# How Effective was the UK Carbon Tax?—A Machine Learning Approach to Policy Evaluation

*By* Jan Abrell, Mirjam Kosch, and Sebastian Rausch*

January 2019

*While carbon pricing is largely viewed as a rationale policy response to climate change by the economics discipline, surprisingly little is known about its performance from an ex-post perspective. This paper evaluates the environmental and cost effectiveness of the UK Carbon Price Support, a carbon tax levied on all fossil-fueled power plants in the market. We propose a novel approach for estimating the treatment effect of a policy intervention in the absence of a control group which leverages machine learning techniques, high-frequency market data, and economic theory. We find that between 2013 and 2016 the carbon tax reduced emissions by 6.2 percent at an average cost of 18.2 Euro per tonne. Simulating the machine-learned model, we characterize the empirical conditions influencing the effectiveness of the tax policy. We find that the ratio of carbon tax-exclusive prices for coal and natural gas is by far the most important driver. (JEL Q41, Q42, Q58)*

To avoid dangerous and costly climate change, the disposal space for carbon dioxide ($CO_2$) in the atmosphere is "scarce" and will soon be exhausted (McGlade and Ekins, 2015; IPCC, 2018). In tackling this major 21$^{\text{st}}$-century challenge, and based on an elementary understanding of how today's market-oriented systems organize economic activity based on scarce resources, economists have long been advocating for carbon pricing as an effective and efficient policy response (Nordhaus, 1994; Goulder and Parry, 2008; Metcalf, 2009). Currently, about 20-25 percent of global $CO_2$ emissions are regulated under some form of carbon pricing (World Bank, 2018). While a plethora of studies offers ex-ante assessments of carbon pricing using theoretical and quantitative simulation-based work,[1] surprisingly little is known about the *ex-post* effects of carbon pricing. This, however, is pivotal for designing effective and efficient climate policies in the future.

This paper contributes by providing an ex-post evaluation of a real-world policy

* Abrell: Center for Energy and Environment of ZHAW Winterthur and Centre for Energy Policy and Economics ETH Zurich, Switzerland (email: jan.abrell@zhaw.ch). Kosch: Center for Energy and Environment of ZHAW Winterthur and Centre for Energy Policy and Economics ETH Zurich, Switzerland (email: mirjam.kosch@zhaw.ch). Rausch: Department of Management, Technology and Economics, ETH Zurich, Switzerland, Center for Economic Research at ETH (CER-ETH), Centre for Energy Policy and Economics ETH Zurich, and Massachusetts Institute of Technology, Joint Program on the Science and Policy of Global Change, Cambridge, USA (email: srausch@ethz.ch).

[1] See, for example, Tavoni et al. (2014), Golosov et al. (2014), Liski and Gerlagh (2016), Goulder, Hafstead and Williams III (2016), Bretschger et al. (2017), and a series of papers from multi-model comparison studies carried out under the framework of the Stanford Energy Modeling Forum for the U.S. (Fawcett et al., 2014) and Europe (Weyant et al., 2013).

experiment of carbon pricing: the UK carbon tax, also known as the *Carbon Price Support* (CPS). The CPS was introduced to enhance economic incentives for $CO_2$ abatement in the heavily fossil-based UK electricity sector. As the CPS affects the output and operating decisions of all fossil-fueled generation facilities in the UK electricity market, the main challenge arises that no suitable control group or counterfactual exists against which the impact on treated units can be evaluated. In order to estimate the causal effects of the CPS policy intervention, it is thus not possible to use standard program evaluation methods based on comparing treated and untreated units.[2] To overcome this problem, we develop and implement a new approach which combines economic theory and machine learning (ML) techniques to establish causal inference of a policy intervention in settings with observational, high-frequency data when no control group exists. We apply our approach to analyze the environmental and cost effectiveness of the UK carbon tax. To our knowledge, this is the first paper in economics to incorporate ML methods to conduct causal inference of carbon pricing.

In a nutshell, the idea of our new approach to estimate treatment effects in the absence of a control group is as follows. First, we derive a structural causal model of the observed outcomes (electricity output by power plant) based on an economic model of wholesale market activity. Second, we use ML techniques—which are optimized for out-of-sample prediction (Mullainathan and Spiess, 2017)—to train the causal model, i.e. to estimate a predictor function for electricity output of each plant in the market given fuel prices, available capacities, and demand. We train the model based on both pre- and post-intervention data combining hourly panel data of electricity output at the plant-level with data on hourly electricity demand, available plant capacity, fuel and carbon prices, and temperature. To overcome the problem of insufficiently large variation in the treatment variable—the UK carbon tax rate is adjusted on an annual basis leaving us with only four CPS levels over our period of analysis—we exploit the variation in the *relative* market prices for coal and natural gas, including the prices for European Emission Allowances (EUA) and the CPS. As the market substitution from carbon-intensive coal-fired power plants to less polluting natural gas plants (coal-to-gas switch) hinges on relative fuel prices, we identify the impact of the CPS on plant output by using the CPS-inclusive fuel price ratio as the pseudo-treatment variable. In a third step, we rely on the concept of the "do-operator" (Pearl, 2009) to obtain the treatment effect for each power plant. Specifically, we estimate the treatment effect as the difference between predicted outcomes with and without policy. By using the difference of predicted outcomes to estimate the treatment effect, we account for the impact of unobserved variables as well as systematic prediction errors in the ML-trained model.

An important feature of our approach is to identify and explicitly represent the channels through which the policy intervention affects the outcome variable. This has three important advantages: First, using a structural causal model, we are able

[2]Such as, for example, difference-in-differences (DiD), regression discontinuity design, and synthetic control methods (Angrist and Pischke, 2008; Athey and Imbens, 2017).

to estimate the treatment effect even in the absence of a control group. Second, it opens up the possibility to go beyond the estimation of the total treatment effect by using the ML-trained model to assess the impact of different (hypothetical) treatment intensities using simulations. Third, as our pseudo-treatment variable—the carbon price-inclusive fuel price ratio, which trivially contains the case of no CPS policy—is already observed before the CPS policy is introduced, we can use observations from both the pre- and post-treatment period to train the model. This improves the basis for learning about the key mechanism between input prices and output through which the policy intervention impacts the economic behavior of firms' output.

Based on our new approach, the ex-post evaluation of the UK carbon tax yields the following main insights. First, our analysis provides strong evidence that a carbon tax is an effective regulatory instrument to reduce $CO_2$ emissions: the CPS induced a substitution away from "dirty" coal to cleaner natural gas-fired power plants—replacing about 15 percent or 46 TWh of coal-based generation and reducing emissions by 6.2 percent between 2013 and 2016. Second, the abatement of one ton of $CO_2$ has brought about on average (over all fossil-based plants and the 2013-2016 period) additional costs of €18.2 in total for consumers and fossil-based electricity producers. Third, simulating the machine-learned model, we characterize the empirical conditions which influence the environmental and cost effectiveness of the tax policy. We find that there is substantial heterogeneity in the carbon tax-induced market impacts over time, which are mainly driven by the level of the CPS and the ratio of carbon tax-exclusive prices for coal and natural gas.

Our paper contributes to the literature in several important ways. First, we add to the recent and emerging literature on the use of ML techniques in economics and quantitative social science. Traditionally, ML methods have been used for pure prediction problems such as demand estimations (Bajari et al., 2015). More recently, ML methods have provided important new tools to improve the estimation of causal effects from observational data in high-dimensional settings as they enable to flexibly control for a large number of covariates (for overview articles see, for example, Varian, 2014; Athey, 2017; Athey and Imbens, 2017; Mullainathan and Spiess, 2017). Burlig et al. (2017) and Cicala (2017) are two recent examples using ML algorithms to estimate causal effects. Our approach differs in two important ways. First, they deal with discrete treatment leading to a change in the data generating process (DGP) between the pre- and the post-treatment period. They, therefore, use the pre-treatment period to train a model predicting the post-treatment outcome without the intervention. In contrast, we deal with a setting with an invariant DGP and continuous treatments. Therefore, we are able to train the model on the full sample, but at the same time have to rely on the continuity of treatment or, alternatively, have to identify a variable with the same causal impact as the treatment variable. Second, ML based predictions have to deal with prediction errors impacted by unobserved variables. Burlig et al. (2017) and Cicala (2017) assume that prediction errors have similar trends across treatment and control groups. Therefore, they use a DiD estimator to eliminate biases caused by prediction errors. In contrast, we eliminate this bias comparing predicted values of observed and counterfactual values, i. e., we assume that prediction errors are independent of treatment levels. While Varian (2016) mentions the possibility of estimating treatment effects by constructing the unobserved counterfactual when no control group is available, to the best of our knowledge this paper provides the first empirical implementation of this idea in economics.

Second, a recent and growing literature, following the U.S. shale gas boom after 2005, uses the variation in natural gas prices to empirically estimate the impact of fuel prices on $CO_2$ and other pollutants stemming from electricity generation (see, for example, Knittel, Metaxoglou and Trindade, 2015; Linn, Muehlenbachs and Wang, 2014; Holladay, Soloway et al., 2016; Holladay and LaRiviere, 2017). Cullen and Mansur (2017) and Lu, Salovaara and McElroy (2012) exploit the fact that the introduction of a carbon price impacts emissions through the same economic mechanism as a change in gas prices. Similar to our approach, these studies use the variation in natural gas prices to estimate the impact of a hypothetical carbon pricing policy on emissions. In contrast, we contribute with an ex-post assessment of a real-world, existing carbon tax policy.

Third, studies investigating the environmental impact of carbon pricing in the electricity sector are abundant but the vast majority of the work relies on numerical simulation methods (see, for example, Delarue, Ellerman and D'Haeseleer, 2010*b*; Delarue, Voorspools and D'Haeseleer, 2008; Rausch and Mowers, 2014; Goulder, Hafstead and Williams III, 2016). Some of the economic mechanisms at work, which we empirically identify in our analysis, have already been analyzed using ex-ante policy analysis based on analytical and simulation models. For example, Kirat and Ahamada (2011) show that the high permit prices induced a switch in the merit order from coal to gas. Delarue, Ellerman and D'Haeseleer (2010*a*) show that abatement does not only depend on the level of carbon prices but also on demand and the ratio between coal and gas prices. Some studies model the fuel switching potential for hypothetical carbon pricing policies as in Pettersson, Söderholm and Lundmark (2012) for the EU ETS and Chevallier et al. (2012) for the UK.

Fourth, there exists only a handful of studies using econometric and program evaluation methods to quantify the environmental impacts of carbon pricing, be it through a tax- or quantity-based approach to regulation. An overview of the work focusing on the EU ETS is provided by Martin, Muûls and Wagner (2016). The paper by McGuinness and Ellerman (2008) represents, to the best of our knowledge, the only analysis of empirically estimating the impact of permit prices on the output of power plants in the UK. Using a panel regression, they quantify the emissions offset in the British power sector for the pilot trading period of the EU ETS. Martin, De Preux and Wagner (2014) analyze the impacts of the *Climate Change Levy* on manufacturing plants in the UK. Using panel data on manufacturing plants in the UK, their identification strategy builds on the comparison of outcomes between plants subject to the full tax and plants paying only 20 percent of the tax. With this paper, we contribute to the scarce empirical evidence on the economic impacts of carbon taxes by applying an estimation strategy which can be used in a setting without a control group.

The remainder of this paper is organized as follows. Section I presents our methodological framework to estimate the treatment effects of a policy intervention in the absence of a control group. Section II details how we apply the framework to assess the effectiveness of the UK carbon tax, including a description of the data and our empirical strategy. It also scrutinizes the validity our approach for estimating the causal effects of the policy intervention. Section III presents and discusses our main findings. Section IV concludes.

## I. Conceptual Framework

### A. Overview

We begin by providing a conceptual description of our proposed framework to estimate the causal effects of a policy intervention when a suitable control group does not exist and when treatment intensity varies over time but not across treated units. The framework comprises three major steps:

*Step 1:* Derive a structural model of the observed outcomes based on economic theory which is invariant to the policy intervention;

*Step 2:* Use machine-learning (ML) techniques to train the causal model, i.e. to estimate a predictor of outcomes based on the causal model;

*Step 3:* Estimate the treatment effect as the difference between predicted outcomes under observed and counterfactual values of the policy intervention (while holding other controls constant).

Before turning to a detailed description of each of the three steps, two general features of our proposed method are important to emphasize. First, it does not rely on the existence of multiple units. The estimation of the predictor function requires sufficient data for a single unit. As a consequence, we derive a time-unit specific treatment effect which can also be computed for a single unit. Second, by relying on ML techniques we allow for non-parametric predictors and therefore for the non-parametric identification of treatment effects.

### B. The Causal Model

Consider a population model according to which the outcome $y_{it}$ of unit $i$ in period $t$ is generated according to

$$y_{it} = f_i\left(x_{it}, h_{it}, z_t\right) + \epsilon_{it}\,, \tag{1}$$

where $z_t$ is the treatment received by all units at time $t$. $x_{it}$ and $h_{it}$ are vectors of observed and unobserved control variables, respectively. $\epsilon_{it}$ is a random noise which is distributed with zero mean, $\mathbb{E}[\epsilon_{it}] = 0$ and variance $\sigma^2_\epsilon$, $\epsilon_{it} \sim (0, \sigma^2_\epsilon)$. $\epsilon_{it}$ is independent of controls and treatment:

$$\epsilon_{it} \perp\!\!\!\perp \left(x_{it}, h_{it}, z_t\right) \quad \forall i, t\,. \tag{2}$$

For each unit $i$, we observe a sample of outcomes $Y_i := (y_{i1}, y_{i2}, \cdots, y_{iT})^\top$ and control variables $X_i := (x_{i1}, x_{i2}, \cdots, x_{iT})^\top$ of size $T$, where $T$ is the number of time periods. While outcomes and controls are observed at the unit level, observed treatment levels are uniform across the population, i.e. we only observe the sample of treatment levels $Z := (z_1, z_2, \cdots, z_T)^\top$—as is, for example, the case for an environmental tax which is levied equally on all units in the market.

We are interested in identifying the causal effect on outcome which is induced by a change in the treatment level from its observed value $z_t$ to a specific value $\overline{z}_t$. To derive the effect of a change in the treatment variables $z$, we make use of an important assumption on the data generating process given by equation (1): the function $f_i$ is invariant to changes in the treatment and control variables (Peters, Bühlmann and Meinshausen, 2016) or, put differently, $f_i$ is assumed to be autonomous (Haavelmo, 1944; Aldrich, 1989).Given the invariance property, we are able to change the treatment variable and use the autonomous process to calculate the outcome under the changed treatment. The treatment effect is then defined as the difference between observed outcomes $y_{it}$, which realized under observed treatment levels $z_t$, and counterfactual outcomes $y_{it}^{\overline{z}}$ under hypothetical treatment levels $\overline{z}_t$:

$$\delta_{it}^{\overline{z}} := y_{it} - y_{it}^{\overline{z}} \quad \forall i, t \tag{3}$$

The *fundamental problem of causal inference* (Holland, 1986), often also referred to as the *missing data problem* (Rubin, 1974), is that we do not observe $y_{it}^{\overline{z}}$ and hence cannot directly calculate the treatment effect. If the treatment level varies across units, for example, matching or difference-in-differences (DiD) methods have been put forward to solve this problem by exploiting the existence of treated and untreated, i. e., control units. The fundamental challenge of our policy evaluation problem is, however, that the treatment (i.e., the carbon tax in the electricity sector) is uniform across the entire population (i.e., it is imposed equally on *all* power plants in the market). We are thus charged with the problem of finding a way to estimate the causal effect of the policy intervention without the possibility of relying on an untreated control group.

We propose to overcome the *missing data problem* by making use of counterfactual simulation which can create the unobserved outcomes $y_{it}^{\overline{z}}$. The main idea of our proposed approach is to predict the outcome for a counterfactual level of the treatment using a causal model for which we can change, i.e. control, the treatment variable. Pearl (2009) conceptualizes such a counterfactual intervention based on a causal model by his $do\,()$ operator. Given the possibility to perform $do\,()$-interventions, we can re-write the treatment effect as:

$$\delta_{it}^{\overline{z}} := y_{it} - f_i\left(x_{it}, h_{it}, do\left(z_t = \overline{z}_t\right)\right) - \epsilon_{it} \quad \forall i, t\ . \tag{4}$$

In order to calculate counterfactual outcomes, the following two assumptions concerning the interaction between controls and the treatment variable have to be satisfied:

ASSUMPTION 1: *Observed controls are independent of the changes in the treatment variable:*

$$x_{it} \perp\!\!\!\perp z_t \,.$$

ASSUMPTION 2: *Unobserved controls are conditionally independent to changes in the treatment variable given the observed controls:*

$$h_{it} \perp\!\!\!\perp z_t | x_{it} \,.$$

Assumption 1 rules out effects of the treatment variable on observed controls. This assumption is necessary as the observed controls are held constant in the counterfactual simulation. Otherwise, if $z$ influences $x$, there would be an indirect effect on the outcome, which would bias our estimate of the treatment effect.

Assumption 2 rules out effects of the treatment variable on unobserved variables after controlling for the observed variables. Again, if $z$ would influence $h$, there would be an indirect effect on the outcome. It is important to note that Assumption 2 does not rule out an effect of unobserved controls. It only implies that once we include all observed controls into the model, the impact of unobserved variables is independent of the treatment level, and, thus, a change in the treatment does not affect the outcome indirectly by changing unobserved variables.

### *C. Using Machine Learning for Prediction Models*

To predict counterfactual outcomes $y_{it}^{\overline{z}}$, we need an estimator $\hat{f}_i$ of the function $f_i$ that produces reliable *out-of-sample* predictions. We harness the power of ML methods which—in contrast to traditional econometric methods focused on consistently estimating *in-sample* parameters of $f$—are optimized to predict the value of the outcome variable (Mullainathan and Spiess, 2017).

Out-of-sample optimization is typically achieved by minimizing the expected prediction error. We use the mean squared error (MSE) as a measure of prediction quality whose expected value can be decomposed as follows:[3]

$$\mathbb{E}[\mathrm{MSE}_i] = \mathbb{E}[\left(y_i - \hat{f}_i\right)^2] = \sigma_\epsilon^2 + \underbrace{\left(\mathbb{E}[\hat{f}_i] - f_i\right)^2}_{=\,\mathrm{Bias}^2(\hat{f}_i)} + \underbrace{\mathbb{E}[\left(\mathbb{E}[\hat{f}_i] - \hat{f}_i\right)^2]}_{=\,\mathrm{Variance}(\hat{f}_i)} \,. \tag{5}$$

The expected prediction error thus consists of three parts: an irreducible population error, which corresponds to the variance of the random noise $\sigma_\epsilon^2$ (Hastie, Tibshirani and Friedman, 2008), and bias and variance terms which are both reducible. Standard econometric techniques, such as ordinary least squares (OLS) regression, aim at minimizing the bias while allowing for high variance. While these methods are thus capable of representing very well the sample data, they are prone to over-fitting and they yield prediction outcomes that are highly dependent on the observed sample.

[3] See, for example, Hastie, Tibshirani and Friedman (2008) and Gareth et al. (2013).

ML methods, in contrast, solve a bias-variance trade-off in order to find the best prediction model. ML methods address the bias-variance trade-off by introducing *hyper- or tuning parameters* in the estimation function. These parameters control for model complexity by decreasing the variance at the cost of a higher bias. The selection of hyper-parameters $\alpha$ is achieved through a process called cross-validation (CV), which makes optimal use of the available data. The CV process starts by splitting the observed sample into several subsets. One of the subsets, called the training set, is then used to estimate the predictor for a given set of hyper-parameters, $\hat{f}_i^{\alpha}$, by minimizing the expected *in-sample* MSE:

$$\hat{f}_i^{\alpha} := \arg\min_{f_i \in \mathcal{F}} \sum_t \left[\left(y_{it} - \hat{f}_i^{\alpha}\left(x_{it}, z_t\right)\right)\right]^2 \tag{6}$$

where $\mathcal{F}$ denotes the set of all possible functions $f_i$. The *out-of-sample MSE* is then computed on the remaining data—called the test or hold-out set—which has not been used for the estimation. Repeating this procedure for all subsets and averaging over all *out-of-sample MSE* yields an estimate of the expected prediction error for a given set of hyper-parameters $\alpha$.

The optimal set of hyper-parameters $\alpha^*$ is the one that minimizes the expected prediction error which is obtained from using a grid search over different candidate sets. Given $\alpha^*$, the final predictor $\hat{f}_i^{\alpha*}$ is obtained by solving the problem in equation (6) on the full sample of data. Finally, the true value of outcome in equation (1) can be written as the the sum of the predicted value and the prediction error $\xi\left(x_{it}, h_{it}, z_t\right)$:

$$y_{it} = \hat{f}_i^{\alpha*}\left(x_{it}, z_t\right) + \underbrace{f_i\left(x_{it}, h_{it}, z_t\right) - \hat{f}_i^{\alpha*}\left(x_{it}, z_t\right)}_{=:\xi(x_{it}, h_{it}, z_t)} + \epsilon_{it}\,. \tag{7}$$

### D. *Estimation of Treatment Effects through Counterfactual Simulation*

In the last step, we can use $\hat{f}_i^{\alpha*}$ to predict the missing outcome under a counterfactual level of the treatment. A simple estimator of the treatment effect would then compare *observed outcomes* under treatment with *predicted outcomes* without treatment as suggested by equation (4). Doing so would, however, result in bias estimates due to the prediction error shown in equation (7).[4] To estimate the treatment effect, we therefore need to eliminate the prediction error. This requires a further assumption:

ASSUMPTION 3: *The prediction error $\xi\left(x_{it}, h_{it}, z_t\right)$ is independent of the treatment:*

$$\xi\left(x_{it}, h_{it}, z_t^0\right) = \xi\left(x_{it}, h_{it}, z_t^1\right) = \xi\left(x_{it}, h_{it}\right) \quad \forall z_t^0, z_t^1$$

[4]Using the definition of the treatment effect (3) and equation (7), an estimator comparing the observed values $y_{it}$ with a predicted counterfactual value you yield: $\hat{\delta}_{it}^{\bar{z}} = y_{it} - \hat{f}_i^{\alpha*}\left(x_{it}, z_t = \bar{z}_t\right) = y_{it} - y_{it}^{\bar{z}} + \xi\left(x_{it}, h_{it}, z_t = \bar{z}_t\right) + \epsilon_{it}$. Therefore, this estimator would be biased by the prediction error.

Assumption 3 implies that the prediction error only depends on observed and unobserved variables, but does not change between the prediction of observed and counterfactual outcomes. Consequently, it allows to estimate the treatment effect $\hat{\delta}_{it}^{\overline{z}}$ as the difference between the *predicted value of observed outcomes* and *predicted values of counterfactual outcomes*:

$$
\begin{aligned}
\hat{\delta}_{it}^{\overline{z}} &= \underbrace{\hat{f}_i^{\alpha *}(x_{it}, z_t)}_{\substack{\text{Prediction based on} \\ \text{observed treatment}}} - \underbrace{\hat{f}_i^{\alpha *}(x_{it}, z_t = \overline{z}_t)}_{\substack{\text{Prediction based on} \\ \text{counterfactual treatment}}} \\
&= y_{it} - \xi(x_{it}, h_{it}) - \epsilon_{it} - \left[ y_{it}^{\overline{z}} - \xi(x_{it}, h_{it}) - \epsilon_{it}^{\overline{z}} \right] \\
&= y_{it} - y_{it}^{\overline{z}} + \phi_{it} ,
\end{aligned} \tag{8}
$$

where $\phi_{it} := \epsilon_{it}^{\overline{z}} - \epsilon_{it}$ is random noise with mean zero. As we only change the treatment variable and as observed and unobserved variables are independent of the treatment (Assumptions 1 and 2), Assumption 3 allows us to eliminate the prediction bias and the impact of unobserved variables in the estimation and , therefore, identification of the treatment effect. In that, it has an analogy to the parallel trend assumption in a DiD setting.[5]

A potential concern for Assumption 3 and therefore the estimation of $\hat{\delta}_{it}^{\overline{z}}$ is the quality of predictions based on *unobserved* counterfactual values. To ensure a valid prediction, two additional assumptions regarding the data need to be satisfied.

First, although ML algorithms are designed to produce reliable out-of-sample predictions, they only locally approximate the true model in the range of observed treatments and covariates. It is thus unclear how the estimated functions behave for covariate and treatment combinations which lie outside of the range of observed combinations. To rule these cases out, we need the *positivity* or *covariate overlap* assumption (Samii, Paler and Daly, 2016):

ASSUMPTION 4: *Each combination of the counterfactual treatment $z$ and covariate level $X$ has been observed, i.e.,*

$$
Pr[z|X] > 0. \tag{9}
$$

While it is highly unlikely that all combinations of $z$ and $X$ have been observed, Assumption 4 requires that these combinations should lie within the range of observed data.

The last assumption is concerned with the variation in the level of treatment and controls which is needed to estimate a valid predictor of the underlying structural

[5] Assumption 3 is one of the main differences to the method used by Burlig et al. (2017) and Cicala (2017). They assume that "*...treated and untreated schools to be trending similar on prediction errors...*" (Burlig et al., 2017, pp. 18) or, likewise, "*Parallel trends in unobservalbes...*" (Cicala, 2017, Assumption 2. p. 23) in the sense that the "*contemporaneous error*", i.e., the prediction error, is behaving similar across regions. Under these assumptions, they are able to differentiated out the prediction error and impact of unobservables using control groups in DiD approach. As we lack the control group, we need to assume that the prediction error is independent of the treatment to differentiate out the impact of unobservables and systematic prediction errors.

process:

ASSUMPTION 5: *The variation in the level of treatment and controls over time is sufficiently large.*

Assumption 5 implies that the impact of a change in the treatment on the outcome can be predicted. For many policy interventions, however, treatments are discrete and do not change often which causes problems for estimating the predictor. A possible remedy is to find a control variable that impacts outcome through the same causal mechanism as the treatment variable. In fact, changing such a control variable implies the same change in outcome as a change in the the treatment variable itself. For example, an archetypical problem in economics is to estimate the impact of imposing an input tax (e. g., a carbon tax). Here, the tax change may be a one-time event or it may comprise a few discrete tax changes. The impact of the tax on input costs follows, however, the same mechanism as a change in input prices. It is then possible to use the variation in input prices to identify the causal mechanism of the input tax—with the consequence that the counterfactual simulation then corresponds to a change in input costs directly implied by the tax change.

## II. Application of the Framework: How Effective Was the UK Carbon Tax?

We apply the proposed framework to assess the impact of the UK carbon tax on carbon abatement. In Sections II.A and II.B we provide information about the policy background and draw on economic theory to derive the causal model (Step 1). Section II.D describes our empirical framework to estimate the treatment effect (Steps 2 and 3). Section II.E scrutinizes the validity of our identifying assumptions given the context of our empirical application.

### A. *The Policy Intervention and Confounding Factors*

The main policy instrument of the UK government to decarbonize the heavily fossil-based UK electricity sector is the *Carbon Price Support* (CPS), an annual constant tax on fossil fuel use in the wholesale electricity market (Department of Energy & Climate Change, 2016). The CPS intends to close the gap between an envisaged minimum carbon price, the so-called *Carbon Price Floor* (CPF) and the price of European Emission Allowances (EUA) traded under the European Emissions Trading System (ETS).[6] Table 1 shows the evolution of the EUA, CPS, and the total carbon price over time. Since the introduction of the CPS in 2013, the CPF always exceeded the EUA price, thus resulting in a positive CPS. In 2013, the modest level of the CPS led to a more than two-fold increase of the total

[6] Prior to the introduction of the CPS, the CPS level was conceptualized to be determined two years in advance as the difference between the EUA future price and the CPF. In 2013, the CPF was announced to increase up to 34.5 (69) €/t$CO_2$ in 2020 (2030). At the end of 2015, however, the UK government fixed the CPS rate to 21.6 €/t$CO_2$ until 2021 (Hirst, 2017). In the 2017 budget, the UK government expressed its confidence that "the Total Carbon Price, currently created by the combination of the EU Emission Trading System and the Carbon Price Support, is set at the right level [. . . ]" (HM Treasury, 2017, Article 3.46), thus indicating that the CPS is likely to stay at its current level in future years.

TABLE 1. Descriptive statistics of UK electricity market: carbon prices, generation and import capacity, fuel prices, output, and demand.

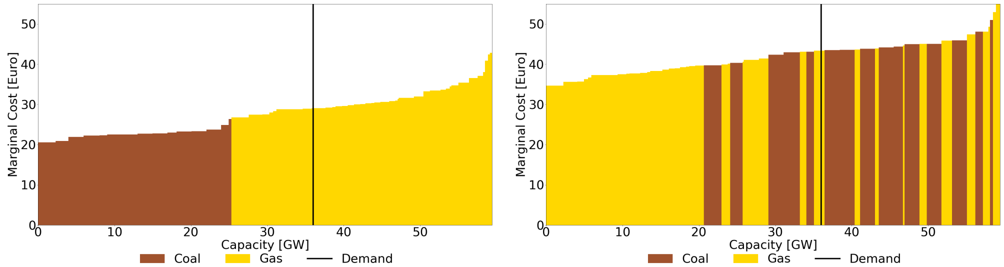
| | Year | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
| *Carbon prices [€ per ton of $CO_2$]*[a] | | | | | | | | |
| EUA | 13.23 | 14.36 | 13.02 | 7.37 | 4.76 | 6.22 | 7.34 | 5.26 |
| CPS (€ per ton) | – | – | – | – | 5.85 | 12.17 | 24.70 | 21.60 |
| Total carbon price (=EUA+CPS) | 13.23 | 14.36 | 13.02 | 7.37 | 10.61 | 18.39 | 32.04 | 26.86 |
| *Capacities [GW]* | | | | | | | | |
| Coal | 25.3 | 25.3 | 25.3 | 24.5 | 19.9 | 18.8 | 19 | 13.8 |
| Gas | 27.3 | 29.5 | 30.2 | 30.3 | 29.3 | 27.4 | 26.6 | 26.1 |
| Import | 2.5 | 2.5 | 3.5 | 3.6 | 4.0 | 4.0 | 4.0 | 4.0 |
| *Fuel prices [€ per MWh thermal energy]* | | | | | | | | |
| Coal | 7.60 | 10.46 | 13.20 | 10.90 | 9.28 | 8.55 | 7.70 | 8.12 |
| | (0.74) | (1.55) | (0.45) | (0.68) | (0.54) | (0.35) | (0.56) | (2.27) |
| Gas | 11.82 | 16.84 | 22.17 | 25.07 | 27.34 | 21.16 | 20.03 | 14.38 |
| | (4.47) | (3.53) | (1.31) | (2.01) | (2.79) | (3.29) | (2.19) | (2.53) |
| Ratio[b] | 0.89 | 0.79 | 0.71 | 0.51 | 0.43 | 0.59 | 0.69 | 0.88 |
| | (0.19) | (0.07) | (0.05) | (0.06) | (0.04) | (0.09) | (0.08) | (0.08) |
| *Hourly demand and generation [GWh]* | | | | | | | | |
| Demand | 27.10 | 28.33 | 25.81 | 24.99 | 23.77 | 22.16 | 20.01 | 19.54 |
| | (6.51) | (6.58) | (6.63) | (6.77) | (6.93) | (6.23) | (6.36) | (6.43) |
| Gas generation | 17.14 | 18.29 | 14.56 | 9.50 | 9.17 | 9.81 | 9.47 | 14.23 |
| | (3.01) | (3.07) | (3.79) | (4.16) | (5.12) | (4.87) | (4.43) | (4.75) |
| Coal generation | 9.81 | 9.97 | 10.70 | 14.35 | 13.11 | 10.13 | 8.17 | 3.27 |
| | (5.80) | (5.29) | (5.14) | (4.04) | (3.18) | (4.10) | (3.45) | (2.88) |

*Notes:* Standard deviations in parentheses. CPS taken from Hirst (2017) and HM Revenue & Customs (2014) converted with exchange rate data from ECB (2017). Daily European Emission Allowances (EUA) spot prices taken from EEX (2017). Further detail about data sources and calculations is provided in Section II.D. [a]As the CPS is adjusted in April of every year, the annual EUA and CPS carbon prices for the years 2013-2016 are calculated based on the period from April to March of the subsequent year. [b]Coal-to-gas fuel price ratio, inclusive of EUA and CPS carbon prices, calculated according to equation (16).

carbon price for the UK electricity industry. In 2016, the CPS was set at the level of €21.60, six times higher than the annual EUA price in this year.

To develop some first intuition for the impacts of the CPS on electricity supply and emissions, Figure 1 plots the short-run supply curve (i.e., ordering marginal cost of fossil-based power plants from low to high) for two situations:[7] a hypothetical situation without the CPS where marginal emissions are only priced at the costs of an EUA (Panel a) and the observed situation with the CPS (Panel b). We observe two main changes. First, the supply curve shifts upward—indicating the increase in the marginal cost of all fossil plants. Second, as natural gas-fired power plants are less carbon-intensive, they are less affected by the carbon price increase and, therefore, become relatively cheaper. Gas plants are thus dispatched into the market and replace emissions-intensive coal-fired plants, in turn reducing emissions. Consistent with this basic mechanism, Figure 2 (Panel (a))shows that

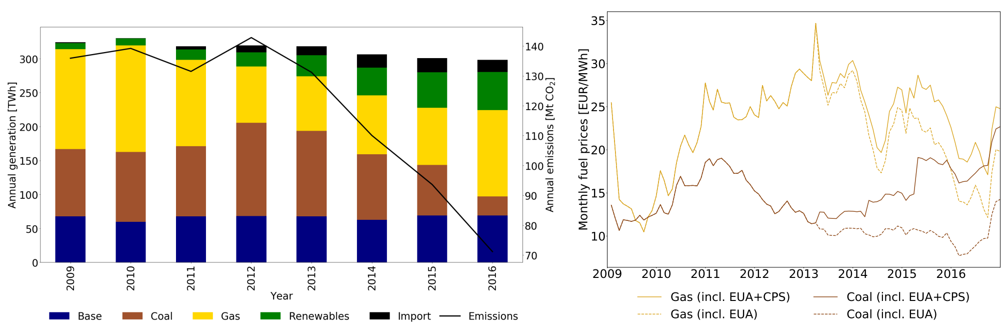
[7]The illustrative calculation shown in the figure is based on one particular hour and assuming average heat efficiencies for plants; it ignores the fact that heat efficiencies, and hence the impact of CPS on individual plants, varies over time depending on temperature and other factors.




(a) Based on hypothetical situation without CPS (b) Based on observational data with CPS

FIGURE 1. Illustrative impact of the UK carbon tax on the short-run market supply curve for electricity

*Notes:* The graph shows the merit order curve of fossil-based power plants on December 19, 2016, at 5:00 p.m. based on the data described in Section II.D. Hydro, nuclear, and renewable power plants are omitted and their total generation is subtracted from demand as they are always dispatched first given that their marginal cost are smaller than those of fossil-based plants. Marginal costs are calculated according to equation 22.




(a) Annual electricity generation by technology and $CO_2$ emissions over time (b) Monthly coal and gas prices with and without CPS

FIGURE 2. Generation, emissions and fuel prices

*Notes:* Own calculations. Electricity generation by fuel is based on ELEXON (2016). "*Base*" comprises electricity generated from hydro and nuclear power plants. "*Renewables*" comprises wind, solar, and other (mainly biomass) generation where generation from wind and solar is corrected for generation embedded in final demand (Nationalgrid, 2016). "*Emissions*" refer to reported values from the EU Transaction Log (European Commission, 2016). Fuel prices for coal and natural gas are taken from EIKON (2007). CPS rates are reported by Hirst (2017) and HM Revenue & Customs (2014), and the EUA price by EEX (2017). Carbon price inclusive fuel prices refer to MWh of *thermal energy*.

starting with the introduction of the CPS in 2013 the annual market share of coal-fired generation sharply decreased while the share of gas-fired plants increased; over the same period, UK's electricity-sector emissions sharply declined.

While Figures 1 and 2 provide some first evidence that the CPS may have led to a reduction in electricity-sector $CO_2$ emissions, there is a host of other factors which are likely to have affected the observed market outcomes.

First, the fraction of electricity demand to be covered with domestic fossil-based generation from coal and natural gas has declined between 2013-2016. This is due to, at least, three factors: *(i)* negative macro-economic shocks and energy efficiency improvements; *(ii)* targeted support policies have likely pushed in zero (or low) marginal-cost generation from renewable energy whenever the underlying natural resource (wind or solar) was available; and *(iii)* UK's electricity imports have slightly increased likely due to both an expansion of newly built inter-connector lines (see Table 1) and the fact that the CPS has increased the domestic cost of generation relative to import prices.

Second, the switch from coal to natural gas was likely also triggered by substantial changes in relative fuel price. Between 2013-2016, natural gas prices declined by nearly 50 percent while coal prices remained largely constant (see Figure 2 Panel (b) and Table 1). This suggests that even without the introduction of the CPS there may have been a marked shift towards gas-fired generation in the UK electricity market.

Third, the decisions to shut down coal-fired plants, reflected in the available production capacity for coal (see Table 1), are likely influenced by factors which are unrelated to the CPS. A main reason for these closures is the European "Large Combustion Plant Directive" (LCPD), which sets specific limits on local pollutant emissions for power plants constructed after the year 1987. The LCPD left electricity firms essentially the choice to either comply with the emissions limits or to "opt out" in which case a maximum operation time of 20'000 hours was granted until the end of 2015 when eventually the plant had to be shut down (European Commission, 2001).

In summary, there is ample evidence that the decline in coal generation and $CO_2$ emissions in the UK power sector which has occurred since the introduction of the CPS in 2013 has likely been the result of a multitude of factors comprising market developments (international fuel prices and electricity demand) and a variety of different policy measures (renewable energy support policies, transmission infrastructure measures, and the CPS). We next present our empirical framework we use to identify and measure the market impacts brought about by the CPS alone.

### *B. Determinants of Wholesale Electricity Market Activity*

We apply microeconomic theory based on a dispatch and peak-load pricing model of the wholesale electricity market (Boiteux, 1960) to pre-select the variables which are potentially important in determining wholesale market outcomes and how they respond to a carbon tax policy. In a subsequent step, these pre-selected variables then enter the ML algorithm to estimate the empirical prediction model which we use for counterfactual analysis to determine the treatment effect of the UK CPS.

COMPETITION IN UK'S WHOLESALE ELECTRICITY MARKET.—The UK wholesale electricity market is a liberalized market based on exchange and over-the-counter trades. In power exchanges, market participants can trade forward and real-time contracts.[8] In the day-ahead market, market participants trade electricity for each hour of the next day. Given the new information in the market, these trades can be revised using the intra-day market which closes one hour before delivery time.

[8]Real-time trading of UK electricity mainly takes place in the EPEX-Spot and Nordpool power exchanges. Forward contracts are traded via the InterContinental Exchange (ICE) and NASDAQ.

In 2014 the UK regulator asked for an investigation of anti-competitive behavior in the UK energy market. In its final report, the "*Competition and Markets Authority*" (CMA, 2016) did not find evidence for anti-competitive behavior in the wholesale electricity market.

A SHORT-RUN EQUILIBRIUM MODEL OF WHOLESALE MARKET ACTIVITY.—We conceptualize the UK wholesale electricity market as being composed of firms which are assumed to operate under perfect competition maximizing profits using production quantities as the decision variable. Generation units of a firm are represented at the plant level where total production of plant $i \in I$ in hour $t \in T$ is denoted by $X_{it}$. The set $I$ comprises thermal carbon-based generation plants (i.e., hard coal, lignite coal, natural gas) and other conventional plants (i.e., nuclear, hydro, pump storage, biomass). Generation from wind and solar is modeled exogenously. Production at any point in time cannot exceed the given effective production capacity $K_{it}$:

$$K_{it} \geq X_{it} \quad \perp \quad \mu_{it} \geq 0 \quad \forall i, t \tag{10}$$

where the time-dependency of capacity mainly reflects maintenance and unscheduled plant outages. $\mu_{it}$ is the shadow price of capacity for technology $i$ at time $t$. The value of capacity in a given hour is zero ($\mu_{it} = 0$) if production is below the capacity limit; it is positive ($\mu_{it} > 0$) if the capacity constraint is binding.[9]

Marginal cost $c_{it}(\boldsymbol{\vartheta_{it}})$ of a generation unit at time $t$ depend on exogenous factors

$$\boldsymbol{\vartheta_{it}} = \{p_t^f, \theta^f, \eta_{it}, p_t^{EUA}, p_t^{CPS}\}$$

comprising the time-dependent price of the fuel $f$ used for electricity generation ($p_t^f$), the carbon content ($\theta^f$), the time-varying EUA and CPS prices on $CO_2$ emissions ($p_t^{EUA}$ and $p_t^{CPS}$), and time-specific heat efficiency ($\eta_{it}$) reflecting ambient temperature ($temp_t$) and potential efficiency losses due to part-load operation.

In equilibrium, the following zero-profit condition, relating unit costs (comprising marginal costs and the opportunity costs for capacity) to unit revenues determines the output of generation unit $i$, $y_{it}$:

$$c_{it}(\boldsymbol{\vartheta_{it}}) + \mu_{it} \geq P_t \quad \perp \quad y_{it} \geq 0 \quad \forall i, t \tag{11}$$

where $P_t$ measures unit profits or the wholesale electricity price at time $t$.[10] If unit cost exceed unit profit, positive generation would lead to losses and thus $y_{it} = 0$. Given perfect competition and no barriers for market entry or exit, zero profits in equilibrium (i.e., unit cost equal to unit profit) determine a positive level of electricity supply $y_{it} > 0$.

The market for electricity in a given hour balances if total supply is equal to

[9] We use the "$\perp$" operator to indicate complementarity between equilibrium conditions and variables. A characteristic of economic equilibrium models is that they can be cast as a complementarity problem, i.e. given a function $F: \mathbb{R}^n \longrightarrow \mathbb{R}^n$, find $z \in \mathbb{R}^n$ such that $F(z) \geq 0$, $z \geq 0$, and $z^T F(z) = 0$, or, in short-hand notation, $F(z) \geq 0 \perp z \geq 0$ (Mathiesen, 1985; Rutherford, 1995).

[10] Equation (11) determines the price as the marginal cost of the marginal generator, i.e. the generation that earns zero capacity rent in the given hour ($\mu_{it} = 0$).

hourly demand $D_t$ which, given our short-run analysis, we assume to be given and price-inelastic:

$$\sum_i y_{it} = D_t \quad \perp \quad P_t \text{ “free”} \quad \forall t\,. \tag{12}$$

Equations (10)–(12) imply that given demand the equilibrium allocation of hourly electricity supplies is determined by the available capacity and the marginal cost ordering of technologies. The equilibrium outcome of each plant $i$ thus depends on demand, and its own as well as the marginal cost and available capacities of all other plants (indicated by $-i$):

$$y^*_{it} = y^*_{it}\left(D_t, c_{it}(\boldsymbol{\vartheta_{it}}), K_{it}, c_{(-i)t}\left(\boldsymbol{\vartheta_{-it}}\right), K_{(-i)t}\right), \forall i, t\,. \tag{13}$$

Equation (13) identifies the major determinants of power plants' outputs (including their responses to a carbon tax policy) by modelling wholesale market activity based on first principles of producer behavior and equilibrium-based market interactions.

Equation (13) should not be viewed as portraying the output decision of a single plant: the dependence of the equilibrium quantities on own and other generators' marginal cost and demand can also be understood in terms of bid functions on the market level. Under perfect competition each generator bids the whole capacity at marginal cost into the market. The market operator then chooses the cheapest bids until demand is fulfilled. Consequently, the acceptance of a bid depends on the ordering of marginal cost in the entire market as well as available capacities and demand.

### *C. Data Sources and Construction.*

To empirically operationalize and embed the model of equation (13) in a framework which enables estimating the treatment effect of the UK CPS on power plants' output decisions, we need measurements for the empirical counterparts of all RHS variables in (13). We thus use a rich and publicly available panel data of hourly generation for each UK fossil-fuel power plant in the period from 2009-2016. In addition, we include data on available hourly capacity, technical characteristics of each plant, non-fossil generation, demand, daily fuel and carbon prices, and weather data.

HOURLY OUTPUT BY PLANT ($y_{it}$).—We use “final physical notification” (FPN) data provided by the operator of the UK electricity balancing system (ELEXON, 2016) as the hourly generation of each fossil power plant unit for the whole sample period. FPN reports the final, 5 minutes before delivery time generation announcement of power plant owners to the grid operator. Although the grid operator might adjust this announcement due to the need for balancing power or re-dispatching measures, these data can be viewed as a reasonable measures for generation (which is not directly observable for UK power plants). As the data on carbon emissions are only available at a plant level, we aggregate power plant units to power plants for

Table 2. Power plant characteristics.

| Plant | Installed capacity [MW] | Average heat efficiency $\eta_i$ [–] | Emissions rate $e_i$ [ton of $CO_2$/MWh] | Opening/ closing date[a] |
|---|---|---|---|---|
| *Natural gas plants* | | | | |
| Pembroke | 2269 | 0.60 | 0.34 | end 2012/– |
| Peterhead | 2134 | 0.55 | 0.36 | –/March 2014 |
| Staythorpe | 1792 | 0.58 | 0.34 | 2010/– |
| Didcot CCGT | 1404 | 0.55 | 0.36 | –/– |
| Connahs Quay | 1380 | 0.48 | 0.42 | –/– |
| West Burton CCGT | 1332 | 0.51 | 0.40 | –/– |
| Grain CHP | 1305 | 0.56 | 0.36 | –/– |
| South Humber | 1239 | 0.50 | 0.40 | –/– |
| Seabank | 1169 | 0.55 | 0.36 | –/– |
| Saltend South | 1164 | 0.52 | 0.38 | –/– |
| Teesside | 1155 | 0.45 | 0.44 | –/Feb. 2013 |
| Immingham CHP | 1123 | 0.44 | 0.46 | –/– |
| Barking | 945 | 0.46 | 0.44 | –/Dec. 2012 |
| Langage | 905 | 0.55 | 0.37 | –/– |
| Marchwood | 898 | 0.58 | 0.34 | –/– |
| Killingholme | 854 | 0.48 | 0.42 | –/March 2015 |
| Severn | 850 | 0.54 | 0.37 | –/– |
| Spalding | 830 | 0.54 | 0.37 | –/– |
| Rocksavage | 800 | 0.53 | 0.38 | –/– |
| Sutton Bridge | 796 | 0.52 | 0.39 | –/– |
| Damhead Creek | 783 | 0.53 | 0.38 | –/– |
| Coryton | 770 | 0.52 | 0.38 | –/– |
| Little Barford | 740 | 0.54 | 0.37 | –/– |
| Rye House | 715 | 0.43 | 0.46 | –/– |
| Keadby | 700 | 0.47 | 0.42 | –/Feb. 2013 |
| Medway | 680 | 0.53 | 0.38 | –/– |
| Baglan Bay | 520 | 0.57 | 0.35 | –/– |
| Deeside | 498 | 0.47 | 0.42 | Dec. 2011/– |
| Great Yarmouth | 420 | 0.56 | 0.35 | –/– |
| Shoreham | 420 | 0.54 | 0.37 | –/– |
| Enfield Energy | 408 | 0.53 | 0.38 | –/– |
| Corby | 401 | 0.39 | 0.51 | –/Oct. 2015 |
| Cottam CCGT | 395 | 0.55 | 0.36 | -/- |
| Kings Lynn | 325 | 0.52 | 0.39 | -/March 2012 |
| Peterborough | 316 | 0.37 | 0.54 | -/Dec. 2011 |
| Average natural gas plant[b] | | 0.51 | 0.40 | |
| *Coal plants* | | | | |
| Longannet | 2304 | 0.42 | 0.81 | –/March 2016 |
| Didcot COAL | 2108 | 0.39 | 0.88 | -/March 2013 |
| Cottam | 2000 | 0.39 | 0.86 | –/– |
| Ratcliffe | 2000 | 0.38 | 0.89 | –/– |
| West Burton COAL | 1972 | 0.38 | 0.90 | –/– |
| Fiddlers Ferry | 1961 | 0.37 | 0.92 | –/March 2016 |
| Ferrybridge | 1960 | 0.38 | 0.89 | –/March 2016 |
| Drax COAL | 1947 | 0.38 | 0.90 | –/– |
| Kingsnorth | 1940 | 0.36 | 0.94 | –/Dec. 2012 |
| Eggborough | 1932 | 0.37 | 0.92 | –/– |
| Aberthaw | 1641 | 0.41 | 0.82 | –/– |
| Cockenzie | 1200 | 0.38 | 0.91 | –/March 2013 |
| Rugeley | 996 | 0.39 | 0.88 | –/June 2016 |
| Ironbridge | 964 | 0.35 | 0.98 | –/March 2012 |
| Uskmouth | 363 | 0.33 | 1.04 | –/– |
| Average coal plant[b] | | 0.38 | 0.89 | |

*Notes:* Installed capacities, fuel type, and plant opening and closure dates are provided by Variable Pitch (2016) and Nationalgrid (2011). For data sources and calculations of heat efficiencies and emission rates see text. [a]"–" indicates that the plants' opening or closure date lies outside of the sample period 2009–2016. [b]Calculated using installed capacities as weights.

our analysis.

FUEL PRICES ($p_t^{fuel}$).—Data on daily fuel prices for coal and natural gas are taken from EIKON (2007). For coal, we use the "ICE CIF ARA Near Month Future". Natural gas prices are "NBP Hub 1st day Futures". All prices are converted to Euro values using daily exchange rates provided by the ECB (2017).

CARBON PRICES ($p_t^{CPS}$ and $p_t^{EUA}$).—CPS rates are reported by Hirst (2017) and HM Revenue & Customs (2014) and the EUA price by EEX (2017). Note that the CPS rate is an annually constant tax in British Pound but reflects exchange rate variations due to conversion to Euro values.

EMISSIONS FACTORS AND PLANT-SPECIFIC HEAT EFFICIENCIES ($\theta^f$ and $\eta_i$).—We take fuel-specific emissions factors from IPCC (2006): 0.34 and 0.20 tons of $CO_2$ per MWh of thermal energy for coal and natural gas, respectively. $CO_2$ emissions for each plant $i$ and year $y$ ($E_{iy}$) are taken from the official registry of the EUTL (European Commission, 2016). Dividing total emissions by total generation per plant, we obtain plant-specific average emissions rates: $e_i = \sum_y E_{iy}/(\sum_t y_{it})$. We then calculate average heat efficiencies for each plant as:

$$\eta_i = \theta^f / e_i \,. \tag{14}$$

Table 2 shows these technical characteristics for each plant in the sample. The average heat efficiency is around 51 percent for natural gas and 38 percent for coal plants. The emission rates, on the other hand, are significantly higher for coal ($0.89\,tCO_2/MWh$) than for gas ($0.40\,tCO_2/MWh$). As we only observe emissions on an annual level, we can only calculate average heat efficiencies. Therefore, hourly changes in heat efficiencies due to, e. g., start-up or ramping constraints, are not considered in our calculations of the emissions impact of the CPS.

AVAILABLE CAPACITY BY PLANT BY HOUR ($\kappa_{it}$).—Installed capacities (shown in Table 2) are provided by Variable Pitch (2016) and Nationalgrid (2011). If observed generation exceeds installed capacity beyond the $95^{th}$ percentile, we set the value of installed capacity equal to the $95^{th}$ percentile of generation.

In addition, data on the maximal output that a plant can provide in a given hour— accounting for permanent and temporary outages due to maintenance or other reasons—the so-called "maximum export limits" (MEL), are provided by ELEXON (2016). Using hourly MEL, we construct a measure of available generation units for each plant. This implies that the availability of the unit is set to zero if MEL is zero; it is set to one otherwise. Summing over all units of a power plant, we obtain a count variable indicating the number of units available per plant, which we use as a proxy for hourly available capacity.

Not all plants in our data run over the entire sample period from 2009–2016 (see Table 2). For years in our sample period during which a plant has been shut down or not yet opened, we set the capacity to zero. In line with this, we also do not predict its counterfactual generation different from zero for these periods, i. e., the impact of the CPS will be zero by assumption. This implies that we assume that plant closures are not caused by the introduction of the CPS. Regarding this assumption it is, however, important to note that the introduction of the

CPS in 2013 coincides with the closure of several coal power plants. While the official reason for the closure is the European "Large Combustion Plant Directive" (LCPD, see Section II.A), we cannot rule out that the shut-down decision for some plants may have also been influenced by the announcement of the CPS. Hence, a fraction of emission reductions due to the LCPD might, in fact, be due to the CPS. However, as our analysis only captures short-run impacts of the CPS on generation, we cannot quantify the long-run impacts of the CPS on capacity. However, we know its direction: While the CPS might have led to a decrease in available coal capacity, it has for sure not led to an increase. As a consequence, our results represent a lower bound estimate of the CPS impact.

DEMAND ($D_t$).—We measure $D_t$ as residual demand, defined as the total output generated by all coal- and natural gas-fired plants using data from ELEXON (2016) on hourly generation aggregated by fuel type.

TEMPERATURE.—We use data on daily temperature provided by ECA&D (2016) to account for time-specific effects on plant-level heat efficiency.

Finally, Table A1 is the Appendix provides descriptive statistics of demand, generation by technology, and imports on an hourly level.

### D. *The Empirical Framework.*

We now turn to the implementation of our conceptual framework established in Section I within the context of the UK CPS. First, we use the information about electricity markets to derive a model of observed outcomes (i.e., generation of each plant), which is invariant to the policy intervention (i.e., the CPS). Second, we use available data and ML algorithms to estimate a predictor of outcomes (i.e., predicted generation). Third, we use the prediction model to estimate the treatment effect as the difference between predicted generation with treatment (i.e. observed CPS) and the unobserved counterfactual (i.e., generation without CPS). Fourth, we present our ex-post calculations to get from the impact on generation to the impact on emissions and abatement cost. Finally, in the subsequent section, we discuss the validity of our approach addressing our four main assumptions.

ESTIMATION EQUATION (STEP 1).—Based on the electricity market model in Section II.B, we know that the equilibrium outcome of each plant $i$ depends on demand $D_t$, and its own as well as the marginal cost and available capacities $K_{it}$ of all other plants. As we do not directly observe plants' (or generators') marginal costs and heat efficiencies ($\eta_{it}$), we exploit the fact that they depend on ambient temperature and thus additionally include daily mean temperature ($temp_t$). The empirical analogue of (13) then becomes:

$$y_{it} = f_i\left[r_t\left(p_t^{coal}, p_t^{gas}, \theta^f, p_t^{EUA}, p_t^{CPS}\right), temp_t, D_t, K_{it}, K_{(-i)t}, \mathbf{\Phi}_t\right] + \epsilon_{it}\,, \tag{15}$$

where we include time fixed effects for each hour of the day and each month of the year ($\mathbf{\Phi}_t$) to account for possible unobserved factors which may impact plant output; and the carbon price inclusive ratio of relative fuel prices:

$$r_t := \frac{p_t^{coal} + \theta^{coal}\left(p_t^{EUA} + p_t^{CPS}\right)}{p_t^{gas} + \theta^{gas}\left(p_t^{EUA} + p_t^{CPS}\right)}\,. \tag{16}$$

While we are interested in the impact of the CPS on plants' output decisions, there is not sufficient variation in the treatment variable ($p_t^{CPS}$) as the CPS changes only in annual steps. As the CPS directly impacts the fuel costs for coal and natural gas, we can, however, exploit the variation in carbon-inclusive fuel prices—instead of including fuel prices ($p_t^{coal}$ and $p_t^{gas}$) and carbon prices ($p_t^{CPS}$ and $p_t^{EUA}$) separately. The implicit assumption here is that a change in fuel prices has the same impact on plants' marginal cost and, hence, output as a change in the carbon price (taking into account the emissions factor of the respective fuel $\theta^{fuel}$. Moreover, using $r_t$ in (16) nicely concords with the view that it is not the absolute but the relative fuel prices determining which plants exit or stay in the market.

MACHINE LEARNING ALGORITHM (STEP 2).—While we know from the theoretical electricity model in (13) and its empirical counterpart in (15) which variables affects plants' output decisions, we do not know the functional form of $f_i$. To obtain an estimator $\hat{f}_i$ of the function $f_i$, we therefore apply ML algorithms, which allow for flexibel functional forms, to produce reliable out-of-sample predictions of each plants' output, $y_{it}$.

We employ the LASSO[11] algorithm (Tibshirani, 1996)—a penalized linear regression model—and use k-fold cross-validation dividing the sample into eight groups (often called folds) to train a prediction model $\hat{f}_i^{\alpha^*}$ for each plant individually.[12] Each prediction model consists of the set of coefficients $\hat{\beta}^{\alpha^*}$ and the optimal regularization parameter $\alpha^*$, which lead to the best possible out-of-sample prediction.[13]

ESTIMATING THE IMPACT OF THE CPS.—To simulate plants' outputs that would have occurred in the absence of the UK carbon tax, we set the CPS treatment variable to zero while leaving all other data unchanged. The counterfactual "no-policy" level of the fuel price ratio is given by:

$$\bar{r}_t = r_t\Big|_{p_t^{CPS}=0}\left(p_t^{coal}, p_t^{gas}, \theta^f, p_t^{EUA}, p_t^{CPS}\right)\,. \tag{17}$$

Based on the estimator in equation (8) detailed in Section I, the impact of the CPS on the output decision of each plant $i$ in each hour $t$ can then be calculated

[11]We also used other algorithms such as random forest. However, the LASSO, an algorithm which is linear in coefficients, lead to the most convincing simulation of the electricity market as a whole: While other algorithms failed at implicitly fulfilling the market clearing condition (see equation (12)), the LASSO algorithm was able to meet this condition—although it was not explicitly modelled (see also II.E).

[12]The LASSO algorithm requires a pre-defined set of input features. In addition to the variables which appear on the RHS of (15), we include (i) interaction terms of all these variables with electricity demand, the coal-to-gas price ratio, and temperature, and (ii) second order polynomials of these three variables.

[13]Appendix B assesses the out-of-sample prediction performance of the ML algorithm as compared to standard regression analysis (OLS) for our data set. We find that the ML (i.e, LASSO) algorithm outperforms the OLS model. This supports the broader insight that ML techniques can be beneficially employed to use prediction to construct an unobserved counterfactual.

as:

$$\hat{\delta}_{it}^{CPS} = \hat{y}_{it}^{\text{with CPS}} - \hat{y}_{it}^{\text{without CPS}}, \tag{18}$$

where

$$\hat{y}_{it}^{\text{with CPS}} = \hat{f}_i^{\alpha*}\left(r_t, temp_t, D_t, K_{it}, K_{(-i)t}, \mathbf{\Phi}_{it}\right) \tag{19}$$

$$\hat{y}_{it}^{\text{without CPS}} = \hat{f}_i^{\alpha*}\left(r_t = \overline{r}_t, temp_t, D_t, K_{it}, K_{(-i)t}, \mathbf{\Phi}_{it}\right)\ . \tag{20}$$

As a closed-form solution of standard errors of the prediction is not available for the LASSO regression (see, for example, Tibshirani, 1996), we use bootstrapping to calculate standard errors. We generate a bootstrap sample with the same length as the original data by using random drawings with replacement. We individually bootstrap by year to get the same amount of values from each year, thus ensuring that all years are equally represented in each sample so as to not violate Assumption 4. We use bootstrapping (with sample size N=1000) to estimate the standard errors of $\hat{\delta}_{it}^{CPS}$ (Venables and Ripley, 2002).

MEASURING $CO_2$ EMISSIONS AND ABATEMENT COST.—To calculate electricity-sector emissions (from combustion of coal and natural gas in electricity generation) at time $t$, we aggregate $CO_2$ emissions from all plants operating in the market:

$$E_t^{\text{without CPS}} := \sum_i \underbrace{e_i \hat{y}_{it}^{\text{without CPS}}}_{\substack{\text{Plant-level} \\ \text{emissions}}}$$

where the emissions of plant $i$ are obtained by multiplying output by the plant-specific emissions rate $e_i$ (see Table 2). Given the estimator for the CPS impact on plant-level output ($\hat{\delta}_{it}^{CPS}$), we can calculate the change in electricity-sector emissions impact due to the CPS as follows:

$$\Delta E_t := \sum_i \underbrace{e_i \hat{\delta}_{it}^{CPS}}_{\substack{\text{Policy-induced change in} \\ \text{emissions of plant } i\ (=:\Delta E_{it})}}. \tag{21}$$

Next to its impact on generation and consequently emissions, the CPS also leads to a change in aggregate production costs. For our ex-post calculations, we assume marginal cost to be linear in fuel and carbon prices. Specifically, based on average heat efficiencies (given by equation (14) and shown in Table 2) marginal cost are calculated as

$$c_{it}(\cdot) = \frac{1}{\eta_{it}}\left(p_t^f + \theta^f(p_t^{EUA} + p_t^{CPS})\right)\ . \tag{22}$$

Aggregate production costs are obtained by summing over marginal generation costs of all plants in the market at time $t$:

$$\Psi_t = \sum_i \hat{y}_{it}^{\text{with CPS}} c_{it}(\cdot) - \hat{y}_{it}^{\text{without CPS}} c_{it}(\cdot) \Big|_{p_t^{CPS}=0} .$$

Using the definition of the treatment effect from equation (18) and plant-specific heat efficiency from equation (14), this can be rewritten as follows:

$$\Psi_t = \underbrace{\sum_i \hat{\delta}_{it}^{CPS} \frac{1}{\eta_{it}} \left( p_t^f + \theta^f p_t^{EUA} \right)}_{\substack{=:T_t \\ \text{Technical abatement} \\ \text{cost}}} + \underbrace{\sum_i p_t^{CPS} e_i \ \hat{y}_{it}^{\text{with CPS}}}_{\substack{=:R_t \\ \text{Tax payments} \\ \text{due to CPS}}} . \tag{23}$$

$\Psi_t$ can thus be decomposed into two parts. $T$ reflects the technical abatement costs for the supply side of the market as the CPS affects plant output by re-ordering the supply or merit order curve. In other words, the CPS leads to an increase in (expensive) natural gas, and a decrease in (cheap) coal generation. This results in higher total production cost for the same amount of electricity generation.

$R$ takes into account the costs incurred due to the CPS tax paid on each unit of generated emissions. While $\Psi$ reflects the costs borne by the supply side of the electricity market, this decomposition is useful as the tax payments by electricity firms are typically recycled in a way which does not destroy the value of $R$. If, for example, the tax revenues from the CPS are fully rebated to electricity consumers, the costs of the CPS aggregated over both sides of the markets amount to $T$ only.

### E. *Scrutinizing the Validity of the Approach: Assumptions 1–5*

As established in Section I.B, the validity of our approach to estimate the treatment effect of the policy intervention as the simple difference between predicted outcomes with and without the policy intervention relies on the existence of a causal model, $f_i$, and a set of assumptions. Before scrutinizing the assumptions, two things are worth noting:

First, the function $f_i$ is independent of the treatment, i. e., the level of the carbon price, as we derive it from the generally valid electricity market model in Section II.B.

Second, by modeling output of one plant depending on the characteristics of all other plants in the market, we do not model a single plants' output decision, but rather the decision of the market maker which plants to use. This implies that our output function is independent of the output function of the other plants and, therefore, the stable unit treatment assumption (SUTVA) is fulfilled. If we would model plant's output depending only on its own marginal cost, i. e., the unit's bid function, then the output of one plant would depend on the treatment applied to other plants as it changes marginal cost and, thus, the ordering of the plants in the supply function (see Imbens and Rubin, 2015, for a further discussion of the SUTVA assumption and unit of analysis).

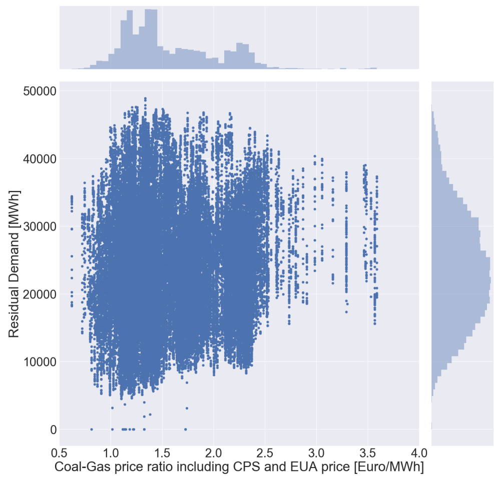
We now turn to the discussion on whether each of the Assumptions 1–5 is met within the specific context of our empirical application. Given that for the discussion of Assumptions 2 and 3 we already need to look at some results, we discuss these assumptions last, right before turing to the results section.

INDEPENDENCE OF CONTROL AND TREATMENT VARIABLES (ASSUMPTION 1)—The validity of our approach to calculate the treatment effect relies on the independence of control and treatment variables as discussed in Section I.B. We believe that these requirements are reasonable for the following reasons: *(i)* the CPS level is determined exogenously at a fixed rate two years in advance; *(ii)* the EUA carbon price is determined by EU ETS market of which the UK electricity sector only covers a negligibly small part; *(iii)* the market share of UK's electricity firms on international fuel markets is not large enough to affect fuel prices; *(iv)* the short-run nature of our analysis means that electricity demand does not react to hourly wholesale electricity prices which may be impacted by a carbon tax; *(v)* likewise, installed capacities cannot be adjusted in the short run and are thus not impacted by a carbon price; and *(vi)* exogenous weather processes determine temperature and are clearly independent of a carbon tax policy.

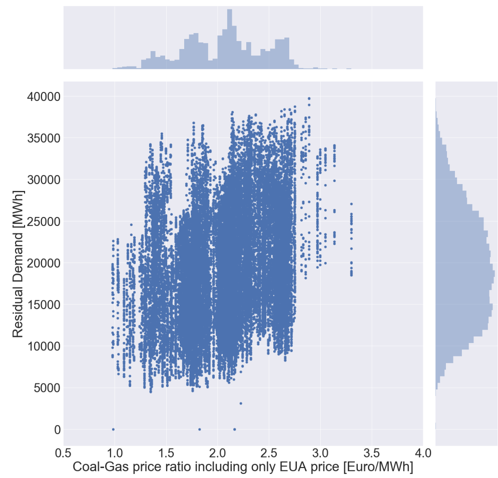
POSITIVITY OR CO-VARIATE OVERLAP (ASSUMPTION 4)—To ensure a high prediction quality of the counterfactual simulation, the positivity assumption requires that the counterfactual fuel price ratio $\overline{r}_{it}$ lies within the range of observed fuel price ratios $r_t$ conditional on observed control variables. Apart from marginal cost, which are expressed through the fuel price ratio, residual demand is the main control variable determining which plant enters the market. The left panel of Figure 3 shows the joint distribution of the observed fuel price ratio and residual demand over the full sample period 2009–2016. The right panel depicts the joint distribution of the counterfactual fuel price ratio and residual demand for the period after the CPS became effective, i.e. from April 2013 until the end of 2016. Comparing both panels, one can see that the imposed counterfactual ratios are well covered by the observed distribution—i.e., only a small fraction of values with a fuel price ratio exceeding a value of 2.5 fall outside the observed sample distribution. We take this as strong evidence that Assumption 4 is satisfied.

VARIATION IN TREATMENT VARIABLE (ASSUMPTION 5)—To train the model and estimate a valid predictor, sufficient variation in treatment and control variables is necessary. While the CPS only varies on an annual level, the distribution of the carbon price inclusive fuel price ratio depicted in Figure 3 shows that there is substantial variation in our modified treatment variable over the sample period.

PREDICTION ERROR INDEPENDENT OF TREATMENT (ASSUMPTION 3)—As we do not observe the no policy counterfactual, we cannot measure the prediction error under no treatment. We can, however, analyze how the observed prediction error evolves over time. Figure 4 uses the the largest coal and gas plants to illustrate our method. Comparing the observed (hollow dots) to the predicted (blue dots) generation under observed covariates, we find that our model predicts observed values rather well. Red dots represent predicted counterfactual generation without the CPS. Our estimator derives the treatment effect as the difference between the predicted value under observed and counterfactual covariates. This can be seen in the figure as




(a) Observed $r_t$ over full sample




(b) Counterfactual $\overline{r}_t$ over CPS period

FIGURE 3. Joint distribution of observed and counterfactual coal-to-gas fuel price ratio ($r_t$ and $\overline{r}_t$, respectively) and residual demand

the distance between blue and red dots. It becomes evident, that the introduction of the CPS led to a decrease in coal and an increase in gas generation.

CONDITIONAL INDEPENDENCE OF UNOBSERVED CONTROLS (ASSUMPTION 2)—As one does not know which unobserved variables may influence plants' output decisions, it is not possible to directly test Assumption 2. By including a number of different fixed-effects specifications, we can, however, test the robustness of our model.

Table 3 reports the impact of the CPS on coal and gas power plant generation from four different model specifications. M1 includes monthly and hourly fixed effects while M2-M4 exclude either monthly or hourly dummies or both.[14] Our finding that the results are robust across model specifications M1-M4 suggests that there do not seem to be significant unobserved variables, with systematic variation at the monthly and/or hourly level, that impact plants' output decisions. We take this as evidence that Assumption 2 is plausibly satisfied in the specific context of our application.

Another important finding of the results shown in Table 3 is, that in all specifications the total net impact of the CPS on generation, i. e., the sum of the impacts on coal and natural gas, does not statistically differ from zero at a 5 % significance level. This implies that the increase in natural gas generation is equal to the decrease in coal generation. This fact is especially noteworthy, as we estimate each plant's output decision separately and do not explicitly impose a structural constraint ensuring market clearing—as depicted by equation (12) in the theoretical model. It is thus not a priori clear that the net impact of the CPS on coal and natural gas is zero. As our empirical model, however, implicitly makes use of all determinants of wholesale market supply and demand, we find that that predicted

[14]Note that we estimate an individual predictor for the generation of each plant rendering it impossible to include unit-specific dummies.

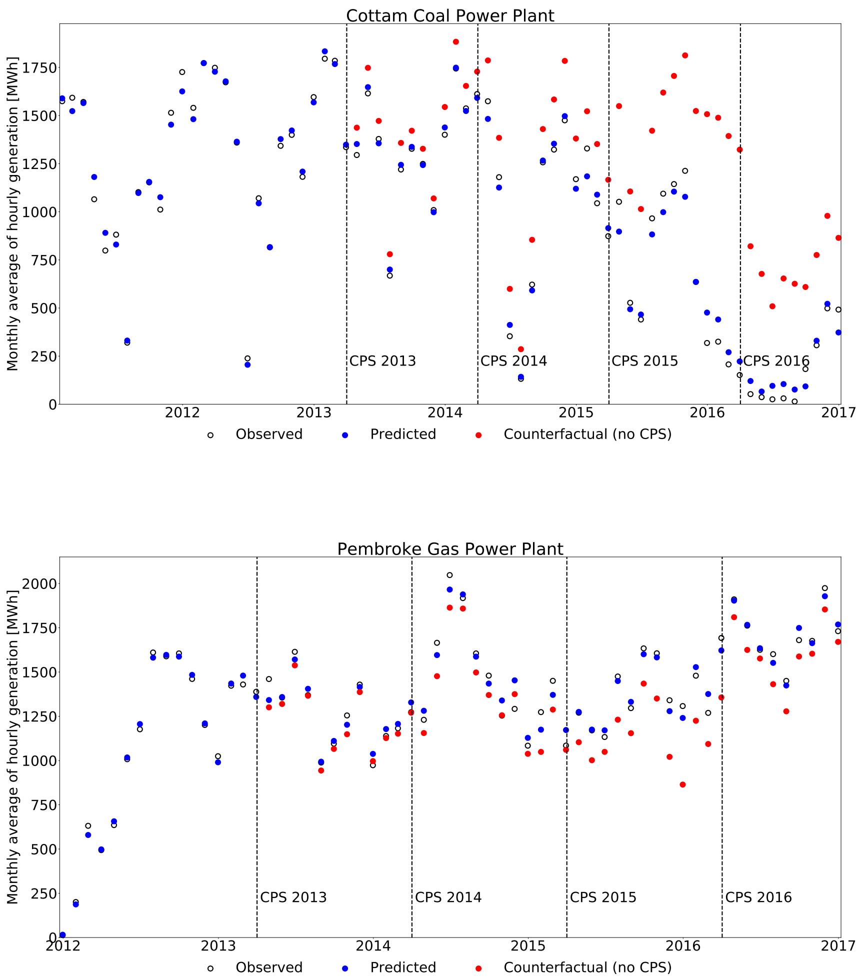



FIGURE 4. Predicted versus observed versus counterfactual values for coal plant (upper row) and gas plant (lower row).

*Notes:* Cottam and Pembroke are the largest coal and gas plants, respectively, which have been in the market during the whole treatment period. In 2013, the CPS was 5.85 €/t, in 2014 it was 12.17 €/t, in 2015 and 2016, 24.70 and 21.60 €/t respectively.

TABLE 3. Assessing unobserved heterogeneity: impact of the UK carbon tax (CPS) on aggregated power plant output by technology category for different model specifications.

| | Model specification | | | |
|---|---|---|---|---|
| | M1 | M2 | M3 | M4 |
| Monthly fixed effects | *yes* | *no* | *yes* | *no* |
| Hourly fixed effects | *yes* | *no* | *no* | *yes* |
| *Coal* | | | | |
| TWh | -46.29 | -42.78 | -43.17 | -42.72 |
| | (1.69) | (1.01) | (1.71) | (1.20) |
| % of total generation[a] | 14.7 | 13.6 | 13.7 | 13.6 |
| *Natural gas* | | | | |
| TWh | 45.55 | 45.00 | 46.01 | 45.23 |
| | (1.06) | (0.92) | (1.07) | (0.75) |
| % of total generation[a] | 15.0 | 14.9 | 15.2 | 14.9 |
| *Total* (TWh) | -0.75 | 2.23 | 2.84 | 2.51 |
| | (2.00) | (1.37) | (2.02) | (1.42) |

*Notes:* Plant-level impacts $\hat{\delta}_{it}^{CPS}$ based on equation (18). [a]Refers to situation without the CPS. Bootstrapped standard errors are shown in parentheses.

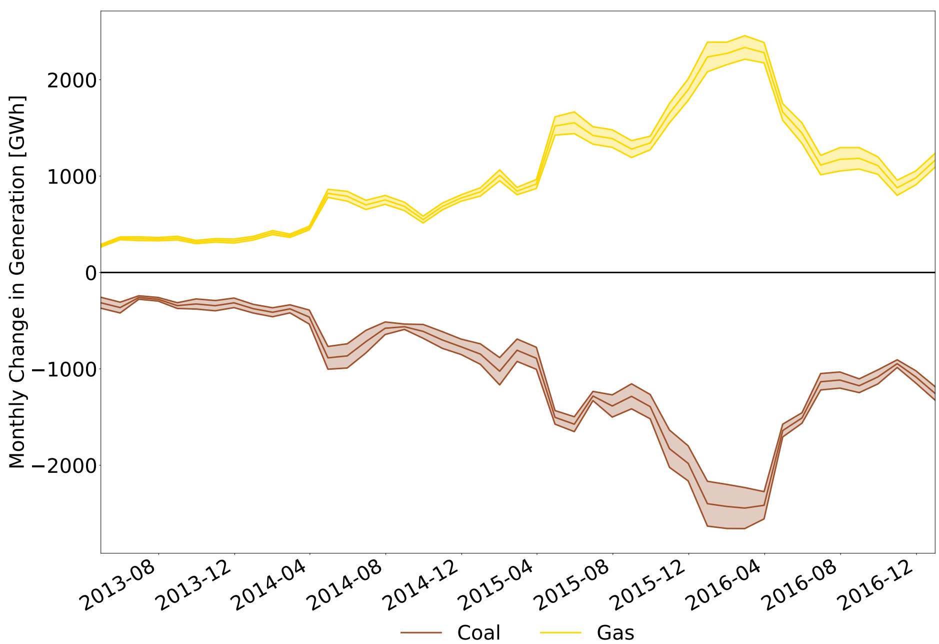



FIGURE 5. Monthly impacts of the UK carbon tax (CPS) on electricity generation

*Notes:* Shaded areas represent 95 % confidence intervals (based on bootstrapped standard errors). Values shown refer to estimated plant-level impacts $\hat{\delta}_{it}^{CPS}$, based on model specification M1 and equation (18), aggregated by technology category and month.

supply matches demand—even though we do not explicitly require market balance. This is not only true for the whole sample period, but can also be shown by the symmetric impacts on coal and natural gas generation on a monthly level (see Figure 5). This implicit market clearing is another evidence, that we are not missing out on important unobserved variables.

Given that under specification M1, which includes hourly and monthly fixed effects, the net impact on coal and gas generation is closest to zero, this is our preferred specification for subsequent analysis.

## III. Results

This section presents our main results. We first report on the aggregated impacts of the CPS on output as well as the emissions and abatement cost impacts. We then scrutinizing the drivers for the heterogeneous impacts over time. Finally, we investigate in detail how both, the environmental effectiveness and abatement costs of the CPS, depend on the prevailing relative market prices for fossil fuels.

### *A. Market impacts of the UK carbon tax*

OUTPUT BY TECHNOLOGY CATEGORY.—Table 4 shows the impact of the CPS on coal and gas generation for each year and the total, i.e. cumulative, impact since its introduction in April 2013 until the end of 2016. We find that, in aggregate over all fossil-based power plants and until the end of 2016, the CPS caused a reduction in the output from coal-fired plants of 46.29 TWh and an increase from gas-fired plants of 45.55 TWh.[15] Relative to a situation without the CPS, these changes correspond to a fuel switch from coal to natural gas of around 15 percent. Plant-specific generation impacts of the CPS are shown in Appendix C.C1.

The impact of the CPS on generation varies substantially over time. The fuel switch was initially low at an absolute level of around 4 TWh in the 2013 period and then increased over the years with the highest value of around 22 TWh in 2015. The absolute impacts for both natural gas and coal are much larger in the 2015 than in the 2016 period (even if we control for the fact that the 2016 period comprises only nine months). In relative terms, coal experienced the largest decrease in the 2016 period. Adding to the heterogeneity in the annually aggregated impacts by CPS period, Figure 5 shows that there is also considerable variation over time in the CPS-induced impacts on monthly output of coal- and gas-fired power plants.

EMISSIONS AND ABATEMENT COSTS.—Table 5 reports on the effects of the CPS on electricity-sector $CO_2$ emissions and abatement cost. We find that in the period 2013–2016, the CPS has reduced emissions by 26.1 million tons—which corresponds to a 6.4 percent reductions of emissions as compared to a situation without a CPS. The highest CPS-induced abatement occurred in the 2015 period with 11.6 million tons or a reduction of 11.6 percent of emissions in that period.

Applying our measure of technical abatement costs $\boldsymbol{T}$ from equation (23), the CPS has reduced one ton of electricity-sector $CO_2$ emissions at an average cost of € 18.2. There is, however, substantial variation in the average technical cost over time, ranging from € 2.7 in 2015 to € 47.5 in 2013 per ton abated $CO_2$.

[15]Note that in none of the CPS periods the net impact on the sum of coal and natural gas does statistically differ from zero at a 5 % significance level.

TABLE 4. Impacts of the UK carbon tax (CPS) on aggregated power plant output by fuel type.

| | Period | | | | Total impact |
|---|---|---|---|---|---|
| | 2013 | 2014 | 2015 | 2016 | 2013-2016 |
| *CPS* [€ per ton of $CO_2$] | 5.85 | 12.17 | 24.70 | 21.60 | – |
| *Output from coal-fired plants* | | | | | |
| TWh | -4.17 | -9.26 | -21.92 | -10.94 | -46.29 |
| | (0.27) | (0.57) | (0.86) | (0.21) | (1.69) |
| % of total generation[a] | -3.7 | -9.8 | -27.0 | -43.6 | -14.7 |
| *Output from gas-fired plants* | | | | | |
| TWh | 4.27 | 9.37 | 21.19 | 10.72 | 45.55 |
| | (0.10) | (0.23) | (0.57) | (0.40) | (1.06) |
| % of total generation[a] | 6.1 | 12.1 | 29.7 | 12.8 | 15.0 |
| *Total* [TWh] | 0.10 | 0.11 | -0.73 | -0.22 | -0.75 |
| | (0.29) | (0.62) | (1.03) | (0.45) | (2.00) |

*Notes:* As the CPS is adjusted in April of every year, all reported variables refer to the period from April to March of the subsequent year. As data is available until December 2016, the 2016 period comprises only nine months. Values shown refer to estimated plant-level impacts $\hat{\delta}_{it}^{CPS}$, based on model specification M1 and equation (18), aggregated by technology category. Bootstrapped standard errors are shown in parentheses. [a]Refers to situation without the CPS.

TAX REVENUE.—Table 5 also reports the tax revenues raised with the CPS. Since the introduction of the CPS until the end of 2016, the British government received around €5.2 billion in tax revenue from the CPS policy. Again, there is temporal heterogeneity in the magnitude of tax revenues collected: the highest tax revenues (around €2 billion) accrued in 2015 when both, emissions and the CPS level, were high; already in the subsequent period, the CPS tax revenue dropped significantly due to the fact that the remaining $CO_2$ emission in the market were considerably lower (controlling for the fact that the 2016 period comprises only nine months).

### *B. Drivers of the heterogeneous market impacts: An overview*

Understanding the drivers and market conditions which determine the effectiveness of a tax policy is crucial for the design of environmental regulation. It is thus important to ask: What explains the heterogeneous impacts over time which, in turn, determine the effectiveness and abatement cost of the UK carbon tax?

In the short-run, the CPS leads to carbon abatement if it induces a replacement of coal-fired by natural gas-fired electricity generation. To this end, the CPS needs to fill the gap between coal and natural gas prices to induce a fuel switch. This can be favored by two main factors: First, a higher level of the CPS. Second, a higher coal-to-gas price ratio, i. e., a smaller the gap between coal and gas prices. A higher fuel-price ratio has, however, two opposing effects. On the one hand, abatement becomes cheaper. On the other, gas-fired generation becomes cheaper even in the absence of carbon policy. It thus becomes more likely that gas plants are already dispatched without any carbon policy and, therefore, the potential for abatement, i. e., the substitution of coal plants, decreases.

To trace out in detail the impacts of these drivers and the interactions between

Table 5. Impacts of the UK carbon tax (CPS) on electricity-sector emissions and abatement cost

| | Period | | | | Total impact |
|---|---|---|---|---|---|
| | 2013 | 2014 | 2015 | 2016 | 2013-2016 |
| *CPS* [€/t] | 5.85 | 12.17 | 24.70 | 21.60 | – |
| *Emissions without CPS* [Mt] | 125.8 | 112.0 | 98.0 | 71.3 | 407.1 |
| *$CO_2$ abatement* | | | | | |
| $\Delta E_t$ [Mt] | 2.1 | 4.7 | 11.6 | 7.6 | 26.1 |
| | (0.25) | (0.53) | (0.81) | (0.24) | (1.60) |
| % of total emissions[a] | 1.7 | 4.2 | 11.9 | 10.7 | 6.4 |
| *Abatement cost* $\Psi_t = T_t + R_t$ | | | | | |
| Technical cost $T_t$ [mio. €] | 101.1 | 129.1 | 195.1 | 20.5 | 445.0 |
| | (9.2) | (18.4) | (29.1) | (16.6) | (58.7) |
| Avg. tech. cost $T_t/\Delta E_t$ [€/t] | 47.5 | 27.2 | 16.8 | 2.7 | 18.2 |
| | (12.5) | (8.7) | (4.0) | (2.3) | (4.0) |
| Tax payments $R_t$ [mio. €] | 725.7 | 1309.6 | 2129.4 | 1372.8 | 5194.3 |

*Notes:* Values shown refer to estimated plant-level impacts $\hat{\delta}_{it}^{CPS}$, based on model specification M1 and equation (18), aggregated by period. As the CPS is adjusted in April of every year, all reported variables refer to the period from April to March of the subsequent year. As data is available until December 2016, we can only estimate the impacts of the CPS for a nine month period. To ensure comparability with previous years, we scale model values for 2016 to a 12-month basis. Bootstrapped standard errors are shown in parentheses. [a] Relative to a situation without CPS.

them, we perform simulations with our ML-trained structural model. Specifically, Figure 6 shows the dependency of hourly abatement (upper panel), total hourly abatement cost (middle panel) and average abatement cost per ton of $CO_2$ (lower panel) with respect to the coal-to-gas price ratio for four different levels of the CPS (5, 10, 20, and 30 € per ton of $CO_2$).[16]

The following insights emerge regarding the two main drivers of heterogeneous market impacts:

level of cps ($p_t^{CPS}$).—We find that for a given coal-to-gas price ratio, the highest CPS level brings about an around six times higher hourly abatement compared to the lowest CPS level. Intuitively, the higher the CPS level, the larger is the increase in the fuel price ratio (for given fuel prices) and, thus, the larger is the impact on the marginal costs and output of fossil-based power plants. Also total technical abatement cost are higher in the case of a high CPS, reflecting two effects. On the one hand, higher abatement naturally implies higher cost. On the other hand, a higher CPS implies higher marginal abatement cost as more coal plants need to be replaced. For CPS levels below 20 €/t $CO_2$ average abatement cost are nearly constant in the fuel price ratio indicating that the difference in total cost is mainly driven by larger abatement. For a level of 30 €/t $CO_2$ the increase in average cost becomes larger indicating a steeper increase in the MAC curve.

coal-to-gas price ratio ($r_t$).—Next to the CPS level, the relative fuel price of coal to natural gas (shown on the x-axis) is a very important driver of CPS impacts on emissions and abatement cost. Abatement shows an inverted U-shaped

[16]Curves are fitted using a least square fit. For abatement (total cost and average cost), a second order polynom (linear function) shows the best fit.

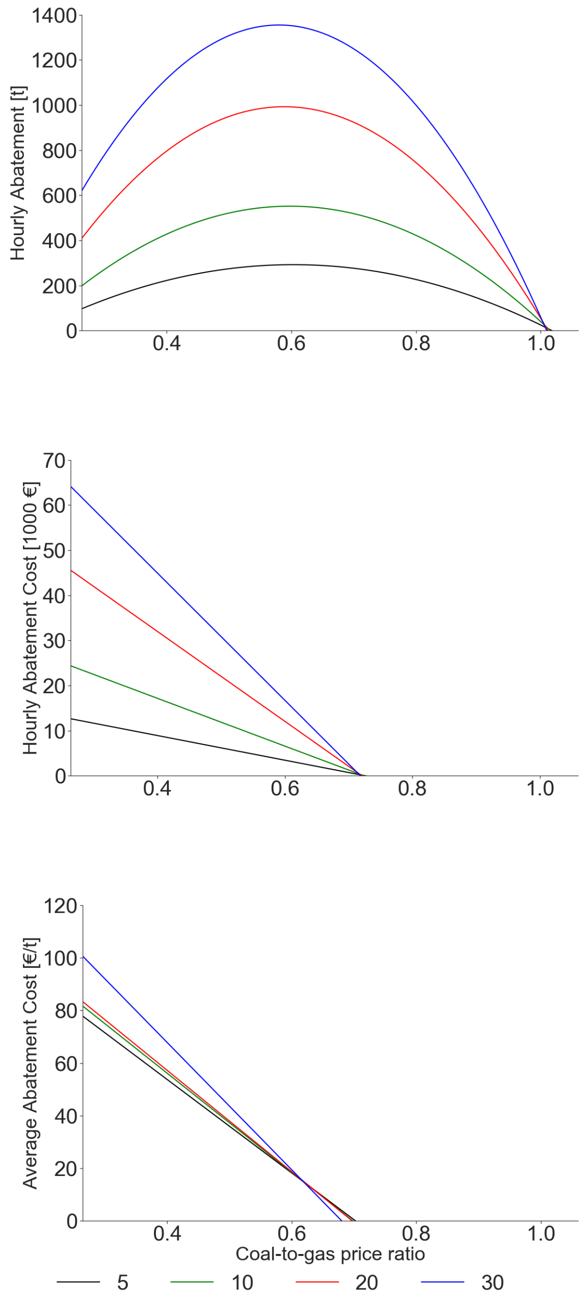



FIGURE 6. Impact of CPS level ($p_t^{CPS}$) and fuel price ratio ($r_t$) on hourly abatement (first row), total abatement cost (middle row), and average abatement cost (bottom row).

*Notes:* Legend refers to CPS levels in €/t.

pattern peaking at a coal-to-gas price ratio of $r_t \approx 0.6$. This is an indication of the two opposing effects of the fuel price ratio. Until a ratio of 0.6, MAC are decreasing and, therefore, abatement increases. Afterwards gas-fired generation becomes favorable even in the absence of the CPS and, thus, the potential for a fuel switch, i. e., abatement, is decreasing. Abatement costs monotonically decrease with an increasing fuel price up to the point where $r_t$ is between 0.6 and 0.8. These numbers are no coincidence: Given the average heat efficiencies of all plants in our sample (see Table 2) and the definition of marginal costs from equation (22), the most efficient gas plant replaces the least efficient coal plant at a fuel price ratio of $r_t = 0.55$. This implies that starting from this ratio, coal and gas prices are roughly equal up to a fuel price ratio of $r_t = 0.88$, where the most efficient gas plant replaces the most efficient coal plant.[17] In the next section, we analyze the impact of the fuel price ratio on abatement and abatement cost in more detail.

### *C. Effectiveness and cost of the carbon tax: zooming in on the role of relative fuel prices*

The effectiveness and cost of the CPS to reduce $CO_2$ emissions majorly depend on the coal-to-gas price ratio $r_t$ which is determined by prevailing market prices for fossil fuels prior to the introduction of the carbon tax. This section further scrutinizes the impacts of the CPS through the fuel price channel on the quantities and costs abated.

Intuitively, an increase in $r_t$ affects the abatement behavior of power plants in two ways. First, given a situation in which coal is cheaper than gas, the carbon price necessary to equalize the after-tax costs of coal and gas is the smaller, the cheaper natural gas becomes relative to coal, i. e., the higher $r_t$. Hence, marginal abatement costs (MAC)—defined as the minimum carbon price which induces a switch between a coal and a natural gas plant—also decrease in $r_t$. We refer to this as the "*MAC effect*". Second, the potential to abate $CO_2$ emissions diminishes as $r_t$ increases. Intuitively, as natural gas gets cheaper relative to coal (i.e., $r_t$ increases), gas-fired plants replace coal plants even without a carbon tax policy, thus implying that the abatement potential (AP) of a carbon price to lower emissions through switching from coal to natural gas—defined as the difference in emissions under the observed market dispatch and the hypothetical emissions in a market situation in which all available gas-fired power plants would be dispatched before coal plants are used—is limited. We refer to this as the "*AP effect*". Note that these two effects work in opposite directions in terms of carbon abatement: as $r_t$ increases, the "*MAC effect*" implies that carbon abatement becomes cheaper, hence increasing abatement, while abatement decreases due to the "*AP effect*".

Figure 7 shows the impact of the (carbon tax-exclusive) fuel price ratio on MAC and the AP.[18] First, MAC are decreasing in $r_t$ as switching from coal to gas plants

[17]The switch is complete at a level of $r_t > 1$ where the least efficient gas plant replaces the most efficient coal plant.

[18]We calculate the MAC and AP for each hour based on observed fuel prices, demand, and heat efficiencies for each plant (see Table 2). MAC are defined as the lowest carbon price inducing a switch between a coal and gas-fired plant. For the AP, we compute a hypothetical market dispatch with all gas-fired plants dispatched first. AP is then defined as the emissions under the realized dispatch under

becomes cheaper when the price difference decreases. Also, as natural gas becomes relatively cheaper, the AP tends to diminish (the dispersion in the AP for a given level of $r_t$ is driven by the considerable variation in electricity demand over the day and the year).

Second, we can identify three distinct regions where the effectiveness and costs of the carbon tax differ systematically depending on the threshold values for $r_t$ where the most efficient gas plant replaces the least (most) efficient coal plant, corresponding to the left and right vertical lines in Figure 7, respectively:

*"Low $r_t$" (high MAC, high AP).*——Coal plants are much cheaper than natural gas plants, implying that the AP is at its maximum as all coal plants are still in the market and can potentially be replaced. The MAC are initially high for lower levels of $r_t$ as the gap between coal and gas prices that has to be overcome by a potential tax in order to induce a fuel switch is high. The MAC then decrease as the coal price increases relative to the gas price.

*"Intermediate $r_t$" (low MAC, high AP).*——When $r_t$ is slightly above the threshold level of 0.55, MAC are low and the AP is high. As $r_t$ increases, gas plants gradually begin to substitute coal plants, implying that the AP decreases. As the production cost of coal and gas plants get closer to one another, a low carbon tax is already sufficient to induce abatement, implying that MAC are low and further decrease.

*"High $r_t$" (low MAC, near-zero AP).*——At a fuel price ratio of 0.88 the most efficient gas plant replaces the most efficient coal plant even in the absence of the CPS. Thus, the natural gas plants becomes the first plant in the merit order curve. As a consequence, gas plants replace nearly all coal capacity in the market, implying that the AP is exhausted. As most gas plants are cheaper than coal-fired plants, MAC are virtually zero. At a fuel price ratio of 1.12 the most inefficient gas plant replaces the most efficient coal plant. As the all gas plants are dispatched before coal, AP is completely exhausted and (short-run) MAC are no longer defined.

Equipped with the intuition from Figure 7, we can finally examine the heterogeneous effects of the CPS in terms of carbon abatement and cost. Figure 8 shows the relation between $r_t$, daily abatement, and average abatement cost as estimated by our ML-trained model. In 2013, the coal-to-gas price ratio was low (stars in the region "Low $r_t$"). Albeit the AP was high, the CPS was too low to trigger substantial abatement due to high MAC. Average abatement cost are rather high again due to the high MAC.

In 2014, coal prices increased relative to gas prices (i. e., circles in the regions "Low $r_t$" and "Intermediate $r_t$") leading to a decrease of MAC. Moreover, the CPS increased. Abatement was therefore higher and average abatement cost decreased.

The largest abatement can be found at a fuel price ratio around 0.55 and a high CPS level, i. e., diamonds. This corresponds to the situation in the year 2015 with low MAC but a high AP. Compared to 2014, the ratio was very similar but due to a higher CPS abatement was higher. Finally, the situation in 2016 is represented by the diamonds in the middle and right part. In this year, both the CPS and $r_t$ were high, i. e., the AP was almost entirely exhausted, implying a lower abatement

fuel prices without the CPS and the hypothetical gas dispatch.

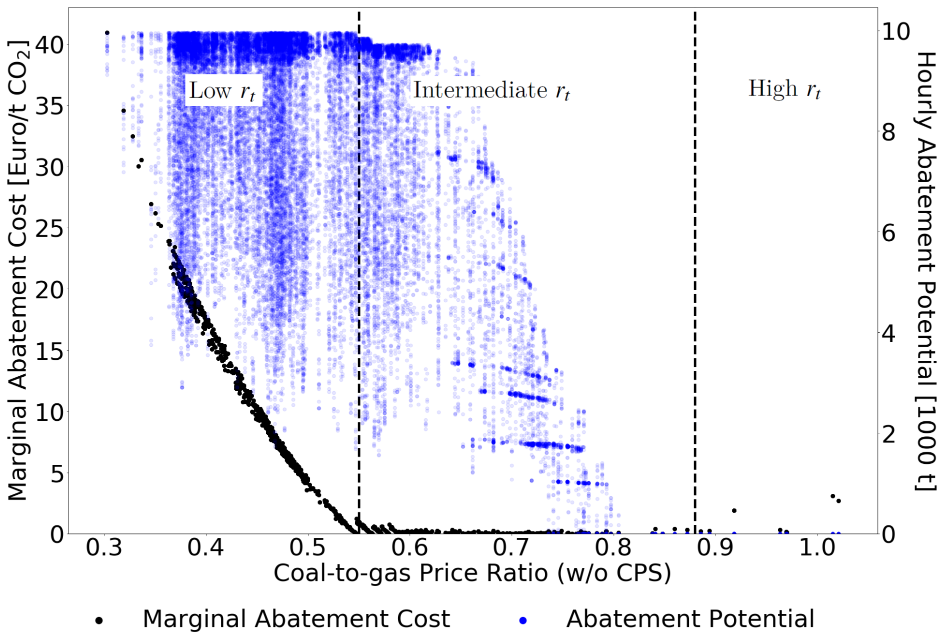



FIGURE 7. Relation between fuel price ratio $r_t$, hourly marginal abatement cost (MAC), and hourly abatement potential (AP)

compared to 2015 but at very low cost. Taken together, Figure 8 clearly shows that the effectiveness of the CPS varied considerably over time depending on the prevailing relative market prices of fossil fuels.

## IV. Concluding Remarks

While economists see carbon pricing (through a tax or an emissions trading system) as the most important regulatory measure for mitigating climate change, surprisingly little is known about its performance in the electricity sector from an ex-post perspective. Causal inference of the market impacts of a broad-based carbon tax, i.e. one which affects all units in the market, is difficult as typically no suitable control group of a counterfactual situation exists.

Against this background, this paper has made two contributions. First, we have developed and implemented a new approach which combines economic theory and machine learning (ML) techniques to establish causal inference of a policy intervention in settings with high-frequency data when no control group exists. Specifically, we exploit economic theory of electricity market dispatch and peak-load pricing to select the variables of a causal model which is then trained using ML to obtain an empirical model for out-of-sample prediction at the firm level. We obtain the treatment effect for each firm as the difference between predicted outcomes with and without policy.

The developed framework rests on several conditions to be applicable. Specifically, we deal with situations in which the underlying structural causal model is constant over time. This allows us to use the full sample to train a predictor

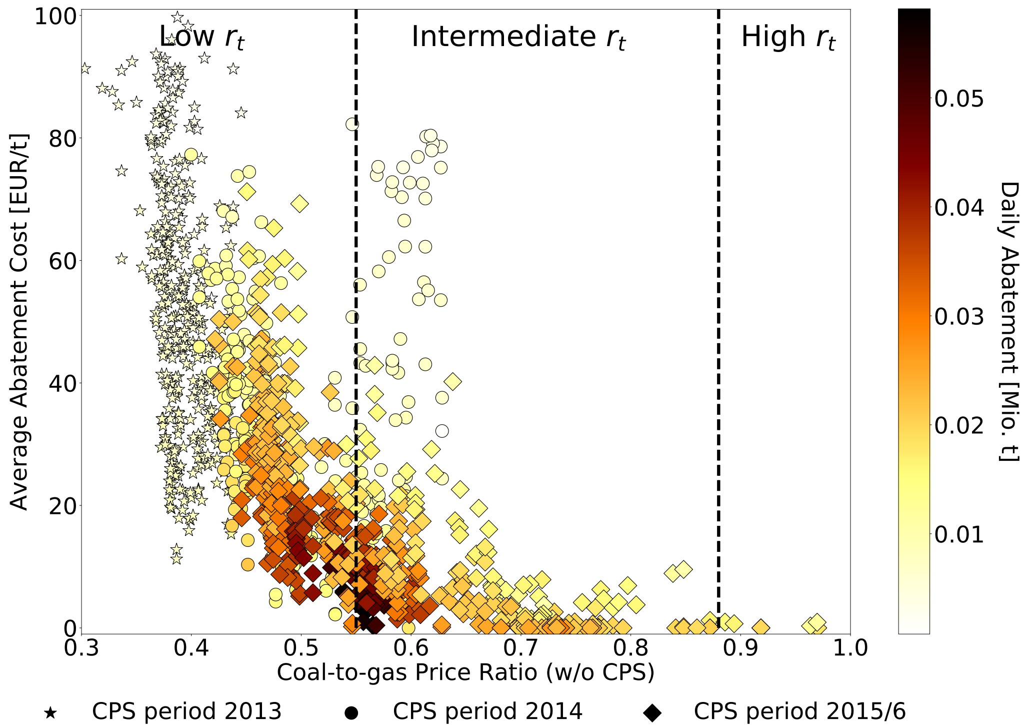



FIGURE 8. Relation between fuel price ratio $r_t$, daily abatement, and daily abatement cost

*Notes:* The color code refers to daily abatement. The marker form represents the level of the CPS corresponding to different years (see also Table 5): the stars refer to 2013 with a a CPS of 6 €/t $CO_2$, the circles refer to 2014 with a CPS of ≈12 €/t $CO_2$, and the diamonds refer to 2015 and 2016 with a CPS in the range of 22–25 €/t $CO_2$.

function ensuring a high prediction quality. Also, the treatment variable has to be variable enough to allow to identify its causal impact in the predictor function. If treatment is not variable enough, as it is the case for the CPS level in our application, one needs to exploit the variation of a control variable with the same causal impact as the treatment variable, in our case relative fuel prices. Furthermore, ML models are local approximations of true functions given the observed sample. This implies that one has to take care that variables used for predictions lie within the support of the sample in order to avoid increasing prediction errors.

Second, employing our new approach, this paper has evaluated the environmental and cost effectiveness of the UK Carbon Price Support—a carbon levy imposed on all fossil-based power plants in the electricity market. To the best of our knowledge, this is the first paper in economics to incorporate ML methods to conduct causal inference of carbon pricing. Our analysis provides empirical evidence for the view that a carbon tax is an effective regulatory instrument to reduce $CO_2$ emissions: the CPS induced a substitution away from "dirty" coal to cleaner natural gas-fired power plants—replacing about 15 percent or 46 TWh of coal-based generation and reducing electricity sector emissions by 6.2 percent between 2013 and 2016. Over that period, we find that the abatement of one ton of $CO_2$ incurred additional total costs of €18.2 for consumers and fossil-based electricity producers. We used simulations with the ML-trained model to identify the empirical conditions which influence the environmental and cost effectiveness of the carbon tax. We find that the ratio of carbon tax-exclusive prices for coal and natural gas is by far the most important driver.

Some limitations of our analysis should be kept in mind. First, we focus on analyzing the short-run market impacts of the CPS. We thus abstract from potential effects of the CPS on energy conservation, i. e., demand reductions, and investments in low-carbon electricity production capacity. Similarly, we do not estimate the impact of the CPS on plant closure—which we assumed to be mainly attributed to the European "Large Combustion Plant Directive". To the extent that such effects increase carbon abatement for a given tax level, our analysis is best viewed as providing a minimum estimate of the environmental effectiveness of the CPS.

Second, by increasing domestic wholesale market prices relative to the costs of electricity imports, the CPS may stimulated imports. This also suggests that our estimate of the domestic $CO_2$ emissions impacts caused by the CPS should be interpreted as providing a lower bound.

Notwithstanding these considerations, our analysis has important implications for the design of carbon pricing policies to mitigate climate change: a carbon tax can but does not necessarily have to be an effective instrument. Its effectiveness in terms of both quantity reductions of $CO_2$ emissions and economic costs significantly depends on the prevailing fuel market conditions.

## REFERENCES


**Aldrich, J.** 1989. "Autonomy." *Oxford Economic Papers*, 41: 15Ű34.

**Angrist, Joshua D., and Joern-Steffen Pischke.** 2008. *Most Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

**Athey, Susan.** 2017. "Beyond Prediction: Using Big data for Policy Problems." *Science*, 335(6324): 483–485.

**Athey, Susan, and Guido Imbens.** 2017. "The State of Applied Econometrics-Causality and Policy Evaluation." *Journal of Economic Perspectives*, 31(2): 3–32.

**Bajari, Patrick, Denis Nekipelov, Stephen P Ryan, and Miaoyu Yang.** 2015. "Machine learning methods for demand estimation." *The American Economic Review: Papers & Proceedings*, 105(5): 481–485.

**Boiteux, M.** 1960. "Peak-Load Pricing." *The Journal of Business*, 33(2): pp. 157–179.

**Bretschger, Lucas, Fillippo Lechthaler, Sebastian Rausch, and Lin Zhang.** 2017. "Knowledge Diffusion, Endogenous Growth, and the Costs of Global Climate Policy." *European Economic Review*, 93: 47–72.

**Burlig, Fiona, Christopher Knittel, David Rapson, Mar Reguant, and Catherine Wolfram.** 2017. "Machine learning from schools about energy efficiency." National Bureau of Economic Research.

**Chevallier, Julien, Erik Delarue, Emeric Lujan, and William D'Haeseleer.** 2012. "A counterfactual simulation exercise of CO2 emissions abatement through fuel-switching in the UK (2008–2012)." *International Journal of Global Energy Issues*, 35(5): 311–331.

**Cicala, Steve.** 2017. "Imperfect Market versus Imperfect Regulation in the U.S. Electricity Generation Generation."

**CMA.** 2016. "Competition and Markets Autority: Energy market investigation."

**Cullen, Joseph A, and Erin T Mansur.** 2017. "Inferring carbon abatement costs in electricity markets: A revealed preference approach using the shale revolution." *American Economic Journal: Economic Policy*, 9(3): 106–33.

**Delarue, Erik D, A Denny Ellerman, and William D D'Haeseleer.** 2010*a*. "Robust MACCs? The topography of abatement by fuel switching in the European power sector." *Energy*, 35(3): 1465–1475.

**Delarue, Erik D, A Denny Ellerman, and William D D'Haeseleer.** 2010*b*. "Short-term CO2 abatement in the European power sector: 2005–2006." *Climate Change Economics*, 1(02): 113–133.

**Delarue, Erik, Kris Voorspools, and William D'Haeseleer.** 2008. "Fuel switching in the electricity sector under the EU ETS: review and prospective." *Journal of Energy Engineering*, 134(2): 40–46.

**Department of Energy & Climate Change.** 2016. "2014 UK Greenhouse Gas Emsissions, Final Figures." *Statistical release.*

**ECA&D.** 2016. "European Climate Assessment & Data. http://eca.knmi.nl/dailydata/predefinedseries.php."

**ECB.** 2017. "European Central Bank: Daily exchange rates. www.ecb.europa.eu/stats."

**EEX.** 2017. "European Energy Exchange: Daily Spot Prices EU Emission Allowances."

**EIKON.** 2007. "Thomson reuters EIKON database."

**ELEXON.** 2016. "ELEXON: www.elexon.co.uk."

**European Commission.** 2001. "Directive 2001/80/EC of the European Parliament and of the Council of 23 October 2001 on the limitation of emissions of certain pollutants into the air from large combustion plant." *Official Journal of the European Community.*

**European Commission.** 2016. "DG Climate Action: Euroepan Union Transaction Log. http://ec.europa.eu/environment/ets."

**Fawcett, Allen A., Leon E. Clarke, Sebastian Rausch, and John P. Weyant.** 2014. "The EMF Study on U.S. Technology and Climate Policy Strategies: Introduction to EMF24." *Energy Journal*, 35(1): 1–9.

**Gareth, James, Daniela Witten, Trevor Hastie, and Robert Tibshirani.** 2013. *An Introduction to Statistical Learning - with Application in R.* Springer Texts in Statistics.

**Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski.** 2014. "Optimal Taxes on Fossil Fuel in General Equilibrium." *Econometrica*, 82(1): 41–88.

**Goulder, Lawrence H., and Ian W. H. Parry.** 2008. "Instrument Choice in Environmental Policy." *Review of Environmental Economics and Policy*, 2(2): 152–174.

**Goulder, Lawrence H., Marc A. C. Hafstead, and Roberton C. Williams III.** 2016. "General Equilibrium Impacts of a Federal Clean Energy Standard." *American Economic Journal: Economic Policy*, 8(2): 186–218.

**Haavelmo, T.** 1944. "The probability approach in econometrics." *Econometrica*, 12: 1Ű115.

**Hastie, Trevor, Robert Tibshirani, and Jerome Friedman.** 2008. *The Elements of Statistical Learning - Data Mining, Inference, and Prediction.* Springer Series in Statistics.

**Hirst, David.** 2017. "Carbon Price Florr (CPF) and the price support mechanism." *House of Commons - Briefing paper (05927).*

**HM Revenue & Customs.** 2014. "Carbon price floor: Reform. www.gov.uk/government/publications/carbon-price-floor-reform."

**HM Treasury.** 2017. "Autumn Budget 2017. www.gov.uk/government."

**Holladay, J Scott, and Jacob LaRiviere.** 2017. "The impact of cheap natural gas on marginal emissions from electricity generation and implications for energy policy." *Journal of Environmental Economics and Management*, 85: 205–227.

**Holladay, J Scott, Steven Soloway, et al.** 2016. "The Environmental Impacts of Fuel Switching Electricity Generators." *The Energy Journal*, 37(4).

**Holland, Paul W.** 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association*, 81(396): 945–960.

**Imbens, Guido W., and Donald B. Rubin.** 2015. *Causal Inference for STatistics, Social, and Biomedical Sciences: An Introduction.* Cambridge University Press.

**IPCC.** 2006. *Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories.* IPCC.

**IPCC.** 2018. "An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty." *https://www.ipcc.ch/sr15/.*

**Kirat, Djamel, and Ibrahim Ahamada.** 2011. "The impact of the European Union emission trading scheme on the electricity-generation sector." *Energy Economics*, 33(5): 995–1003.

**Knittel, Christopher R., Konstantinos Metaxoglou, and Andre Trindade.** 2015. "Natural Gas Prices and Coal Displacement: Evidence from Electricity Markets." National Bureau of Economic Research.

**Linn, Joshua, Lucija Muehlenbachs, and Yushuang Wang.** 2014. "How Do Natural Gas Prices Affect Electricity Consumers and the Environment?" *Resources for the Future Discussion Paper No. 14-19.*

**Liski, Matti, and Reyer Gerlagh.** 2016. "Carbon Prices for the Next Hundred Years." *Economic Journal*, 128(609): 728–757.

**Lu, Xi, Jackson Salovaara, and Michael B. McElroy.** 2012. "Implications of the Recent Reductions in Natural Gas Prices for Emissions of CO2 from the US Power Sector." *Environmental Science & Technology*, 46(5): 3014–3021.

**Martin, Ralf, Laure B De Preux, and Ulrich J Wagner.** 2014. "The impact of a carbon tax on manufacturing: Evidence from microdata." *Journal of Public Economics*, 117: 1–14.

**Martin, Ralf, Mirabelle Muûls, and Ulrich J Wagner.** 2016. "The impact of the European Union Emissions Trading Scheme on regulated firms: what is the evidence after ten years?" *Review of environmental economics and policy*, 10(1): 129–148.

**Mathiesen, Lars.** 1985. "Computation of Economic Equilibria by a Sequence of Linear Complementarity Problems." *Mathematical Programming Study*, 23: 144–162.

**McGlade, Christophe, and Paul Ekins.** 2015. "The geographical distribution of fossil fuels unused when limiting global warming to 2 degree C." *Nature*, 517: 187–190.

**McGuinness, Meghan, and A Denny Ellerman.** 2008. "CO2 abatement in the UK power sector: evidence from the EU ETS trial period." *MIT CEEPR Working Paper*, 08-010.

**Metcalf, Gilbert E.** 2009. "Market-Based Policy Options to Control U.S. Greenhouse Gas Emissions." *Journal of Economic Perspectives*, 23(2): 5–27.

**Mullainathan, Sendhil, and Jann Spiess.** 2017. "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives*, 31(2): 87–106.

**Nationalgrid.** 2011. "Electricity Then Year Statement 2011 - Appendix F. www.nationalgrid.com/uk/publications."

**Nationalgrid.** 2016. "Final electricity demand. www.nationalgrid.com/uk."

**Nordhaus, William D.** 1994. *Managing the Global Commons: The Economics of Climate Change.* MIT Press, ISBN: 9780262140553, 223 pp.

**Pearl, Judea.** 2009. *Causality.* Cambridge University Press.

**Peters, Jonas, Peter Bühlmann, and Nicolai Meinshausen.** 2016. "Causal inference by using invariant prediction: Identification and confidence intervals." *Journal of the Royal Statistical Society: Series B*, 78(5): 947–1012.

**Pettersson, Fredrik, Patrik Söderholm, and Robert Lundmark.** 2012. "Fuel switching and climate and energy policies in the European power generation sector: A generalized Leontief model." *Energy Economics*, 34(4): 1064–1073.

**Rausch, Sebastian, and Matthew Mowers.** 2014. "Distributional and efficiency impacts of clean and renewable energy standards for electricity." *Resource and Energy Economics*, 36(2): 556 – 585.

**Rubin, Donald B.** 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology*, 66(5): 688–701.

**Rutherford, Thomas F.** 1995. "Extension of {GAMS} for complementarity problems arising in applied economic analysis." *Journal of Economic Dynamics and Control*, 19(8): 1299 – 1324.

**Samii, Cyrus, Laura Paler, and Sarah Zukerman Daly.** 2016. "Retrospective Causal Inference with Machine Learning Ensemles: An Application to Anti-recidivism Policies in Colombia." *Political Analysis*, 24: 434–456.

**Tavoni, Massimo, Elmar Kriegler, Keywan Riahi, Detlef P. van Vuuren, Tino Aboumahboub, Alex Bowen, Katherine Calvin, Emanuele Campiglio, Tom Kober, Jessica Jewell, Gunnar Luderer, Giacomo Marangoni, David McCollum, Mariesse van Sluisveld, Anne Zimmer, and Bob van der Zwaan.** 2014. "Post-2020 climate agreements in the major economies assessed in the light of global models." *Nature Climate Change*, 5: 119–126.

**Tibshirani, Robert.** 1996. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society. Series B (Methodological)*, 58: 267–288.

**Variable Pitch.** 2016. "Variable Pitch. www.variablepitch.co.uk."

**Varian, Hal R.** 2014. "Big data: New tricks for econometrics." *The Journal of Economic Perspectives*, 28(2): 3–27.

**Varian, Hal R.** 2016. "Causal inference in economics and marketing." *Proceedings of the National Academy of Sciences*, 113(27): 7310–7315.

**Venables, W.N., and B.D. Ripley.** 2002. *Modern Applied Statistics with S - Fourth Edition.*

**Weyant, John, Brigitte Knopf, Enrica De Cian, Ilkka Keppo, and Detlef P. van Vuuren.** 2013. "Introduction to the EMF28 Study on Scenarios for Transforming the European Energy System." *Climate Change Economics*, 4(1): 1–3.

**World Bank.** 2018. "State and Trends of Carbon Pricing 2018." *World Bank, Washington, DC. Doi: 10.1596/978-1-4648-1292-7.*

# Appendix A: Additional data

Table A1. Descriptive statistics: annual means and standard deviations of observed hourly electricity demand, generation, and imports by technology category.

| | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
|---|---|---|---|---|---|---|---|---|
| Residual demand | 27.10 | 28.33 | 25.81 | 24.99 | 23.77 | 22.16 | 20.01 | 19.54 |
| | (6.51) | (6.58) | (6.63) | (6.77) | (6.93) | (6.23) | (6.36) | (6.43) |
| Total demand | 36.55 | 37.27 | 35.79 | 35.88 | 35.89 | 34.56 | 34.21 | 33.70 |
| | (7.76) | (8.15) | (7.68) | (7.52) | (7.74) | (7.40) | (7.47) | (7.74) |
| Gas | 17.14 | 18.29 | 14.56 | 9.50 | 9.17 | 9.81 | 9.47 | 14.23 |
| | (3.01) | (3.07) | (3.79) | (4.16) | (5.12) | (4.87) | (4.43) | (4.75) |
| Coal | 9.81 | 9.97 | 10.70 | 14.35 | 13.11 | 10.13 | 8.17 | 3.27 |
| | (5.80) | (5.29) | (5.14) | (4.04) | (3.18) | (4.10) | (3.45) | (2.88) |
| Nuclear | 7.41 | 6.67 | 7.39 | 7.51 | 7.53 | 6.82 | 7.50 | 7.60 |
| | (1.03) | (1.12) | (1.13) | (0.83) | (0.97) | (1.04) | (0.61) | (0.66) |
| Hydro | 0.41 | 0.24 | 0.42 | 0.37 | 0.33 | 0.45 | 0.47 | 0.38 |
| | (0.22) | (0.17) | (0.21) | (0.22) | (0.24) | (0.27) | (0.26) | (0.26) |
| PSP | -0.13 | -0.11 | -0.09 | -0.11 | -0.11 | -0.11 | -0.10 | -0.12 |
| | (1.14) | (1.01) | (0.95) | (0.96) | (0.92) | (0.93) | (0.90) | (0.96) |
| Other | 0.00 | 0.00 | 0.00 | 0.24 | 0.44 | 0.85 | 1.29 | 1.62 |
| | (0.00) | (0.00) | (0.00) | (0.25) | (0.34) | (0.26) | (0.53) | (0.46) |
| Wind | 1.02 | 1.16 | 1.74 | 2.00 | 2.80 | 3.24 | 3.70 | 3.63 |
| | (0.66) | (0.82) | (1.15) | (1.43) | (1.79) | (2.17) | (2.26) | (3.08) |
| Solar | 0.00 | 0.00 | 0.02 | 0.14 | 0.35 | 0.57 | 0.96 | 1.11 |
| | (0.00) | (0.00) | (0.03) | (0.21) | (0.56) | (0.85) | (1.48) | (1.64) |
| Imports | 0.15 | 0.06 | 0.54 | 1.13 | 1.49 | 2.22 | 2.37 | 2.03 |
| | (1.28) | (1.44) | (1.17) | (1.13) | (0.86) | (0.51) | (0.65) | (1.20) |

*Notes:* Standard deviations in parentheses. Data for generation by fuel type is based on ELEXON (2016). Nationalgrid (2016) provides data for final demand and embedded wind and solar generation.

Table A2. Descriptive statistics: installed annual generation capacities by technology category [GW].

| | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
|---|---|---|---|---|---|---|---|---|
| Gas | 20.9 | 23.0 | 23.4 | 25.0 | 24.2 | 24.1 | 23.7 | 23.6 |
| Coal | 25.3 | 25.3 | 25.3 | 24.5 | 19.9 | 19.1 | 19.1 | 15.3 |
| Hydro | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 | 0.9 |
| Nuclear | 11.2 | 11.2 | 11.2 | 11.2 | 11.2 | 11.2 | 11.2 | 11.2 |
| OCGT | 1.4 | 1.4 | 1.4 | 1.4 | 1.3 | 1.3 | 1.3 | 1.3 |
| Oil | 3.7 | 3.7 | 3.7 | 3.7 | 3.7 | 3.7 | 3.7 | 3.7 |
| Other | 2.9 | 2.9 | 2.9 | 2.9 | 2.9 | 2.9 | 2.9 | 2.9 |
| PSP | 2.7 | 2.7 | 2.7 | 2.7 | 2.7 | 2.7 | 2.7 | 2.7 |
| Imports | 2.5 | 2.5 | 3.5 | 3.6 | 4.0 | 4.0 | 4.0 | 4.0 |

*Notes:* Installed capacities are provided by Variable Pitch (2016) and Nationalgrid (2011). Plant characteristics of individual coal and gas plants, i. e., heat efficiencies, emission rates, installed capacities as opening and closure dates are shown in Table 2.

## Appendix B: Machine Learning (LASSO) Algorithm versus OLS

This section compares the out-of-sample performance of the LASSO algorithm versus a standard linear OLS regression model. The comparison of both models is based on the same input variables (and data) as specified in equation (15).

To assess model performance, we proceed in three steps. First, we split out data into eight different pairs of train- and hold-out samples, i.e. each time we use all but one year to train the model and use the remaining year as a hold-out set. Consequently, each of the years 2009 to 2016 is used once as a hold-out set while the rest of the sample is used to train the model. Second, we use each train set to build the models which predict hourly generation $y_{it}$ on a set of input features $x_{it}$ and $z_t$ for each $i \in I$, separately. In this step, we perform cross-validation to tune the regularization parameter $\alpha$. The final step compares different types of models with respect to their in-sample and out-of-sample performance. We can assess for each plant the predictive performance by hold-out year and model type. We use the coefficient of determination—defined as $1 - \sum_i (y_i - \hat{y}_i)^2 / (\sum_i (y_i - \overline{y_i})^2)$—as the score function to evaluate model performance. A test score of 1.0 indicates that the model perfectly predicts the observed data. Note that, in contrast to the commonly reported $R^2$, the test score can be negative because the model can be arbitrarily poor .

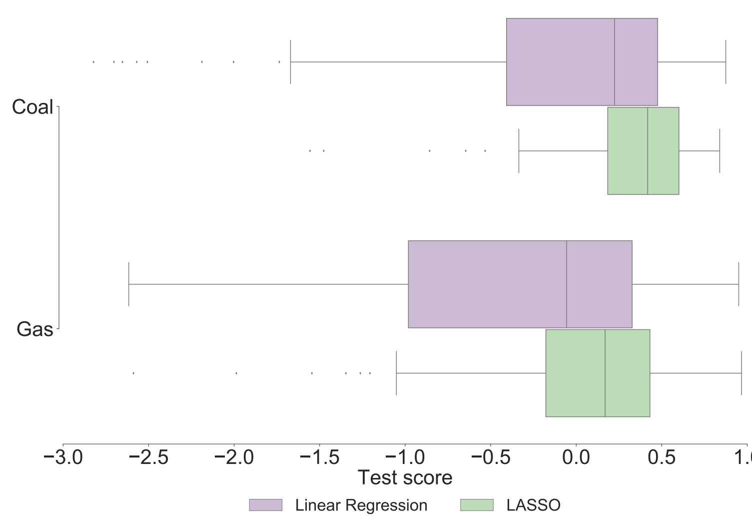



Figure B1. Comparison of the distribution of plant-specific performance scores by fuel type for LASSO vs. OLS models.

Figure B1 compares the test scores of the LASSO and OLS algorithms assessing the prediction of the hold-out set. It is evident that the LASSO outperforms the OLS model in terms of out-of-sample prediction: both average mean scores for coal- and gas-fired plants are higher for LASSO and the respective inter-quartiles ranges are significantly smaller under LASSO as compared to OLS. While from a conceptual perspective the qualitative ranking of LASSO and OLS models in terms of out-of-sample performance are not surprising, Figure B1 makes the important point that in the context of the suggested framework for policy evaluation (and given the specific empirical context), the use of a ML method is advantageous.

## Appendix C: Additional results

### *C1. Plant-specific impacts of UK carbon tax*

Table C1. Impacts of UK carbon tax (CPS) on power plant output [TWh].

| | Period | | | | Total impact |
|---|---|---|---|---|---|
| | 2013 | 2014 | 2015 | 2016 | 2013-2016 |
| *Natural gas plants* | | | | | |
| Pembroke | 0.38 | 0.84 | 2.01 | 0.70 | 3.94 |
| Peterhead | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Staythorpe | 0.19 | 0.65 | 1.40 | 0.29 | 2.53 |
| Didcot CCGT | 0.52 | 0.85 | 2.26 | 1.05 | 4.68 |
| Connahs Quay | 0.28 | 0.58 | 1.04 | 0.68 | 2.58 |
| West Burton CCGT | 0.04 | 0.36 | 0.91 | 0.32 | 1.63 |
| Grain CHP | 0.21 | 0.66 | 1.39 | 0.37 | 2.63 |
| South Humber | 0.17 | 0.35 | 0.63 | 0.41 | 1.55 |
| Seabank | 0.36 | 0.76 | 1.36 | 0.88 | 3.36 |
| Saltend South | 0.07 | 0.17 | 0.67 | 0.49 | 1.41 |
| Immingham CHP | 0.18 | 0.37 | 0.66 | 0.43 | 1.64 |
| Langage | 0.23 | 0.29 | 1.00 | 0.83 | 2.35 |
| Marchwood | 0.04 | 0.08 | 0.14 | 0.09 | 0.35 |
| Severn | 0.12 | 0.25 | 0.44 | 0.28 | 1.09 |
| Spalding | 0.29 | 0.66 | 1.67 | 0.76 | 3.38 |
| Rocksavage | 0.05 | 0.11 | 0.46 | 0.29 | 0.92 |
| Sutton Bridge | 0.08 | 0.18 | 0.31 | 0.20 | 0.77 |
| Damhead Creek | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Coryton | 0.11 | 0.24 | 0.43 | 0.28 | 1.07 |
| Little Barford | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Rye House | 0.06 | 0.11 | 0.17 | 0.09 | 0.43 |
| Medway | 0.18 | 0.61 | 1.23 | 0.34 | 2.36 |
| Baglan Bay | 0.05 | 0.22 | 0.42 | 0.33 | 1.02 |
| Deeside | 0.07 | 0.15 | 0.26 | 0.17 | 0.65 |
| Great Yarmouth | 0.23 | 0.28 | 0.91 | 0.44 | 1.86 |
| Shoreham | 0.01 | -0.05 | 0.17 | 0.12 | 0.25 |
| Enfield Energy | 0.10 | 0.21 | 0.37 | 0.26 | 0.94 |
| Corby | 0.08 | 0.14 | 0.13 | 0.00 | 0.35 |
| Cottam CCGT | 0.18 | 0.32 | 0.79 | 0.63 | 1.92 |
| Fellside | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Fawley Cogen | 0.00 | -0.01 | -0.02 | -0.01 | -0.04 |
| Grangemouth | -0.01 | -0.01 | -0.02 | -0.02 | -0.06 |
| *Coal plants* | | | | | |
| Longannet | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Cottam | -0.88 | -2.15 | -6.95 | -3.47 | -13.46 |
| Ratcliffe | -0.39 | -0.82 | -1.46 | -0.95 | -3.61 |
| West Burton COAL | -1.10 | -2.47 | -5.98 | -3.33 | -12.89 |
| Fiddlers Ferry | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Ferrybridge | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Drax COAL | -0.69 | -1.64 | -3.71 | -2.22 | -8.25 |
| Eggborough | -0.83 | -1.77 | -2.74 | -0.59 | -5.93 |
| Aberthaw | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Rugeley | -0.18 | -0.40 | -0.71 | -0.14 | -1.43 |
| Uskmouth | -0.09 | -0.01 | -0.36 | -0.26 | -0.72 |

*Notes:* Values shown refer to estimated plant-level impacts $\hat{\delta}_{it}^{CPS}$, based on model specification M1 and equation (18). As the CPS is adjusted in April of every year, all reported variables refer to the period from April to March of the subsequent year. As data is available until December 2016, the 2016 period comprises only nine months. The plants are ordered from high to low according to their installed capacity (see Table 2).