Renewables and Others | – | – | 0.0 | 0.0% |
| **Profits** (Billion $/Year) | 5.2 | 6.0 | -0.9 | -14.4% |
| Nuclear | 3.1 | 4.7 | -1.6 | -33.9% |
| Lignite | 2.6 | 2.2 | 0.4 | 17.0% |
| Hard Coal | 0.9 | 0.5 | 0.3 | 63.6% |
| Gas | -0.3 | -0.4 | 0.0 | 8.1% |
| Oil | -1.3 | -1.3 | 0.0 | -0.8% |
| Net Electricity Imports | 0.2 | 0.2 | 0.0 | -5.9% |
| Renewables and Others | – | – | 0.0 | 0.0% |

**Notes:** This table reports average annual operating revenues, costs and profits with versus without the nuclear phase-out, as estimated using our machine learning algorithm. All values are annualized averages based on predictions from after the nuclear shutdowns in March 2011 to the end of 2017. Operating revenues are the product of each plant's hourly production with the hourly wholesale electricity price. We ignore any additional revenues plants may receive, such as capacity payments, ancillary services payments, subsidies etc. Operating costs are the product of each plant's hourly production with its hourly marginal cost. Operating profits are operating revenues minus operating costs. Other sources such as renewables are excluded from this table as we avoid making explicit assumptions about their marginal costs or their revenues (e.g., additional non-market subsidies).

**Table 5** Estimated Impact of the Nuclear Phase-Out on $CO_2$ Emissions and Local Air Pollution Mortality Damages

| | **Average with Phase-Out** | **Average w/out Phase-Out** | **Change** | **Change (%)** |
|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** |
| $CO_2$ Emissions (Mt/Year) | 316.6 | 280.3 | 36.3 | 13.0% |
| Lignite | 182.8 | 175.9 | 6.9 | 3.9% |
| Hard Coal | 108.0 | 82.2 | 25.8 | 31.4% |
| Gas | 17.0 | 13.6 | 3.3 | 24.5% |
| Oil | 8.9 | 8.6 | 0.3 | 3.6% |
| $SO_2$ Emissions (Kt/Year) | 151.7 | 135.8 | 15.9 | 11.7% |
| Lignite | 94.7 | 91.4 | 3.2 | 3.5% |
| Hard Coal | 49.5 | 37.2 | 12.3 | 33.0% |
| Gas | 1.2 | 1.0 | 0.2 | 18.4% |
| Oil | 6.3 | 6.2 | 0.2 | 2.5% |
| $NO_x$ Emissions (Kt/Year) | 213.4 | 189.7 | 23.7 | 12.5% |
| Lignite | 121.5 | 116.8 | 4.7 | 4.0% |
| Hard Coal | 69.0 | 52.5 | 16.5 | 31.5% |
| Gas | 12.1 | 10.0 | 2.2 | 21.8% |
| Oil | 10.7 | 10.4 | 0.3 | 2.9% |
| PM Emissions (Kt/Year) | 5.5 | 4.9 | 0.6 | 12.2% |
| Lignite | 3.3 | 3.2 | 0.1 | 3.9% |
| Hard Coal | 2.0 | 1.5 | 0.5 | 30.3% |
| Gas | 0.1 | 0.1 | 0.0 | 24.6% |
| Oil | 0.2 | 0.1 | 0.0 | 3.3% |
| Mortality (Excess Deaths/Year) | 8,549.7 | 7,407.2 | 1,142.4 | 15.4% |
| Lignite | 4,142.9 | 3,988.1 | 154.9 | 3.9% |
| Hard Coal | 3,776.2 | 2,870.9 | 905.3 | 31.5% |
| Gas | 366.1 | 293.0 | 73.1 | 25.0% |
| Oil | 264.4 | 255.3 | 9.2 | 3.6% |
| Pollution Damages ($bn/Year) | 65.3 | 56.6 | 8.7 | 15.4% |
| Lignite | 31.6 | 30.5 | 1.2 | 3.9% |
| Hard Coal | 28.8 | 21.9 | 6.9 | 31.5% |
| Gas | 2.8 | 2.2 | 0.6 | 25.0% |
| Oil | 2.0 | 1.9 | 0.1 | 3.6% |

**Notes:** This table reports estimates for emissions of $CO_2$ as well as three local pollutants: $SO_2$, $NO_x$, and PM. The final row presents estimates of the mortality damages from all three of these local air pollutants. All values are annualized averages based on predictions from immediately after the March 2011 to the end of 2017. Emissions are the product of each plant's hourly generation with our estimate of their emissions rate. Emissions rates are the product of (a) the amount of fuel required to produce one unit of electricity, and (b) the emissions intensity of the fuel. Emissions estimates are limited to fossil-fuel-fired plants in Germany. We ignore other potential sources of emissions in the electricity sector, such as emissions from smaller biomass, landfill gas or waste plants. We also focus on emissions and damages in Germany and so do not estimate changes in emissions in neighboring countries due to changes in net imports. For the pollution damages reported in the last row of the table, we present only the monetary costs associated with premature mortality due to air pollution exposure in order to ensure consistency with the complementary analysis using pollution monitor data.

**Table 6** Impact of the Phase-Out on Local Air Pollution Mortality Damages

| | **Average with Phase-Out (1)** | **Average w/out Phase-Out (2)** | **Change (3)** | **Change (%) (4)** |
|---|---|---|---|---|
| $NO_2$ Emissions ($^{ug}/_{m^3}$) | 28.3 | 27.7 | 0.6 | 2.2% |
| Lignite | 24.9 | 24.4 | 0.5 | 2.1% |
| Hard Coal | 29.6 | 27.9 | 1.6 | 5.9% |
| Gas | 29.3 | 29.2 | 0.1 | 0.3% |
| Oil | 29.5 | 29.3 | 0.2 | 0.6% |
| $PM_{10}$ Emissions ($^{ug}/_{m^3}$) | 21.0 | 20.6 | 0.4 | 1.9% |
| Lignite | 21.7 | 21.3 | 0.4 | 1.7% |
| Hard Coal | 21.3 | 20.2 | 1.0 | 5.2% |
| Gas | 20.7 | 20.6 | 0.1 | 0.3% |
| Oil | 20.4 | 20.3 | 0.1 | 0.5% |
| $PM_{2.5}$ Emissions ($^{ug}/_{m^3}$) | 14.1 | 13.8 | 0.3 | 2.2% |
| Lignite | 15.1 | 14.8 | 0.3 | 2.3% |
| Hard Coal | 13.6 | 12.8 | 0.8 | 5.9% |
| Gas | 13.8 | 13.8 | 0.1 | 0.4% |
| Oil | 13.9 | 13.8 | 0.1 | 0.6% |
| Mortality ($^{\text{Excess Deaths}}/_{\text{Year}}$) | – | – | 493.0 | – |
| Lignite | – | – | 124.9 | – |
| Hard Coal | – | – | 315.7 | – |
| Gas | – | – | 20.2 | – |
| Oil | – | – | 32.3 | – |
| Pollution Damages ($^{\$\text{bn}}/_{\text{Year}}$) | – | – | 4.7 | – |
| Lignite | – | – | 1.2 | – |
| Hard Coal | – | – | 3.0 | – |
| Gas | – | – | 0.2 | – |
| Oil | – | – | 0.3 | – |

**Notes:** This table reports estimates of the monetary damages associated with the premature mortality resulting from the additional air pollution exposure as a consequence of the nuclear phase-out. The changes in daily concentrations of $PM_{2.5}$, $PM_{10}$, and $NO_2$ are obtained by panel regressions of air pollution at the monitor-level on daily, plant-level electricity production; these regressions include plant fixed effects, month-of-year fixed effects and year-of-sample fixed effects. The coefficients from these regressions give us an estimated relationship between electricity production and pollution concentration levels for each pollutant and each fuel type. We multiply the relevant estimated relationship by our predicted changes in production by each plant due to the phase-out. The resulting changes in air pollution concentrations are converted to a change in premature mortality using dose-response estimates from the ESCAPE project (Lancet 2014). We monetize this additional premature mortality using a value of statistical life of \$7.9 million for Germany taken from Viscusi and Masterman (2017). We assume that only the population residing within 20 km of Germany's fossil power plants is exposed to the air pollution from these plants (approximately 7.5% of the total population). We do not report the absolute levels of mortality or damages, only the change due to the phase-out, because the baseline levels of pollution recorded at monitors are not attributable entirely to power plant activity; for example, industrial facilities, cars, and trucks also emit these pollutants.

**Table 7** Overall Estimated Impact of the Nuclear Phase-Out on Total Costs

| | Average with Phase-Out (1) | Average w/out Phase-Out (2) | Change (3) | Change (%) (4) |
|---|---|---|---|---|
| **Total Costs ($^{\$bn}/$Year)** | 97.4 | 85.2 | 12.2 | 14.3% |
| **Private Costs** | | | | |
| Operating Costs | 14.2 | 12.6 | 1.6 | 12.7% |
| **External Costs** | | | | |
| $CO_2$ Climate Damages | 15.8 | 14.0 | 1.8 | 13.0% |
| Mortality from Local Pollution | | | | |
| Method 1: Reported Emissions | 65.3 | 56.6 | 8.7 | 15.4% |
| Method 2: Pollution Monitors | – | – | 4.7 | – |
| Local Pollution Morbidity | 1.9 | 1.6 | 0.2 | 14.1% |
| Nuclear Waste and Accidents | 0.3 | 0.4 | -0.2 | -38.2% |

**Notes:** This table reports the estimates of the different intensive margin costs incurred with versus without the phase-out. Private costs are the operating costs of the power plants in our analysis plus any changes in net imports (valued at the electricity price). We assume that the production costs of renewable and other sources are equal to zero when calculating these operating costs. External costs consist of climate damages from carbon emissions, mortality and morbidity costs from air pollution emissions, as well as the costs associated with nuclear accident risk and nuclear waste disposal. For the total costs row in bold, we use the estimates from the reported emissions method when adding in the external costs of local pollution on mortality.

# List of Tables



# List of Figures

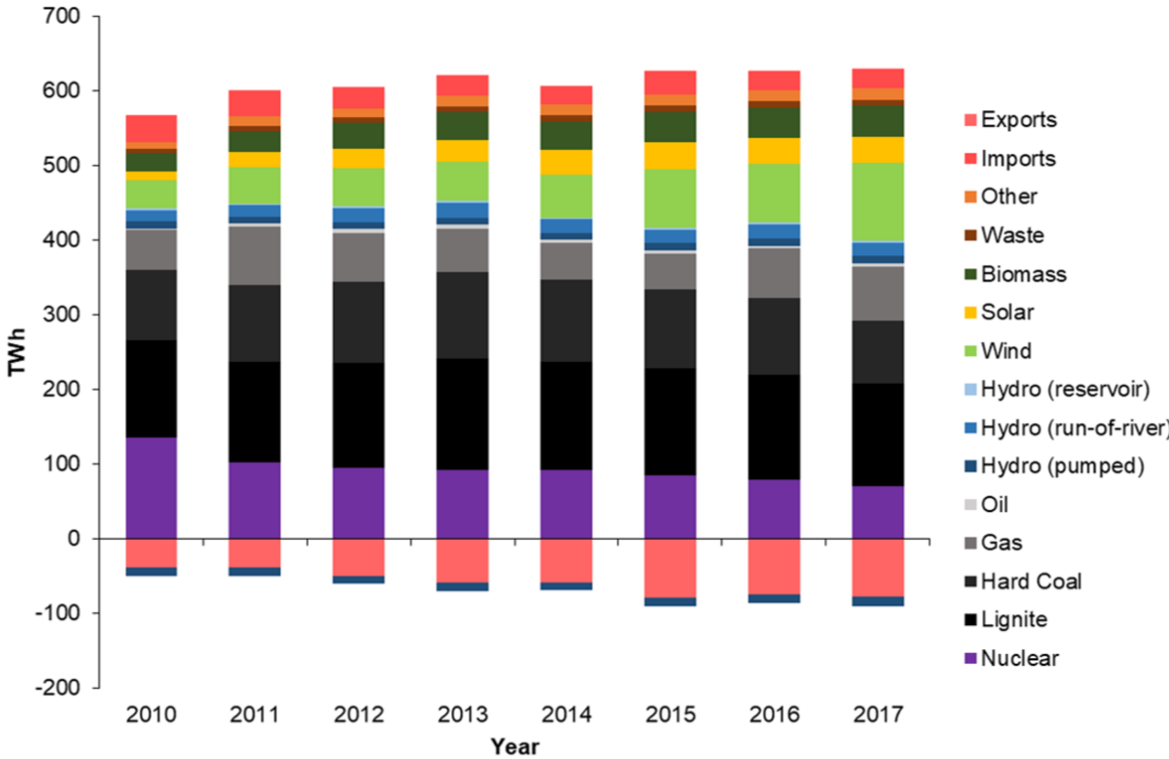
# Appendices

## A Appendix Tables and Figures

Appendix Figure A.1 presents annual total electricity production in Germany by source as well as total imports and exports. This figure documents the precipitous drop in nuclear production following the 2011 closure of nine reactors as well as the rapid growth in production from wind and solar resources over our 2010-2017 sample period.

[Figure A.1 about here.]

**Figure A.1:** Electricity Production by Source: 2010-2017




**Notes:** This figure plots the annual total quantity of electricity produced by different types of sources in Germany from 2010-2017. We also plot the annual total quantity of electricity imports and exports for this same sample period. The data underlying this figure are from BNetzA Monitoring Reports.

# B Further Detail on the Predictive Dispatch Model

Studies of the electricity sector traditionally utilize some form of electricity dispatch model that combines engineering and economic modeling tools to simulate the operation of the power grid. These models must explicitly specify firm incentives (ex: whether/how firms exercise market power) as well as operational constraints such as transmission congestion and plants' start-up/ramping costs.

We opt to employ an empirical approach instead. Specifically, our approach seeks to recover how plants are dispatched based on a host of different variables pertaining to plant operations, demand, and electricity transmission. The primary benefit of this empirical approach is that it requires fewer assumptions regarding firm incentives or operational constraints. We allow the data to tell us how these factors impact plant operations.

That being said, this empirical approach has limitations as well. First, we can only examine scenarios that are sufficiently similar to observed outcomes. This is why other empirical models of wholesale electricity markets tend to focus either on ex-post policy assessments or identifying how marginal changes in electricity demand impact plant operations. Indeed, our paper focuses on an ex-post evaluation of the nuclear phase-out in Germany on aggregate market outcomes.

We want to highlight that empirical approaches such as our typically do not offer robust insights for a given plant in a given hour. As such, our empirical modeling should be seen as a complement rather than a substitute for more explicit simulation modeling of electricity markets. This is particularly true when the behavior of individuals plants or short-term physical constraints are of interest rather than aggregate market outcomes.

Our paper utilizes a Random Forest algorithm. This algorithm has a number of useful properties. First, the relationships between our predictors and aggregate market outcomes are likely to be highly non-linear, including many complex interactions. Random Forests are well-suited to letting the data inform where these complex interactions lie rather than having to make strong ex-ante assumptions (e.g. no need to pre-specify polynomials, splines and interactions within a linear regression framework). Second, the structure of the Random Forest regression algorithm means that the support of possible outcome predictions is bounded by the support of the outcome values in the training data-set. This means that the predictions from our model will have a natural bounding of 0-1, thus avoiding the risk of making erroneous predictions (e.g. operating rates above 100% or below 0%).

Third, using Quantile Regression Forests allows us to make predictions regarding the full conditional distribution of our outcomes rather than just the conditional expectation of these outcomes. This is important because there is clearly uncertainty about whether a given plant will operate in a given hour conditional on the covariates for that plant-hour. However, being able to characterize the distribution of potential outcomes means we can (a) examine the uncertainty in our results, and (b) adjust our final estimation to calculate the most likely changes to in plant-level production that still meet physical requirements (i.e. that demand equals supply). For example, though our primary specifications report the conditional averages of predicted outcomes, we find that both the mean and median of the potential predictions produced by our model perform reasonably well (see Figure 4).

Our Random Forest model is estimated using a training dataset of roughly 4.5 million observations. The most important independent variables for our analysis are:

- **Net Load**. Net load is defined as total electricity demand minus production from low marginal cost or non-dispatchable sources. Specifically, we subtract production from renewables (wind, solar, hydro, biomass, waste) and nuclear. This net load variable thus measures the amount of production required by "flexible" (typically fossil-fuel fired) sources.[22]

- **Marginal Cost**. A plant decides whether to produce primarily based on whether its marginal cost is less than the electricity price it will be paid for its output. In electricity markets such as Germany's, the electricity price is typically set by the highest marginal cost plant necessary to meet demand (i.e. the clearing plant that is on the margin). Consequently, we first construct estimates of each plant's marginal cost over time. We then estimate the marginal cost of the clearing plant: the last fossil plant (or border point) necessary to meet net load in a given hour. Finally, we construct a "standardized" marginal cost for each plant as the plant's marginal cost minus the marginal cost of the clearing plant for that hour. Plants

[22]We also include lags and leads of net load to capture the fact that many power plants have dynamic production constraints (e.g. the speed at which they can "ramp up" their output, or the minimum amount of time they have to be offline before they can restart).

typically produce (don't produce) if this standardized marginal cost is negative (positive).

- **Available Capacity**. Where the "marginal cost" variable captures the position of a plant in the supply curve in terms of price, the "available capacity" variable captures the position of a plant in the supply curve in terms of quantity. For each plant, we calculate the total amount of capacity from other fossil plants (or border points) with a lower marginal cost. Our "available capacity" variable is then calculated as the total amount of capacity with a lower marginal cost than the plant minus net load for that hour. Once again, plants with negative available capacity are likely to produce, while plants with positive available capacity are unlikely to produce.

Figure B.1a illustrates the relative importance of each of our covariates. As expected, net demand, marginal cost and available capacity are all particularly important covariates. However, it is noteworthy that the two most important covariates are the type of source (i.e. lignite, hard coal, gas, oil or border point) and whether a fossil-fuel-fired plant is combined-heat-and-power. This reflects the fact that different types of electricity generators face different operational constraints. For example, many natural gas plants in Germany are combined-heat-and-power. As such, whilst they may have higher marginal costs than coal plants, they receive revenues both for their electricity output and from providing heating services. Consequently, combined-heat-and-power plants operate more frequently than would be suggested by simply comparing their marginal cost to electricity prices.

[Figure B.1 about here.]

The machine learning application we use is designed to predict how dispatchable flexible sources such as fossil-fuel plants and border flows increase or decrease their output in order to meet the residual demand left after accounting for output from renewables and nuclear sources. Net load, the relative marginal cost of each plant, and the amount of alternative available capacity are key predictors in the analysis not only because they play a significant role in explaining plant operating decisions, but also because they are the variables we modify in order to construct the counterfactual scenario. For the scenario with the phase-out, the net load variable is the observed net load given the phase-out decision as shown in Figure B.2a. For the counterfactual scenario without the phase-out, nuclear production would have been higher and so net load would have been lower, as shown in Figure B.2b. This reduction in net load also changes the marginal cost and available capacity variables. Specifically, if net load is lower, the marginal cost of the clearing plant would also be lower. Moreover, the amount of capacity below net load is also lower for lower values of net load. This is illustrated in Figures B.2c and B.2d.

[Figure B.2 about here.]

When making out-of-sample predictions using a predictive model such as this, it is important to ensure that the training data-set provides sufficient support across the predictor variables. This is because our algorithm is ill-suited to extrapolate beyond the economic conditions seen in the training data. We are confident that assessing the impacts of nuclear phase-out is an interpolation exercise rather than extrapolation exercise in part because the portfolio of fossil-fuel power plants and the underlying transmission grid does not change very much over our 2010-2017 sample period.

Rescaling certain variables can also help to ensure that our out-of-sample prediction is not extrapolating too far outside the support of the training data.[23] The three main variables we use to approximate the interaction between supply and demand are net load, plant marginal costs, and the amount of available capacity. Almost by definition, the counterfactual no-phase-out scenario will contain some periods where these variables fall outside the range in the training dataset. Even so, there is such wide variation in electricity demand, production from renewables and marginal costs that the overlap in support between these variables in the factual versus counterfactual scenarios is very good. This can be seen in Figures B.1b, B.1c and B.1d.

Figure B.3 shows the median model predictions for how the nuclear phase-out impacted aggregate plant-level electricity production in Germany. As expected, points on this figure tend to lie above the horizontal axis; the nuclear phase-out reduced nuclear generation, with fossil-fuel-fired production filling the gap. The largest response to the phase-out comes from the hard coal plants.

[23]For example, we rescale the marginal costs of each plant by the marginal cost of the last plant needed to clear the market. Even if fuel costs doubled from 2010-20177, for example, the rescaling would ensure that the rescaled marginal costs fed into our algorithm stay within a reasonable range over our sample period.

[Figure B.3 about here.]

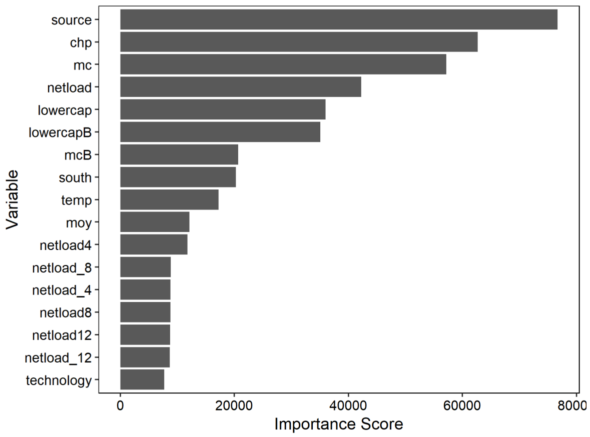
Using the median predictions displayed in Figure B.3 we find around 40 TWh per year of additional supply from higher fossil-fuel plants and net imports. However, it is important to note that there is no constraint in our estimation process that the total amount of estimated replacement production should match the lost nuclear output. In fact, the amount of lost nuclear production is around 50 TWh per year and so using the median predictions actually leads us to under-estimate the level of replacement generation. To remedy this, we utilize the information our quantile regression model provides us on the full conditional distribution of potential changes to output. Specifically, we generate predictions for the 10th, 25th, 50th, 75th and 90th percentiles of each of our outcomes. We then find the combination of these percentiles that fully replaces the lost nuclear generation with the most likely set of plant-level changes (i.e. closest to the median). Put another way, we find the percentiles closest to the median that produce a change in annual total generation equal to the annual lost nuclear output. Ensuring that additional supply exactly meets lost nuclear output only requires moving a few percentiles from the median.

Finally, Figure B.4 illustrates which plants and border points increased production to meet the reductions in nuclear output due to the phase-out. Most of the fossil-fuel generation comes from the industrial regions in the west and south of the country. Changes to net imports come primarily at the borders with Denmark, France and the Czech Republic.

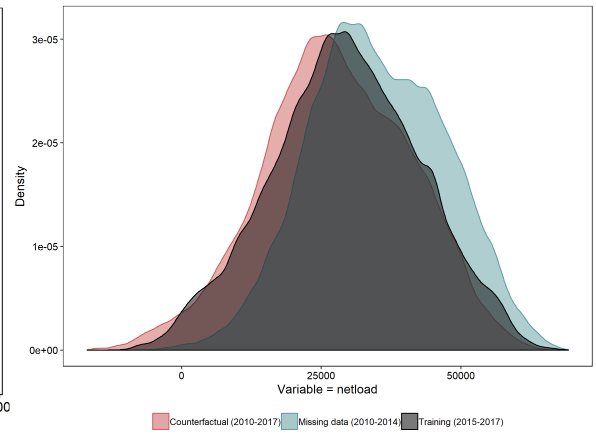
[Figure B.4 about here.]

**Figure B.1:** Machine Learning Model Diagnostics

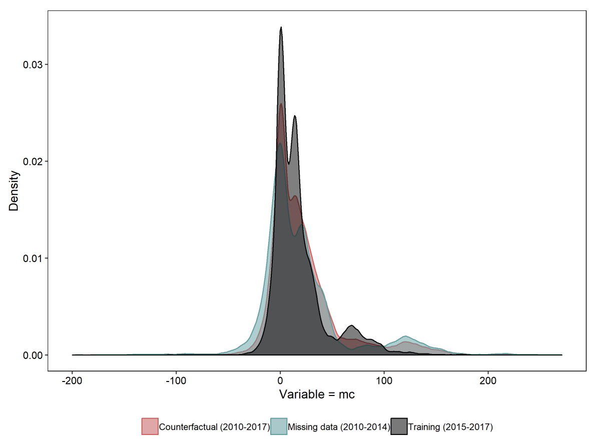
**(a)** Variable importance scores



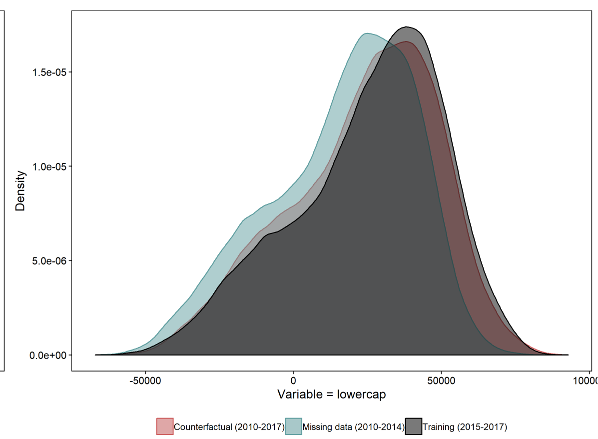

**(b)** Support of Net Demand




**(c)** Support of Marginal Cost



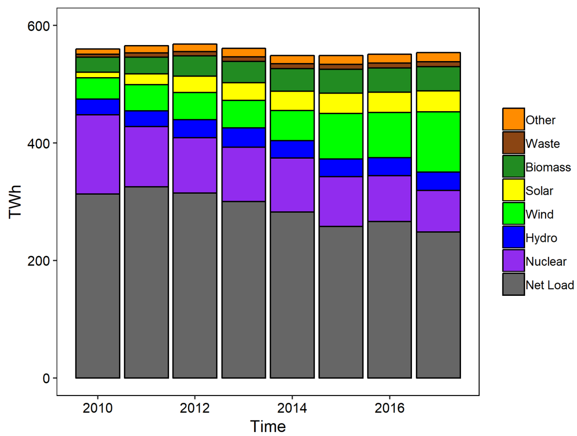

**(d)** Support of Available Capacity



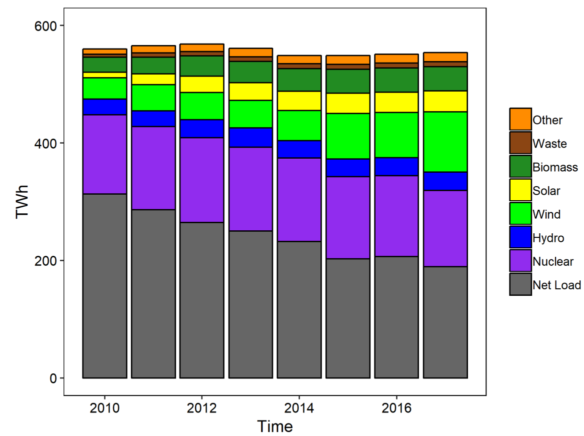

**Notes:** This figure illustrates a range of key model diagnostics related to the machine learning estimation. Panel (a) shows the importance scores for each of the variables included in the estimation. Importance scores indicate the relative importance of each variable in predicting the outcome of interest. The abbreviated names in the figure are as follows: source = source type (e.g. lignite, hard coal, gas, oil or border); mc = marginal cost relative to the clearing unit; mcB = marginal cost relative to the clearing unit including border capacity; lowercap = amount of capacity with a lower marginal cost; lowercapB = amount of capacity with a lower marginal cost including border capacity; chp = presence and scale of combined-heat-and-power capability; technology = technology type (e.g. steam turbine, combined cycle turbine or transfer); temp = local temperature; south = indicator for whether the plant or border point is located in the south of the country; moy = month-of-year; dow = day-of-week; hod = hour-of-day; netload = electricity load minus production from wind, solar, hydro and nuclear sources; netloadX = difference between current net load and net load X hours ago; netload_X = difference between current net load and net load X hours ahead. Panels (b-d) show the support of three key variables: net demand, standardized marginal cost and available capacity. The grey area shows the distribution of observations in the 2015-2017 training data-set (i.e.: where we have hourly, plant-level production data). The blue area shows the distribution of observations in the missing 2010-2015 data (i.e.: where we only have hourly data on production by fuel type). The red area shows the distribution of observations in the counterfactual scenario (i.e.: without the nuclear phase-out) across the full 2010-2017 analysis period.

**Figure B.2:** Net Demand and Scenario Implementation

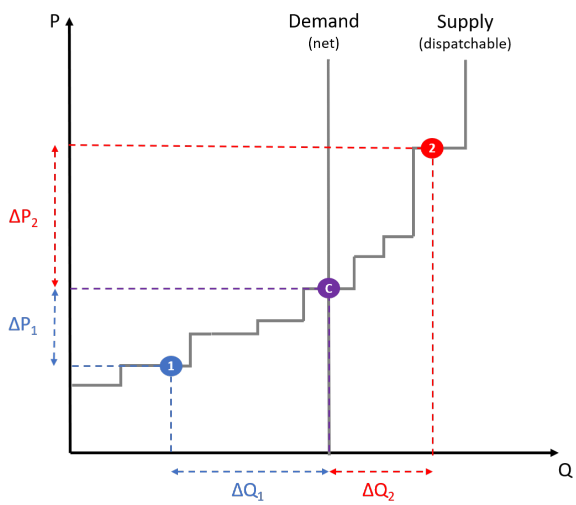
**(a)** Estimated Net Demand (With Phase-Out)



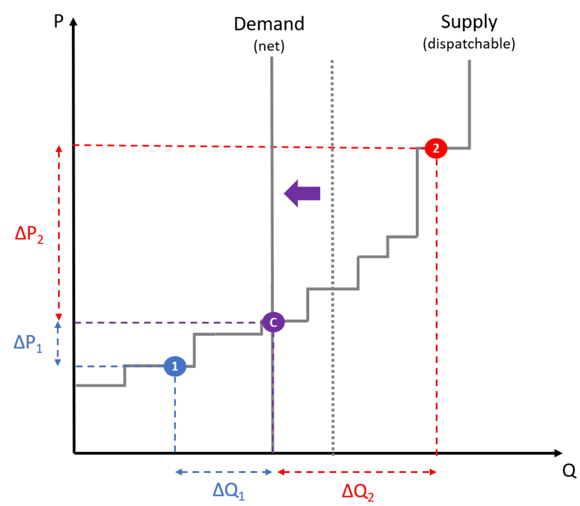

**(b)** Estimated Net Demand (Without Phase-Out)




**(c)** Net Demand Illustration (With Phase-Out)

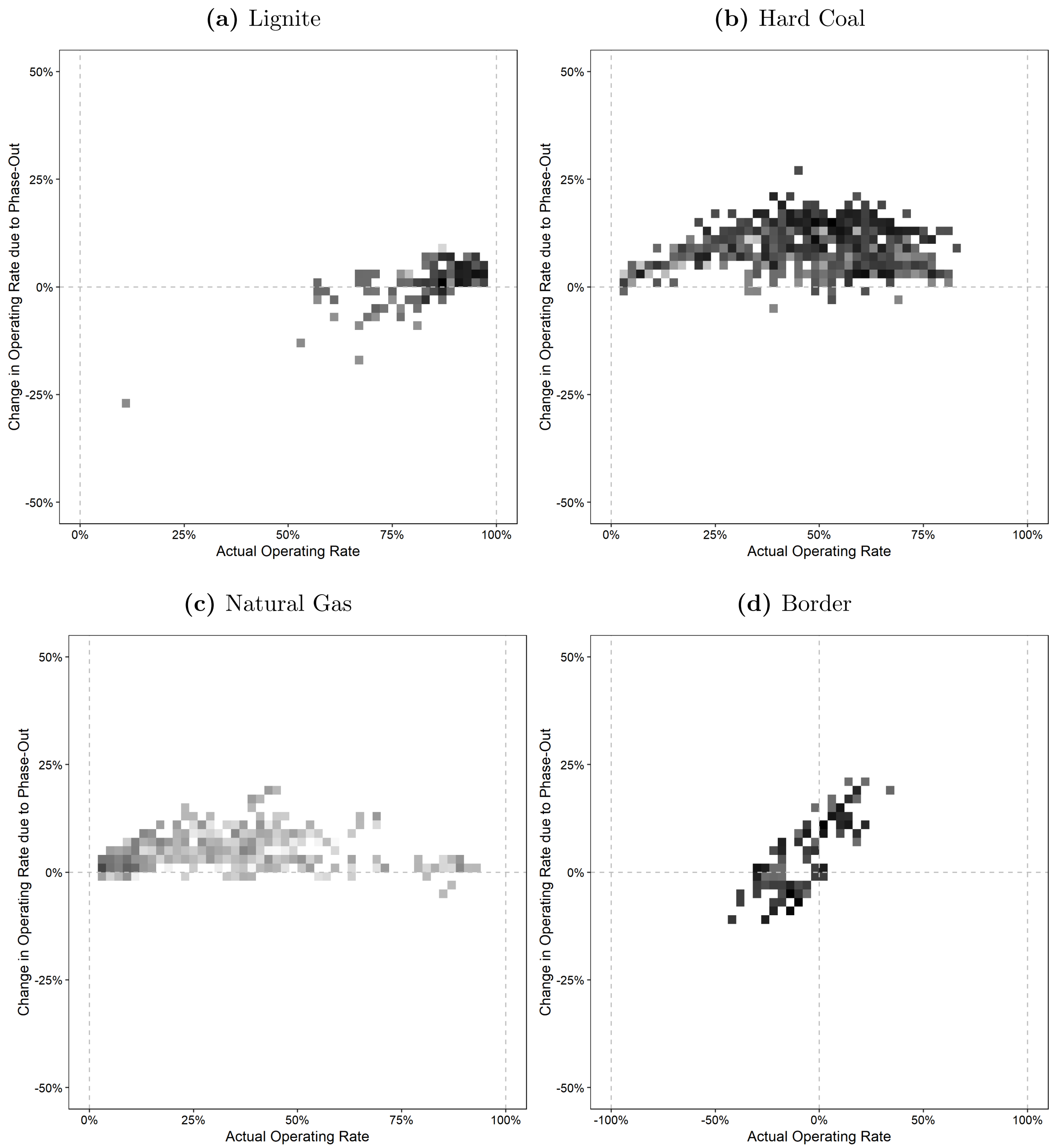



**(d)** Net Demand Illustration (Without Phase-Out)



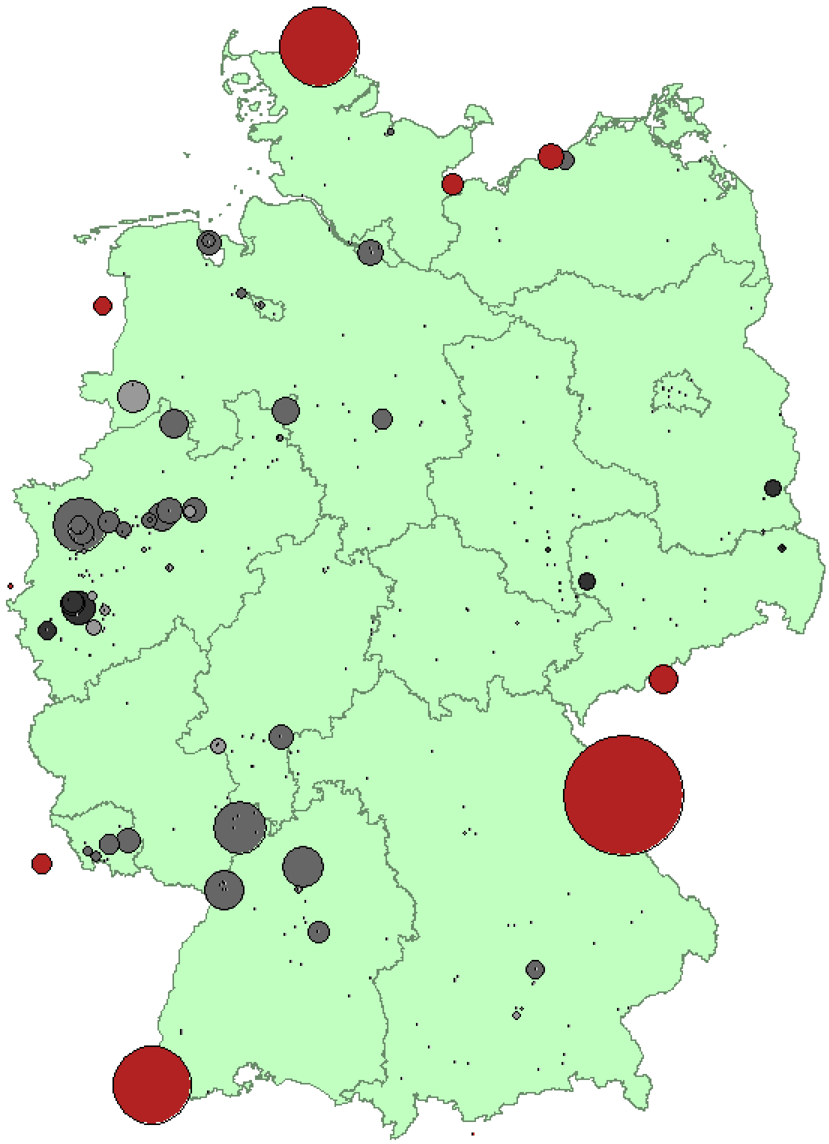

**Notes:** This figure illustrates the role of the net electricity demand variable in the analysis. Net demand is defined as total electricity demand minus production from low marginal cost or non-dispatchable sources. Specifically, we subtract production from renewables (wind, solar, hydro, biomass, waste) and nuclear. Panels (a) and (b) show the level of net demand both with and without the phase-out respectively. Note that production from renewables is growing over time, which results in less net demand to be satisfied by flexible sources such as fossil-fuel fired plants. Comparing panel (a) to panel (b) shows that more nuclear production without the nuclear phase-out leads to less net demand to be satisfied in this scenario. Panels (c) and (d) provide an illustration of how changing net demand impacts the estimation process. This happens because altering net demand alters the position where net demand intersects with the supply curve of dispatchable capacity. This intersection point is indicated by the clearing fossil-fuel plant (or border point) that is "on-the-margin" (purple). Altering the clearing fossil plant (or border point) affects the relative marginal cost ($\Delta P$) and available capacity ($\Delta Q$) values for all dispatchable supply. These two variables are illustrated for a high marginal cost plant (red) and a low marginal cost plant (blue).

**Figure B.3:** Plant-level Changes in Production due to the Phase-Out




**Notes:** This figure illustrates the plant-level disaggregation of the machine learning prediction model results. The model predicts the operating rate of each power plant in each hour, where a value of 0% means the plant is offline and a value of 100% means it is running at maximum capacity. These figures plot plant-level annual average operating rates. The x-axis corresponds to each plant's operating rate in the baseline scenario with the phase-out. The y-axis corresponds to the impact of the phase-out on plant-level operations. This is determined by the difference between the predictions in the scenario with the phase-out versus the scenario without the phase-out. Darker areas indicate higher numbers of plant-year observations. Each panel refers to a different type of dispatchable electricity source. Panel (a) covers lignite plants, panel (b) covers hard coal plants, panel (c) covers gas plants and panel (d) covers border points. Oil plants are not shown because they are a very small portion of total capacity and are largely invariant to the phase-out.

**Figure B.4:** Map of Plant-Level Changes in Production due to the Phase-Out



**Notes:** This map illustrates the location of the fossil-fuel-fired plants or border points that increased their electricity production as a result of the nuclear phase-out policy. The size of the circle reflects the amount of additional production provided by the fossil-fuel plant or border point. Points in red are border points and points in grey are fossil-fuel plants. Lignite plants are depicted in the darkest grey, followed by hard coal, then natural gas, and finally oil plants are depicted in the lightest grey.