iEnergy
ARTICLE
https://doi.org/10.23919/IEN.2025.0027

Do we actually understand the impact of renewables on electricity prices? A causal inference approach

Davide Cacciarelli[1] ✉, Pierre Pinson[1,2,3,4], Filip Panagiotopoulos[5], David Dixon[5] and Lizzie Blaxland[5]

ABSTRACT

Understanding how renewable energy generation affects electricity prices is essential for designing efficient and sustainable electricity markets. However, most existing studies rely on regression-based approaches that capture correlations but fail to identify causal relationships, particularly in the presence of non-linearities and confounding factors. This limits their value for informing policy and market design in the context of the energy transition. To address this gap, we propose a novel causal inference framework based on local partially linear double machine learning (DML). Our method isolates the true impact of predicted wind and solar power generation on electricity prices by controlling for high-dimensional confounders and allowing for non-linear, context-dependent effects. This represents a substantial methodological advancement over standard econometric techniques. Applying this framework to the UK electricity market over the period 2018–2024, we produce the first robust causal estimates of how renewables affect day-ahead wholesale electricity prices. We find that wind power exerts a U-shaped causal effect: at low penetration levels, a 1 GWh increase reduces prices by up to £7/MWh, the effect weakens at mid-levels, and intensifies again at higher penetration. Solar power consistently reduces prices at low penetration levels, up to £9/MWh per additional GWh, but its marginal effect diminishes quickly. Importantly, the magnitude of these effects has increased over time, reflecting the growing influence of renewables on price formation as their share in the energy mix rises. These findings offer a sound empirical basis for improving the design of support schemes, refining capacity planning, and enhancing electricity market efficiency. By providing a robust causal understanding of renewable impacts, our study contributes both methodological innovation and actionable insights to guide future energy policy.

KEYWORDS

Causal inference, electricity prices, renewable energy, wind power, solar power, double machine learning.

Increasing the share of electricity from renewable energy sources, e.g., wind energy and solar energy, is one of the key actions towards climate change mitigation. Projections from the International Energy Agency (IEA) indicate that by 2030, wind and solar power generation assets will jointly contribute to 40% of global electric power generation[1]. In the UK, wind energy alone represented 29.4% of total electricity generation in 2023[2]. Solar energy, while currently in a less prevalent position, is increasing its market share rapidly, driven by low technology costs and generous support policies. The integration of these variable and less predictable energy sources is profoundly affecting electricity markets, requiring a deeper understanding of their impact on market dynamics. Although the environmental benefits of renewable energy generation, such as reduced CO_2 emissions, are well-documented[3], their increasing penetration introduces both opportunities and challenges for the operation of the electricity market[4]. From a market perspective, the inherent lower marginal costs associated with wind and solar power generation are expected to exert downward pressure on wholesale day-ahead prices (also referred to as spot prices) by shifting the equilibrium point through the merit-order effect[5]. However, from an operational point of view, the variability and limited predictability of renewable energy generation induce an additional need for system balancing and ancillary services[6,7]. Spot prices represent the primary economic signal of real-time supply–demand balance in electricity markets. They are directly impacted by variable renewable generation and are thus a natural focus for evaluating short-term market impacts. However, analysing their response to renewables is particularly challenging due to strong non-linearities, time-varying effects, and the presence of confounding variables such as demand shifts and fuel prices. In this study, we aim to move beyond descriptive and correlational analyses by applying a causal inference framework capable of estimating context-specific effects of renewable generation on electricity prices. Our approach focuses on the UK electricity market from 2018 to 2024, where wind and solar penetration levels have steadily increased and where accurate market and forecast data are available. The methodology and empirical insights are, however, relevant to broader electricity systems facing similar transitions.

1 Literature review

The market impact of renewable energy sources, particularly wind energy, has been the subject of extensive research over the past two decades. Early studies were conducted for the Danish electricity market (part of Nord Pool). Although actual wind power generation was found to have limited impact on prices[8], forecast wind power penetration emerged as a primary determinant, reflecting the information available to market participants at the time of gate closure[9]. Since these pioneering studies, researchers have conducted similar investigations in numerous other electricity markets, such as in Australia[10–12], Germany[13–17], Ireland[18,19], Italy[20],

[1]Dyson School of Design Engineering, Imperial College London, London SW7 2AZ, UK; [2]Department of Technology, Management and Economics, Technical University of Denmark, 2800 Kongens Lyngby, Denmark; [3]Halfspace, 1306 Copenhagen, Denmark; [4]Center for Research in Energy: Economics and Markets, Aarhus University, Aarhus, Denmark; [5]National Energy System Operator, Wokingham, UK
Address correspondence to Davide Cacciarelli, d.cacciarelli@imperial.ac.uk or d.cacciarelli@gmail.com

Korea[21], the Netherlands[22], the Nordics[23–25], Spain[26–28], the UK[29,30], and the US[31–34]. More recently, the cannibalisation effect, in which the market value of renewable energy declines as their penetration levels increase by displacing more expensive forms of generation, has been investigated for several European markets[35]. These studies highlight common trends observed across electricity markets globally, such as the price reduction resulting from the merit-order effect and the increased price volatility linked to higher levels of installed wind capacity. Despite these relevant insights, most of the aforementioned works rely on traditional regression-based approaches, which are often limited in their ability to fully capture the complexity of the impact of renewable energy on electricity markets. Here though, the diagnostic analytics approaches to be developed and the insights to be derived ought to be causal. Specifically, simpler analyses might struggle to unravel the influence of confounding factors, such as demand fluctuations, gas prices, and hourly price profiles. In addition, they often overlook the non-linear dynamics that govern the impact of renewable energy penetration on electricity prices, leading to biased or incomplete conclusions. Indeed, traditional econometric approaches are often constrained by their fully parametric nature, requiring strong assumptions about the functional form of the relationships being modelled. While traditional methods such as multiple regression models have long been used to estimate the impact of renewables on prices, these rely on parametric assumptions that may bias estimates in high-dimensional and non-linear settings. Indeed, while panel regressions or machine learning-based forecasting models improve predictive performance, they generally do not provide valid counterfactual estimates needed to understand what would happen to prices if renewable generation were to change. This gap motivates the use of causal inference frameworks.

Only a limited number of studies have attempted to rigorously identify causal effects in electricity markets. The double machine learning (DML) framework proposed by Chernozhukov et al.[36] offers a flexible solution by orthogonalising treatment and outcome with respect to confounders, while leveraging machine learning models to estimate nuisance functions. This allows for valid causal effect estimation even in high-dimensional and non-linear settings. However, standard DML only estimates average treatment effects and does not capture local heterogeneity. We extend the DML framework to estimate non-linear, context-specific causal effects of renewable generation on prices by introducing a local partially linear DML approach based on boxcar kernels. Our method allows treatment effects to vary across the renewable penetration spectrum while maintaining robustness through cross-fitting, bootstrapping, and high-dimensional residualisation. To our knowledge, this is the first study to apply such a framework to the question of how wind and solar power influence UK electricity prices, and among the first to do so in any market using observational data. This contribution advances the literature in two key ways: (i) methodologically, by enabling the estimation of heterogeneous causal effects without relying on strong parametric assumptions; and (ii) empirically, by providing robust evidence on the changing marginal value of wind and solar power over time and across market conditions.

2 Methodology

This study employed different layers of analysis to explore the impact of predicted wind and solar power penetration on electricity prices. To provide a baseline using methodologies commonly employed in the literature, we first analysed the relationship between predicted renewable penetration and spot prices using regression analysis. Then, we developed a robust causal inference approach to estimate the evolving and non-linear impact of renewables while controlling for a large set of confounding variables. The data and code used in this study are publicly available for reproducibility (https://github.com/dcacciarelli/market-impact-renewables).

The dataset for this study comprises electricity market data, wind and solar generation forecasts, and various other market-relevant variables for the UK, spanning the period from 2018 to 2024. A comprehensive list of data sources, detailed descriptions of the variables, and the preprocessing steps used to derive additional variables are provided in Table 1. That table also describes the role of each variable in our models. The dataset integrates observations from the EnAppSys platform (https://www.enappsys.com/), direct inputs from the National Energy System Operator (NESO), and publicly available data from the NESO Data Portal (https://www.neso.energy/data-portal).

2.1 Regression-based analysis

To establish a benchmark and provide intuitive insights into the relationship between renewable energy penetration and electricity prices, we first employ regression-based techniques commonly found in the literature. These methods help characterise mean trends and distributional features of price behaviour before advancing to a more robust causal analysis.

2.1.1 Mean smoothing

Given the well-established non-linear nature of the relationship between electricity prices, predicted renewable energy penetration, and time of day, we employed locally weighted polynomial regression models[9]. These models assume that although the overall relationship is non-linear, it can be locally approximated by a series of linear models. So, by choosing a set of m fitting points, we can locally estimate m linear models representing wholesale electricity prices as a function of predicted wind or solar penetration and time of the day. For this study, we choose 576 fitting points arranged in a 24×24 grid. At each fitting point $\boldsymbol{x}_u$, given that we have n observations $(\boldsymbol{z}_i, p_i)$ available, the price p_i is modelled as

$$p_i = \boldsymbol{\beta}_u^\top \boldsymbol{z}_i + \varepsilon_i, \quad i = 1, \cdots, n, \tag{1}$$

where $\boldsymbol{\beta}_u$ represents the vector of regression coefficients tailored specific to the fitting point $\boldsymbol{x}_u$, $\boldsymbol{z}_i$ comprises a vector of input features, and ε_i denotes the error term (centred, and with finite variance). While with a higher number of fitting points we could adequately approximate non-linear functions with simple first-order models, using polynomial features enables the modelling of more complex relationships. If we let r_i and h_i represent the normalised renewable power level (wind or solar power) and the hour of the day, we can obtain a second-order model by including quadratic and interaction terms, after expanding each data point $\boldsymbol{x}_i = [r_i\ h_i]^\top$ to $\boldsymbol{z}_i = [1\ r_i\ h_i\ r_i^2\ r_i h_i\ h_i^2]^\top$. Locally weighted polynomial regression then estimates the model coefficients $\boldsymbol{\beta}_u$ by weighting observations based on their proximity to the fitting point (using a distance $d(\boldsymbol{x}_u, \boldsymbol{x}_i)$). This is achieved through a kernel function that assigns higher weights to nearby observations, reducing influence as the distance increases. Here, we used the tri-cube weight function given by

$$w_{u,i} = \begin{cases} (1 - d(\boldsymbol{x}_u, \boldsymbol{x}_i)^3)^3, & \text{if } d(\boldsymbol{x}_u, \boldsymbol{x}_i) < 1 \\ 0, & \text{otherwise} \end{cases}, \tag{2}$$

Table 1 Dataset description with variables used in the DML framework. The check marks (√) indicate whether the variable in the row was used to residualise the variable in the column. For example, if a variable is marked with a √ under "Price (wind model)", it means this row's variable was used as a confounder in residualising the electricity price in the wind model

Variable	Description	Source	Price (wind model)	Wind forecast	Price (solar model)	Solar forecast
Date	Date of the observation (datetime index). One observation for each 30-minute settlement period.	EnAppSys	×	×	×	×
Year	Year of observation.	Derived	√	×	√	×
Month	Month of the year (1–12).	Derived	√	√	√	√
Day	Day of the week (0–6, Monday to Sunday).	Derived	√	×	√	×
Hour	Hour of the day (0–23).	Derived	√	√	√	√
Daylight hours	Number of hours of daylight based on geographical location. It has been computed using the Python Astral package (https://astral.readthedocs.io/).	Derived	√	√	√	√
APX price	Day-ahead electricity price on the Amsterdam power exchange (APX) (GBP/MWh).	EnAppSys	×	×	×	×
NordPool price	Day-ahead electricity price on the NordPool exchange (GBP/MWh).	EnAppSys	×	×	×	×
Intraday price	Intraday or within-day (MIDP) electricity price (GBP/MWh).	EnAppSys	×	×	×	×
Actual load	Initial transmission system demand outturn (MW).	EnAppSys	×	×	×	×
Estimated load	Estimated electricity load, generated from actual demand with noise (MW).	Derived	√	×	√	×
Gas price	National balancing point (NBP) price for natural gas [GBP/MWh].	EnAppSys	√	×	√	×
Carbon permits	MID CO_2 prices from the EU Emissions Trading Scheme (GBP/tCO_2e).	EnAppSys	√	×	√	×
Wind capacity	Installed wind capacity (MW).	NESO	×	√	×	×
Wind forecast	Predicted wind production (MW).	NESO	×	×	√	×
Predicted wind penetration	Obtained as: wind forecast / estimated load * 100 (%).	Derived	×	×	×	×
Solar capacity	Installed solar capacity (MW).	NESO	×	×	×	√
Solar forecast	Predicted solar production (MW).	NESO	√	×	×	×
Predicted solar penetration	Obtained as: solar forecast / estimated load * 100) (%).	Derived	×	×	×	×

where d is chosen as the normalised Euclidian distance between an observation $\boldsymbol{x}_i$ and the fitting point at hand $\boldsymbol{x}_u$, i.e., $d(\boldsymbol{x}_u, \boldsymbol{x}_i) = ||\boldsymbol{x}_i - \boldsymbol{x}_u||_2 / h_u$. An adaptive bandwidth h_u for each fitting point $\boldsymbol{x}_u$, can be used to ensure that observations beyond a certain distance from the current point exert no influence on parameter estimation. We calculated h_u using the 30th percentile of distances from all data points to a given fitting point, accommodating to local data density and variability. At each fitting point, the estimated parameter vector $\hat{\boldsymbol{\beta}}_u$ is obtained by minimising the weighted sum of squared residuals, as in

$$\hat{\boldsymbol{\beta}}_u = (\boldsymbol{Z}^\top \boldsymbol{W}_u \boldsymbol{Z})^{-1} \boldsymbol{Z}^\top \boldsymbol{W}_u \boldsymbol{p}, \tag{3}$$

where $\boldsymbol{W}_u$ is a diagonal matrix containing the weights $w_{u,i}$, $\boldsymbol{Z}$ is the design matrix expanded to the quadratic model form (i.e., whose rows are all input feature observations $\boldsymbol{z}_i^\top$), and $\boldsymbol{p}$ is the vector of observed response values p_i. This methodology effectively captures local variations by emphasising the influence of nearby observations, thus providing a detailed and smooth representation of the relationship between wind power production, time of day, and wholesale prices.

2.1.2 Quantile modelling

Understanding the relationship between predicted renewable penetration, time of the day, and mean wholesale prices provides interesting insights on the behaviour of the market. However, to cope with the inherent uncertainty of renewable generation, it is equally important to describe the distribution of data around the mean trend. Using quantile regression, we can get insights into the conditional distribution of the price, indicating the value below which a certain proportion of observations fall for a given quantile with nominal level q. For example, if $q = 0.9$, the quantile regression line will show the value below which 90% of the observed data points lie. Locally weighted polynomial regression can be extended to model quantiles by minimising a cost function that differently weights prediction errors depending on whether they fall below or above the specified quantile. This approach captures the asymmetric risk associated with underestimations or overestimations. The model parameters for a specific quantile q are determined by

$$\hat{\boldsymbol{\beta}}_{u,q} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{n} w_{u,i} \rho_q (p_i - \boldsymbol{z}_i^\top \boldsymbol{\beta}), \tag{4}$$

where p_i represents the observed price, $\boldsymbol{z}_i$ is the i-th observation expanded to the model form and w_i is its weight (computed using (2)), $\boldsymbol{\beta}$ is the vector of coefficients to be estimated, and ρ_q is the quantile loss function defined as $\rho_q(\varepsilon) = \varepsilon(q - 1\{\varepsilon < 0\})$, with $1\{\varepsilon < 0\}$ being an indicator function that takes the value 1 if the residual e is negative and 0 otherwise, and q represents the nominal level of the quantile of interest (e.g., 0.1 for the 10% quantile and 0.9 for the 90% quantile). Unlike the case for mean modelling, there is no closed-form solution for the parameter vector $\hat{\boldsymbol{\beta}}_{u,q}$, which is estimated at each fitting point using an iterative gradient-based optimisation method to minimise the quantile loss function in (4). This methodology provides a detailed representation of the behaviour of prices, beyond average effects, for varying levels of predicted wind power production or penetration.

2.2 Causal inference using double machine learning

Traditional regression models may fall short in disentangling causation from confounding influences, especially when the data involve complex, non-linear interactions or high-dimensional confounders like fuel prices, demand fluctuations, and seasonal patterns. The DML framework[36] leverages machine learning algorithms to flexibly model the relationships between confounders, treatments, and outcomes. This approach enables the estimation of causal effects in the presence of complex, high-dimensional data while mitigating the spurious effects introduced by confounders.

2.2.1 Partially linear DML

The partially linear DML framework[37] assumes that the response p (here, the electricity price) is a function of the treatment t (e.g., predicted renewable energy penetration) and other confounding variables $\boldsymbol{a}$, as in

$$p_i = \beta t_i + f(\boldsymbol{a}_i) + \varepsilon_i, \quad i = 1, \cdots, n, \tag{5}$$

where β represents the (linear) effect of t on the outcome p, f is a potentially non-linear function of the confounders $\boldsymbol{a}$, and ε is an error term (centred, and with finite variance). Similarly, we assume that the treatment variable itself can be modelled as a function of the confounders, or a subset of them, i.e.,

$$t_i = g(\boldsymbol{a}_i) + \eta_i, \quad i = 1, \cdots, n, \tag{6}$$

where g represents the modelled relationship between t and $\boldsymbol{a}$, and η is an error term (centred, and with finite variance). The DML framework is designed to isolate the causal effect of the treatment variable β by adjusting for confounding influences in two stages. First, we estimate the nuisance parameters using two machine learning models to approximate the functions f and g. Then, we use the trained models to “remove” the influence of confounders from both the treatment and the response variables. Then, it regresses the residualised response $\tilde{p}_i = p_i - f(\boldsymbol{a}_i)$, $i = 1, \cdots, n$, on the residualised treatment $\tilde{t}_i = t_i - g(\boldsymbol{a}_i)$, $i = 1, \cdots, n$, to estimate β. Here, β coincides with the average treatment effect (ATE), or average causal effect, which measures the overall impact of a treatment on the outcome across the entire population. Formally, for a continuous treatment variable t, the ATE is defined as

$$\text{ATE} = \mathbb{E}[p \mid do(t+1)] - \mathbb{E}[p \mid do(t)], \tag{7}$$

where $p|do(t)$ represent the potential outcome with treatment t and $p|do(t+1)$ is the potential outcome after a unit increase in t. ATE can be interpreted as the expected increase in the response p, resulting from a unit increase in the treatment t. The relationship is hence linear and not conditional on any contextual variable.

The main steps of DML-based effect estimation for the partially linear case are reported in Algorithm 1, where $\boldsymbol{a}_i$ represents the ith vector of confounding variables, t_i is the value of the treatment for the ith observation, and p_i is the corresponding price value. The function f represents the modelled relationship between p and $\boldsymbol{a}$, while g represents the modelled relationship between t and $\boldsymbol{a}$. It should be noted that the models trained in steps 5 and 6 of Algorithm 1 can be any machine learning model, allowing for flexibility in capturing complex relationships between variables. In our implementation, we employed LightGBM regression models[38].

2.2.2 Local partially linear DML

While the partially linear DML framework provides reliable causal estimates under the assumption of linear treatment effects, it may not fully capture the complexities of the electricity market, where the effect of renewables may vary based on the penetration level. To address this issue, we extend the standard DML approach by introducing a local partially linear DML framework based on a boxcar kernel. Rather than estimating a global ATE across the entire dataset, we assume that the relationship between treatment (renewable energy generation) and outcome (electricity prices) can be locally approximated using a partially linear model within small, predefined subsets of the data. This enables us to estimate the so-called conditional average treatment effect (CATE), which can vary depending on the segment of the data under consideration. Within this context, the boxcar kernel acts as a window function to isolate a subset of the data around each point of interest (e.g., for a given level of penetration, time). The idea is similar in essence to the locally weighted polynomial regression models used previously. For a conditioning (and possibly multidimensional)

Algorithm 1 Partially linear DML

1: **Input:** dataset $D = \{(\boldsymbol{a}_i, t_i, p_i)\}_{i=1}^{n}$, number of folds K

2: **Output:** linear ATE estimate $\hat{\beta}$

3: Randomly partition the dataset into K folds. ▹ Split the data

4: **for** each fold $k = 1, \cdots, K$ **do** ▹ Train predictive models

5: Train a model $\hat{f}_{-k}$ using $K-1$ folds to predict the response p from the confounders $\boldsymbol{a}$.

6: Train a model $\hat{g}_{-k}$ using $K-1$ folds to predict the treatment t from the confounders $\boldsymbol{a}$.

7: Use $\hat{f}_{-k}$ and $\hat{g}_{-k}$ to predict $\boldsymbol{p}$ and $\boldsymbol{t}$ in the held-out fold k.

8: Compute residuals ▹ Orthogonalisation

$$\tilde{\boldsymbol{p}}_k = \boldsymbol{p}_k - \hat{f}_{-k}(\boldsymbol{A}_k)$$
$$\tilde{\boldsymbol{t}}_k = \boldsymbol{t}_k - \hat{g}_{-k}(\boldsymbol{A}_k)$$

9: Regress $\tilde{\boldsymbol{p}}_k$ on $\tilde{\boldsymbol{t}}_k$ using OLS ▹ Effect estimation

$$\hat{\beta}_k = \text{OLS}(\tilde{\boldsymbol{p}}_k, \tilde{\boldsymbol{t}}_k) = (\tilde{\boldsymbol{t}}_k^\top \tilde{\boldsymbol{t}}_k)^{-1} \tilde{\boldsymbol{t}}_k^\top \tilde{\boldsymbol{p}}_k$$

10: **end for**

11: Compute the final estimate as the mean:

$$\hat{\beta} = \frac{1}{K} \sum_{k=1}^{K} \hat{\beta}_k$$

variable $\boldsymbol{x}$ and a fitting point $\boldsymbol{x}_u$, the boxcar kernel is defined as

$$K_u(\boldsymbol{x}) = \begin{cases} 1, & \text{if } \|\boldsymbol{x}-\boldsymbol{x}_u\|_2 \leqslant \dfrac{h}{2}, \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

where h is the width of the window. The kernel takes the value 1 within the bandwidth parameter h around $\boldsymbol{x}_u$ and 0 otherwise. This kernel is applied to partition the data into smaller, rectangular regions where the partially linear model is assumed to hold, yielding a subset $\mathcal{W}(\boldsymbol{x}_u) = \{(\boldsymbol{x}_i, \boldsymbol{a}_i, t_i, p_i), K_u(\boldsymbol{x}_i) = 1\}$. For each segment of the data within the window, we estimate the CATE by applying the partially linear DML method.

Formally, at each fitting point $\boldsymbol{x}_u$ (i.e., a given predicted wind penetration level), we define the local CATE as

$$\begin{aligned} \mathrm{CATE}(\boldsymbol{x}_u) = & \mathbb{E}[p \mid do(t+1), \boldsymbol{x} \in \mathcal{W}(\boldsymbol{x}_u)] \\ & - \mathbb{E}[p \mid do(t), \boldsymbol{x} \in \mathcal{W}(\boldsymbol{x}_u)], \end{aligned} \tag{9}$$

where $\mathcal{W}(\boldsymbol{x}_u)$ is the subset of data within the window centred around the fitting point $\boldsymbol{x}_u$, and the expectation is taken within this window. We refer to $\mathcal{W}(\boldsymbol{x}_u)$ as the contextual subset of data. The procedure follows two main steps, similar to the standard DML, but with the additional local partitioning. For each contextual subset $\mathcal{W}(\boldsymbol{x}_u)$, we use machine learning models to estimate the nuisance models f and g, and residualise both the treatment variable t and the outcome p. This removes the confounding effects, leaving us with the contextual residualised variables $\tilde{t}_u$ and $\tilde{p}_u$, with vectors of values $\tilde{\boldsymbol{t}}_u$ and $\tilde{\boldsymbol{p}}_u$ that corresponds to the contextual subset $\mathcal{W}(\boldsymbol{x}_u)$. The variables used for residualising wind power production, solar power production, and spot prices are listed in Table 1. Then, at each window we estimate the CATE with

$$\hat{\beta}_u = \mathrm{OLS}(\tilde{\boldsymbol{p}}_u, \tilde{\boldsymbol{t}}_u \mid \boldsymbol{x} \in \mathcal{W}(\boldsymbol{x}_u)) = (\tilde{\boldsymbol{t}}_u^\top \tilde{\boldsymbol{t}}_u)^{-1} \tilde{\boldsymbol{t}}_u^\top \tilde{\boldsymbol{p}}_u. \tag{10}$$

To ensure robustness despite data sparsity at extreme penetration levels, we used a fixed-size boxcar kernel (10,000 observations) that guarantees a consistent number of data points for each local effect estimation. This choice ensures estimation stability, but limits the analysis to the penetration range with sufficient support. Consequently, our estimates cover up to approximately 50% wind penetration and 8% solar penetration, as also visible from Figures 3 to 8. On top of this, we use a bootstrap procedure for each contextual subset $\mathcal{W}(\boldsymbol{x}_u)$ to ensure robust results and to capture the distributional uncertainty of our estimates. Additionally, a Gaussian filter was applied to reduce the noise in the CATE estimates and to provide a smoother representation of the non-linear effects.

The key steps of the estimation procedure are described in Algorithm 2, where $\boldsymbol{x}$ is the conditioning variable used to contextualise the estimation of the causal effect. In our implementation, we used kernels with a size h of 10,000 observations, and a step size s of 1,000 observations. The smoothing parameter of the Gaussian filter was set to 1.5 times the standard deviation of the

Algorithm 2 Locally partially linear DML

```
Input: dataset D = {(x_i, a_i, t_i, y_i)}_{i=1}^n, kernel size h, step size s, number of bootstrap replications B, number of folds K
Output: nonlinear CATE estimate
for each fitting point x_u do
    Select subset W(x_u) = {(x_i, a_i, t_i, p_i), ||x - x_u||_2 <= h/2}      ▷ Contextual subset of the data
    for each bootstrap iteration b = 1, ..., B do
        Draw a resampled dataset W^b(x_u) from W(x_u) with replacement      ▷ Bootstrap resampling
        Randomly partition the dataset into K folds.                        ▷ Split the data
        for each fold k = 1, ..., K do                                      ▷ Train predictive models
            Train a model f^b_{-k} using K-1 folds to predict the response p from the confounders a.
            Train a model g^b_{-k} using K-1 folds to predict the treatment t from the confounders a.
            Use f^b_{-k} and g^b_{-k} to predict p and t in the held-out fold k.
            Compute residuals                                               ▷ Orthogonalisation
                p~^b_k = p^b_k - f^b_{-k}(A^b_k)
                t~^b_k = t^b_k - g^b_{-k}(A^b_k)
            Regress p~^b_k on t~^b_k using OLS                              ▷ Effect estimation
                β^b_k(x_u) = OLS(p~^b_k, t~^b_k) = (t~^b_k^T t~^b_k)^{-1} t~^b_k^T p~^b_k
        end for
        Compute the mean estimate for the current bootstrap sample W^b(x_u) as:
            β^b(x_u) = (1/K) Σ_{k=1}^K β^b_k(x_u)
    end for
    Compute the mean estimate for the current kernel W(x_u) as:
        β(x_u) = (1/B) Σ_{b=1}^B β^b(x_u)
end for
Aggregate β(x_u) across all kernel centres x_u to construct the nonlinear CATE estimate
Smooth the aggregated CATE estimates to reduce noise using a Gaussian filter    ▷ Optional
```

In mathematical notation, the steps above read:

- Input: dataset $D = \{(\boldsymbol{x}_i, \boldsymbol{a}_i, t_i, y_i)\}_{i=1}^n$; Select subset $\mathcal{W}(\boldsymbol{x}_u) = \{(\boldsymbol{x}_i, \boldsymbol{a}_i, t_i, p_i), \|\boldsymbol{x}-\boldsymbol{x}_u\|_2 \leqslant \frac{h}{2}\}$
- $\tilde{\boldsymbol{p}}_k^b = \boldsymbol{p}_k^b - \hat{f}_{-k}^b(\boldsymbol{A}_k^b)$, $\quad \tilde{\boldsymbol{t}}_k^b = \boldsymbol{t}_k^b - \hat{g}_{-k}^b(\boldsymbol{A}_k^b)$
- $\hat{\beta}_k^b(\boldsymbol{x}_u) = \mathrm{OLS}(\tilde{\boldsymbol{p}}_k^b, \tilde{\boldsymbol{t}}_k^b) = (\tilde{\boldsymbol{t}}_k^{b\top} \tilde{\boldsymbol{t}}_k^b)^{-1} \tilde{\boldsymbol{t}}_k^{b\top} \tilde{\boldsymbol{p}}_k^b$
- $\hat{\beta}^b(\boldsymbol{x}_u) = \frac{1}{K} \sum_{k=1}^{K} \hat{\beta}_k^b(\boldsymbol{x}_u)$
- $\hat{\beta}(\boldsymbol{x}_u) = \frac{1}{B} \sum_{b=1}^{B} \hat{\beta}^b(\boldsymbol{x}_u)$

mean effect estimates. The number of bootstrap replications within each boxcar kernel was set to 100.

2.2.3 Robustness and assumptions

To ensure the credibility and robustness of our findings, we employ several methodological safeguards throughout the causal inference pipeline. First, we use K-fold cross-fitting within the double machine learning (DML) framework to mitigate overfitting and reduce bias when estimating nuisance functions. This means that for each fold, the models used to predict both the treatment and the outcome are trained on disjoint subsets of the data from those used for estimating the treatment effect. Second, we apply a bootstrapped sliding window approach, combining residualisation with repeated resampling to construct conditional average treatment effect (CATE) estimates along the renewable penetration distribution. This allows us to generate robust pointwise causal estimates along with empirical confidence intervals, highlighting the stability of the results despite variations in data density. Third, we perform orthogonalisation via residualisation: we remove confounding influences from both the treatment (wind or solar production) and the outcome (day-ahead price) using LightGBM-based regressors trained on a rich set of domain-informed covariates. This includes time-of-day effects, load, gas prices, and carbon permits. By separating the predictive models from the final causal regression, our approach ensures orthogonality between residualised variables, a key condition for valid causal inference in high-dimensional settings. The reliability of our method depends on several intrinsic data requirements: (i) the availability of accurate renewable generation forecasts and market prices; (ii) sufficient coverage of key confounders such as load and fuel prices; and (iii) enough observations across the penetration spectrum to support local effect estimation. While the method assumes conditional ignorability—i.e., that all relevant confounders are observed—we address this by incorporating a comprehensive and expert-curated set of variables. Moreover, we apply kernel-based smoothing and quantile aggregation to manage data sparsity in high-penetration regions. These combined strategies ensure that our estimates remain credible, robust, and interpretable across diverse temporal conditions and penetration levels.

3 Relationship between renewable energy penetration and electricity prices

We first examine the relationship between predicted renewable energy penetration and wholesale electricity prices in the UK over a period from 2018 to 2024. Figure 1 presents bar plots of the mean day-ahead APX prices in the UK across varying levels of predicted wind and solar power penetration. For wind power penetration (Figure 1(a)), a clear downward trend emerges, where higher penetration levels are associated with lower electricity prices. This aligns with the well-known merit-order effect, as increased predicted renewable energy generation displaces higher-cost conventional generation, reducing market prices. In contrast, for solar power penetration (Figure 1(b)), the pattern is more complex. While moderate solar penetration levels are linked to lower prices, an unexpected "bump" in the price distribution emerges at around 4%–7% penetration. This counter-intuitive phenomenon may be influenced by external factors such as higher demand during summer months, energy supply disruptions, or other market conditions. Without careful causal analysis, these fluctuations could be mis-attributed solely to solar penetration rather than a broader set of intertwined effects within electricity markets.

To reinforce these initial findings, we performed a regression-based analysis using techniques commonly employed in the literature (using local polynomial (quantile) regression[9]). Figure 2 illustrates the relationship between forecast wind and solar power penetration and spot prices, over the same period. For wind power (Figure 2(a)), the results corroborate earlier insights, showing that higher predicted wind power penetration generally leads to lower wholesale electricity prices. Specifically, as predicted wind power penetration increases, daily price patterns undergo a noticeable transformation: the pronounced peaks during morning and evening hours at low penetration levels are significantly smoothed. This indicates that wind power forecasts not only reduces average prices but also dampen daily price spikes. Results from the quantile regression model (Figure 2(c)) provide further insights into the distributional effects of wind power generation. With higher levels of predicted wind penetration, the range between the 10% and 90% quantiles narrows, suggesting reduced price variability. However, the lighter colour shading at high penetration levels indicates limited data in these regions, cautioning against overgeneralising these findings.

The analysis of predicted solar power penetration (Figures 2(b) and 2(d)) reveals a more complex relationship. As noted earlier (Figure 1(b)), higher levels of solar penetration appear to be linked with rising spot prices (which is counter-intuitive). The discrepancy between observed behaviour for wind and solar power also highlights the importance of penetration levels. With wind energy

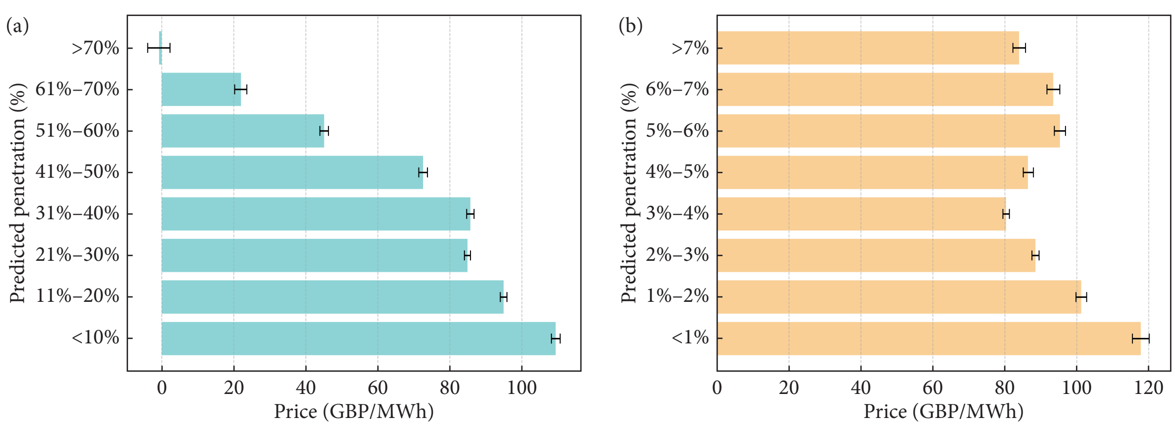

Figure 1 Price variation across predicted renewable penetration levels. Bar plots show the mean APX prices for wind and solar penetration levels, over 2018–2024, with error bars representing associated 95% confidence intervals. The predicted wind power penetration intervals **(a)** are defined in 10% increments, while those for predicted solar power penetration **(b)** are defined in 1% increments, reflecting the lower installed solar power capacity in the UK.

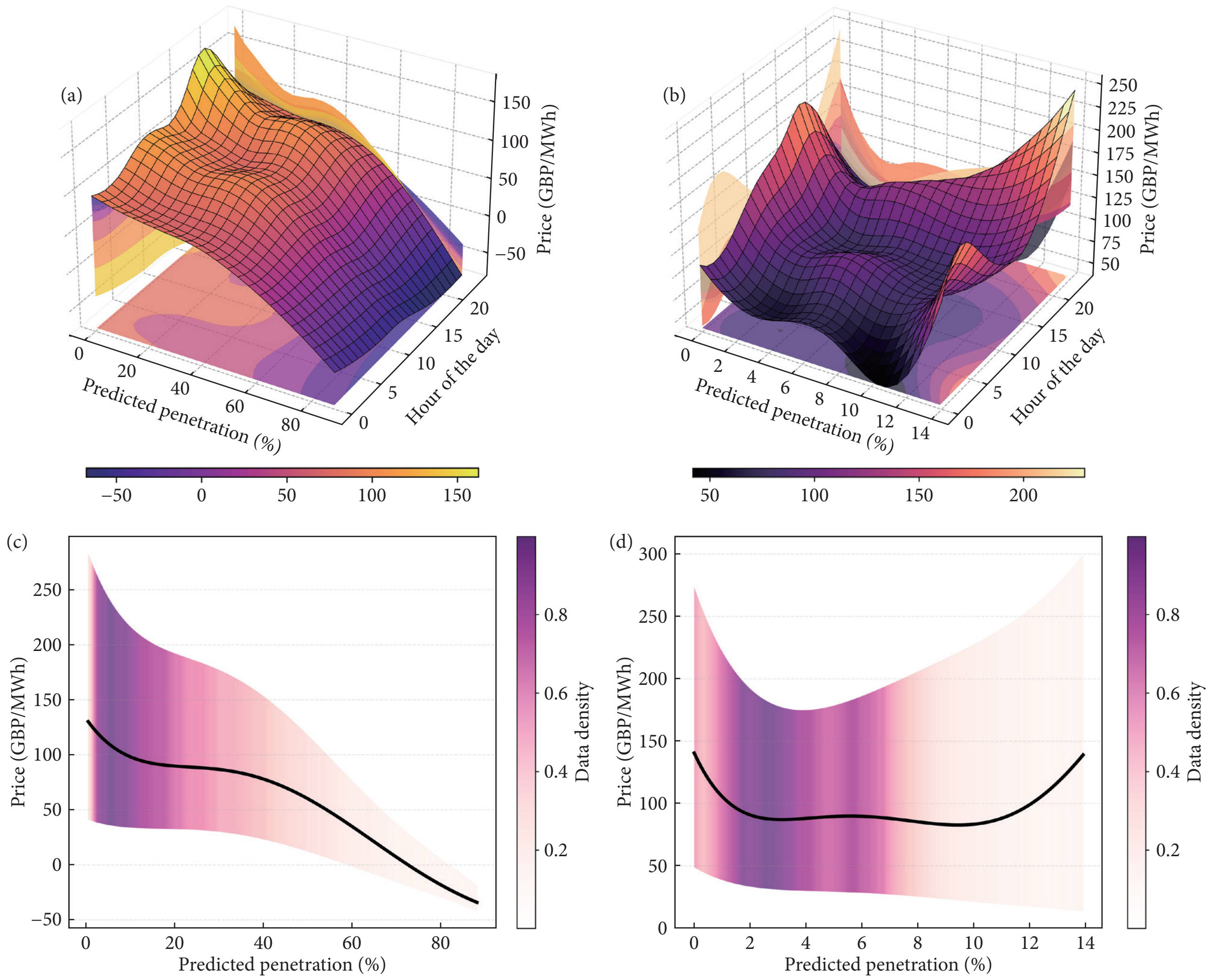

Figure 2 Regression analysis of predicted renewable penetration and APX wholesale prices. The top row shows the mean spot price modelled as a function of the hour of the day, as well as predicted **(a)** wind power penetration and **(b)** solar power penetration. The bottom row presents the results of quantile regression models, with the solid line representing the mean spot price, and the shaded region representing central 80%-coverage intervals, where the colour intensity indicates the data density (in terms of penetration data). These models capture the distributional effects of predicted **(c)** wind power penetration and **(d)** solar power penetration, on spot prices.

being more pervasive in the UK market, its effect is easier to discern, whereas the impact of solar energy is potentially overshadowed by external factors. These results underscore the limitations of regression-based approaches in disentangling the causal effects of renewable energy penetration. Without accounting for confounding variables, we risk drawing misleading conclusions, such as inferring that higher solar power penetration drives higher wholesale prices. This emphasises the necessity of devising and employing robust causal inference methods to accurately attribute changes in electricity prices to renewable energy penetration.

4 Causal impact of renewables on electricity prices

Understanding the impact of wind and solar energy generation on electricity prices is inherently a causal question: what happens to electricity prices if we predict one more GWh of wind or solar energy? This inquiry goes beyond a simple statistical association; it seeks to uncover the direct, underlying effect of additional renewable energy integration. Historically, experimentation has been central to scientific discovery, providing controlled environments to isolate causes and effects. However, in many real-world environments, such as electricity markets, conducting controlled experiments is neither feasible nor practical. In these cases, deriving causal insights from observational data alone remains one of the most formidable challenges in science and statistics[39–40]. Given the limitations of traditional regression models, we employ an approach based on DML to move beyond simple associations and uncover the true effects of renewable integration. Specifically, our local partially linear DML framework enables us to isolate the nonlinear impact of wind and solar power production on wholesale electricity prices. This methodology addresses confounding factors that often distort observational analyses, providing a robust foundation for understanding how renewable integration affects market dynamics.

4.1 Estimated causal effects

Figure 3 shows the results of our DML framework on the APX prices, for the period 2018–2024. In particulat, it provides a visualisation for the individual CATE estimates obtained from each bootstrap iteration of our causal inference framework. These estimates represent the causal impact, in GBP/MWh, of a 1 GWh

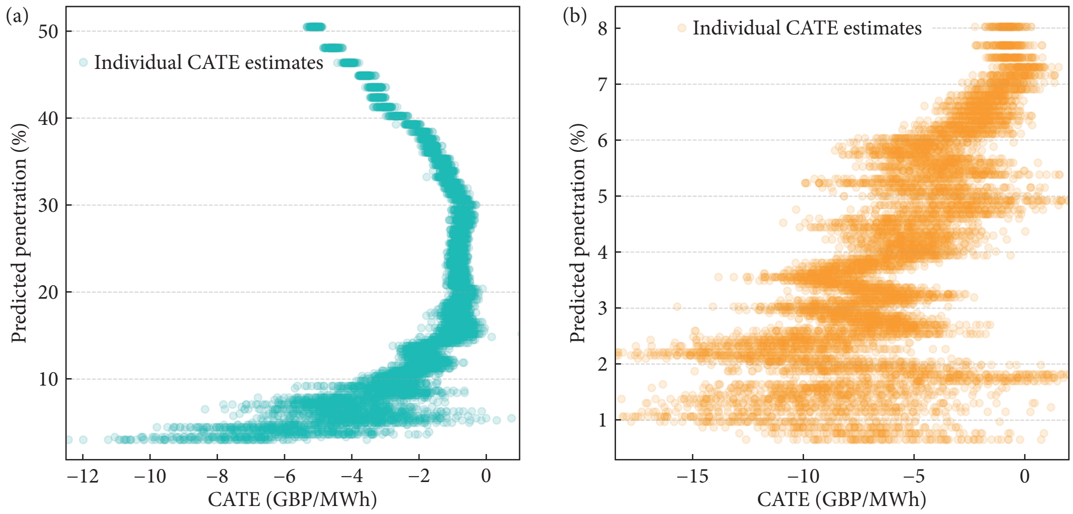

Figure 3 Individual CATE estimates prior to smoothing. Individual CATE estimates of the effect of wind **(a)** and solar **(b)** power production on electricity prices.

increase in renewable energy generation. Each data point corresponds to the output of our locally partially linear DML method prior to any smoothing.

The results after applying the smoothing step are shown in Figure 4. The same results for NordPool and intraday prices are depicted in Figures 5 and 6. The causal estimate is the conditional average treatment effect (CATE), which represents the impact of a 1 GWh increase in predicted renewable energy generation on wholesale electricity prices, while controlling for a wide set confounders. Instead, the observational mean reflects the raw association between renewable penetration and electricity prices at the same penetration level, without accounting for confounding factors. The full list of potential confounding variables to consider is given in Table 1. It includes variables like varying installed capacity (for wind and solar energy), time of day, time of year, load, gas price, carbon permits, etc. Importantly, while we use predicted renewable energy generation as the treatment variable in our causal analysis, we explicitly account for demand (load) in two ways: (i) it is included as a confounder in the residualisation step, ensuring its influence on prices is not mistakenly attributed to renewables; and (ii) we use predicted renewable penetration as a contextual variable to condition the estimation of effects, thus allowing us to study how the marginal impact of an additional GW varies across different penetration levels. This dual strategy enables a robust separation of the effects of renewables from those of load.

The causal analysis reveals distinct effects for wind and solar power. Wind power exhibits a U-shaped causal effect on electricity prices (Figure 4(a)): the price-reducing effect from wind energy generation is strongest at low penetration levels, diminishes at moderate levels, and increases again at higher penetration levels. This pattern likely reflects the interaction between wind energy generation and the slope of the supply curve in the wholesale electricity markets, at various penetration levels. At low wind power penetration, wind power rapidly displaces the most expensive marginal generators, leading to substantial price reductions. As penetration increases, this effect diminishes temporarily, only to re-emerge as wind displaces more entrenched higher-cost generators. The reduced effect in the moderate range (20%–30%) could arise because the supply curve flattens as gas providers with similar cost structures enter the market, reducing the sensitivity of prices to

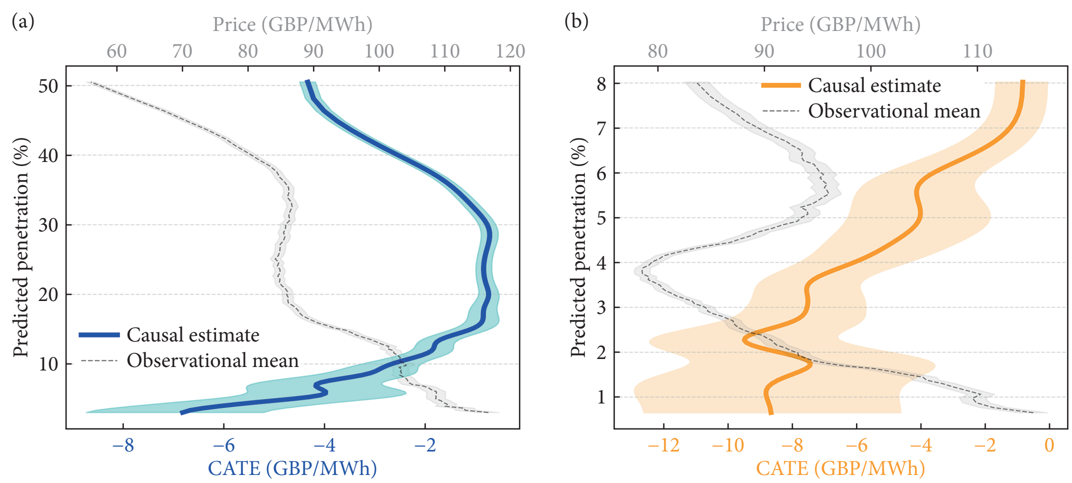

Figure 4 Causal impact of predicted renewable power production on electricity prices (APX). Comparison of observational mean and causal effects of wind **(a)** and solar **(b)** power production on wholesale electricity prices. The solid lines represent the non-linear CATE estimates derived using our local partially linear DML framework, capturing the price impact (in GBP/MWh) of a 1 GWh increase in renewable energy generation. The dashed lines show the observational mean trends as a function of renewable penetration levels, illustrating the differences between raw associations and the true causal effects. Shaded areas denote 80% confidence intervals.

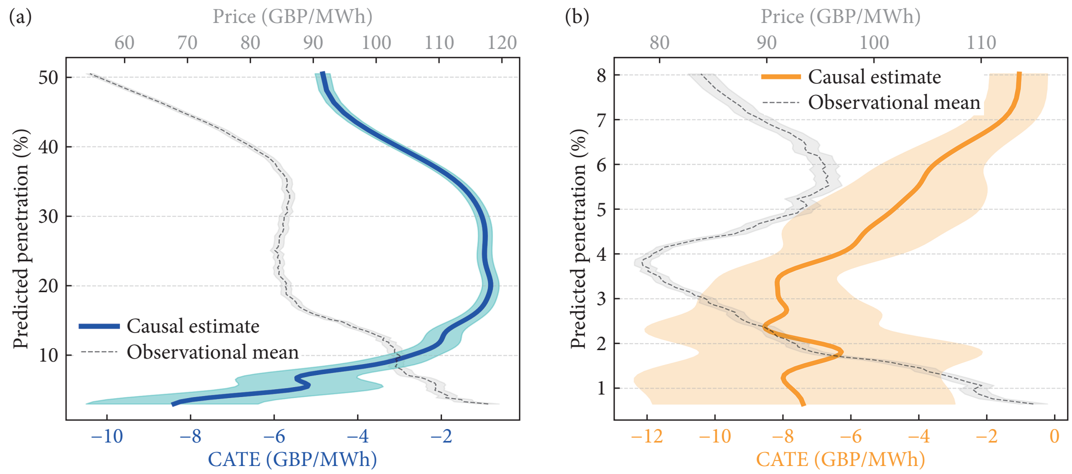

Figure 5 Causal impact of predicted renewable power production on electricity prices (NordPool). Comparison of observational mean and causal effects of wind **(a)** and solar **(b)** power production on wholesale electricity prices.

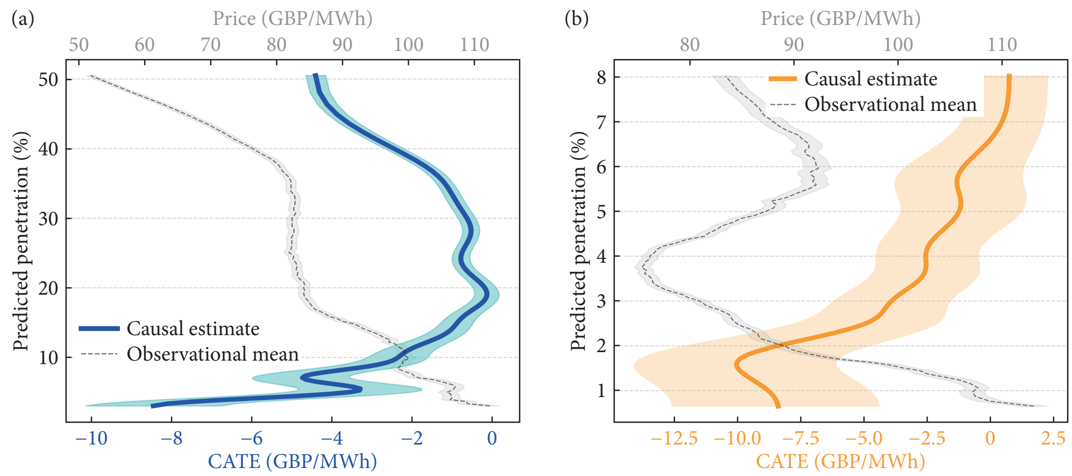

Figure 6 Causal impact of predicted renewable power production on electricity prices (intraday). Comparison of observational mean and causal effects of wind **(a)** and solar **(b)** power production on wholesale electricity prices.

additional wind power generation. For solar power (Figure 4(b)), our causal inference approach uncovers a consistent price-reducing effect across all penetration levels. This finding stands in sharp contrast to the observational mean trends, which suggest a "bump" in electricity prices at moderate solar penetration (around 5%). This bump, observed in simpler analyses, likely arises from unaddressed confounding factors such as demand shifts or major events occurring during periods of high solar output. The causal analysis effectively isolates the intrinsic merit-order effect of solar power, demonstrating its role in reducing electricity prices even when confounders might obscure this relationship. Finally, we hypothesise that the effect of solar power on electricity prices might also exhibit a U-shaped pattern, similar to wind, since the effect fundamentally comes from the overall shape of the supply curve in the market. However, this potential trend is not fully observable in our analysis due to the limited attainable penetration of solar power, which reflects the reduced installed solar capacity in the UK. As solar power penetration increases in the future, it is plausible that a more pronounced U-shaped relationship could emerge, driven by dynamics analogous to those seen for wind power. Comparable patterns and causal relationships have been observed while analysing the effects on the NordPool (Figure 5) and intraday (Figure 6) electricity prices. In Figure 7, we explored an alternative approach by using predicted renewable penetration not only as a contextual variable but also as the main input to the model. This analysis provides a complementary view by estimating the causal effect of increasing renewable penetration (%) directly, rather than focusing on energy generation (GWh).

These findings have significant implications for analysts, policy makers and market participants. The U-shaped effect observed in Figure 4(a) underscores the importance of considering how renewable energy influences market structures across varying levels of market penetration. For solar power, the discrepancy between observational and causal trends underscores the risks of relying on simpler regression methods, which may lead to disproportionately erroneous conclusions about the economic impacts of renewables. By providing an accurate assessment of the true effects of renewable generation, our approach supports better-informed market designs and policy interventions aimed at maximising the benefits from renewable energy integration.

We provide an alternative analysis where we attempt to estimate the CATE using predicted renewable penetration as the input variable. Previously, predicted penetration was used as a contextual variable within the boxcar kernel, slicing the data into subsets to estimate localised effects. In that framework, the CATE represented the price impact of adding 1 extra GWh of renewable energy gen-

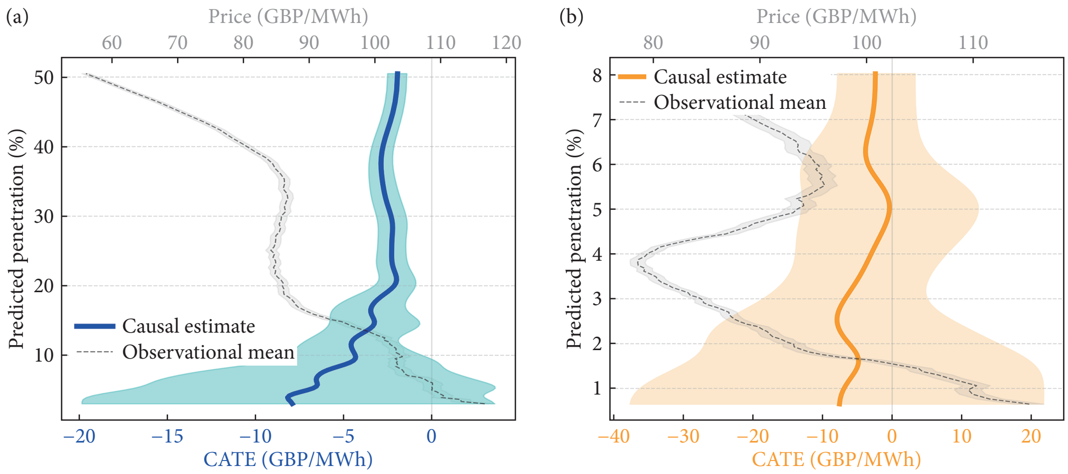

Figure 7 Effect of the predicted penetrations. Nonlinear CATE estimates derived using our locally partially linear DML framework, capturing the price impact [GBP/MWh] of a 1% increase in renewable penetration.

eration to the system. Here, we present an alternative analysis where the CATE reflects the price impact of increasing renewable penetration by 1 percentage point (%). Figure 7 shows the results of applying our DML framework with predicted renewable penetration serving both as the input variable for the final model and as a contextual variable for the boxcar kernel. This approach introduces potential challenges, particularly regarding additional confounding factors. Normalising the predicted renewable production by the total estimated load to compute predicted penetration may introduce dependencies that are not fully accounted for by the standard confounders used in residualising production. This added confounding underscores the need for caution when interpreting these results. Unlike penetration, renewable production forecasts are primarily driven by well-understood factors such as weather conditions and seasonal patterns, making them less susceptible to hidden confounders. To mitigate these issues, we applied residualisation to predicted penetration using the same confounders employed for renewable production—namely, wind (or solar) capacity, daylight hours, hour of the day, and month of the year. Additionally, we included estimated load and predicted renewable production (wind or solar) as confounders to address dependencies introduced by the penetration calculation.

In Figure 7, the smoothing factor of the Gaussian filter was set equal to the standard deviation of the mean CATE. For both wind and solar, we observe increased variability in the estimates. For wind, this variability is particularly pronounced below 10% penetration. For solar, the variability remains consistently higher across the entire range of predicted penetration. This heightened variability is likely due to the smaller effect size, making the estimates more susceptible to confounding factors introduced during the normalisation process. Nevertheless, the mean causal estimates remain negative throughout the range, reaffirming the merit order effect, whereby increased renewable penetration reduces electricity prices.

4.2 Evolution of causal effects over time

The impact of renewables on electricity prices is not static. It evolves alongside market conditions, technological advancements, and changes in renewable penetration levels driven by additional installed generation capacity. To investigate these dynamics, we analyse the non-linear causal effect over time using a sliding window approach spanning two financial years. This approach provides a temporally resolved view of how the price effects of wind and solar energy have changed over the period 2018–2024, while ensuring a sufficient number of observations to apply our DML framework. The findings in Figure 8 reveal a growing impact of renewable energy on electricity prices over the years. In particular, the magnitude of wind's price-reducing effect (Figure 8(a)) has significantly increased in the last four years, reflecting its expanding role in the electricity market and the growing relevance of the merit-order effect. This trend underscores the profound influence of wind power on market outcomes, highlighting its policy implications as governments and regulators aim to integrate higher levels of renewable energy into the grid. The evolution of the effect of solar power on electricity prices (Figure 8(b)) also appears to be increasing, indicating that the impact of solar energy may become substantial in a near future with the further deployment of solar power generation capacity in the UK.

5 Conclusions

5.1 Discussion

This study addresses a fundamental question in electricity market research: what is the true causal impact of renewable energy generation on wholesale electricity prices? By advancing and applying a robust causal inference framework based on local partially linear DML, we move beyond conventional regression-based analyses that can conflate correlation with causation. Our results provide clear insights into the market effects of wind and solar power. We show that wind power exerts a U-shaped impact on prices where price reductions are strongest at low penetration levels, moderate at mid-levels, and intensify again at high penetration. This pattern likely reflects the structure of the supply curve and the varying marginal costs of displaced generation. In contrast, solar power consistently reduces prices across the observed penetration range, reinforcing the merit-order effect even in the presence of complex confounding factors. Importantly, we demonstrate that both effects have become more pronounced over time, indicating an increasing influence of renewable energy on price formation in the UK electricity market between 2018 and 2024. While we focus on day-ahead prices, the proposed causal inference framework is readily extensible to other electricity price signals (e.g., intraday, balancing, ancillary services) or market outcomes (e.g., volatility, reserve costs, emissions). The key requirement is the availability of sufficient observational data and relevant confounders for the spe-

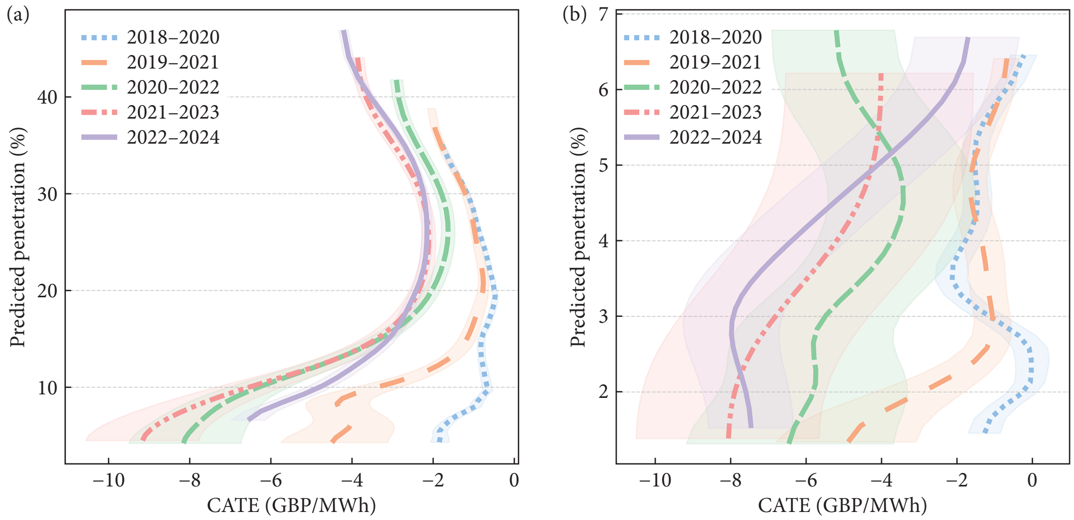

Figure 8 Temporal evolution of the causal impact of renewable power production on electricity prices. non-linear CATE estimates of (a) wind and (b) solar power production over time, obtained using a sliding window spanning two financial years, with each window containing approximately 35,000 observations. Shaded regions represent 80% confidence intervals.

cific time scale and outcome of interest.

These findings carry important implications for analysts, market operators, and policymakers. Traditional econometric approaches, often used in energy policy design, may yield misleading results by failing to control for confounding variables or to account for non-linear dynamics. Our causal estimates offer more reliable evidence to inform market interventions, capacity planning, and regulatory reforms aimed at accommodating rising shares of variable renewable generation. Although our study focuses on the UK, the methodological approach is widely applicable. Many electricity markets face similar challenges as they integrate increasing levels of renewables. Our framework can be adapted to these contexts, supporting comparative analyses and global policy benchmarking. Furthermore, as solar and wind penetration continue to grow, the observed causal relationships may become even more significant, potentially amplifying both the economic and system-level impacts of renewables. Indeed, although our empirical analysis is UK-specific, the methodology is general and can be applied to any electricity market with appropriate data. As in this study, domain expertise can support confounder selection and model specification, though data-driven or hybrid causal discovery methods may be employed in settings where expert knowledge is limited. Moreover, the observed non-linear effects of renewables on prices suggest that one-size-fits-all market mechanisms may be suboptimal. Our findings support policies that dynamically adjust support schemes or capacity incentives based on penetration levels and temporal factors. For instance, as the value of renewables fluctuates non-linearly with penetration, subsidy mechanisms may benefit from being adaptive to market signals. Moreover, the increasing effect of renewables over time reinforces the need for robust capacity planning, integration with carbon markets, and investments in system flexibility.

5.2 Limitations and future research directions

While this study advances the understanding of the causal effects of renewable energy on electricity prices, several limitations merit discussion and point to fruitful avenues for future research. First, although our model accounts for a comprehensive set of observed confounders, including fuel prices, demand, time-of-day effects, and installed capacity, unobserved confounding remains a potential source of bias. Moreover, the DML framework assumes that observations are independently and identically distributed (i.i.d.), yet electricity market data inherently exhibit time-series dependencies. While we mitigate this through the inclusion of temporal controls, future work should investigate causal inference methods specifically designed for time-dependent or panel data structures, or address issues like concept drift and structural breaks to deal with evolving market scenarios.

Beyond these methodological considerations, several substantive research directions emerge. A natural extension of this work is to examine how renewable energy integration affects other critical market outcomes, such as price volatility, balancing costs, and reserve requirements. These system-level dimensions are increasingly relevant as variable renewable energy penetration grows and system flexibility becomes more valuable. Similarly, the interactions between renewables and emerging technologies, including long-duration storage and demand-side flexibility, warrant causal investigation. These technologies will likely play a central role in shaping future market dynamics and mitigating renewable-induced variability.

Finally, expanding the scope of causal analysis beyond prices to include broader societal outcomes such as emissions reduction and reliability would contribute to a more holistic understanding of renewable energy's value. Applying causal frameworks to these dimensions could improve the design of energy policies aimed at not only efficiency, but also equity and sustainability. In summary, while this work provides a solid foundation for understanding the causal price effects of renewables, it also opens a rich agenda for methodological and applied research that will be essential in navigating the next phases of the energy transition.

Article history

Received: 2 June 2025; Revised: 20 July 2025; Accepted: 12 September 2025

Additional information

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] IEA (2023). Net zero roadmap: A global pathway to keep the 1.5 °C goal in reach. Technical Report, International Energy Agency (IEA). Available at https://www.iea.org/reports/net-zero-roadmap-a-global-pathway-to-keep-the-15-c-goal-in-reach

[2] National Grid (2023). How much of the UK's energy is renewable? Available at https://www.nationalgrid.com/stories/energy-explained/how-much-uks-energy-renewable

[3] IPCC (2023). Summary for policymakers. In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Lee, T. H. and Romero, J. Eds.]. IPCC, Geneva, Switzerland). Available at https://www.ipcc.ch/report/ar6/syr/summary-for-policymakers/

[4] Morales, J. M., Conejo, A. J., Madsen, H., Pinson, P., Zugno, M. (2013). *Integrating Renewables in Electricity Markets: Operational Problems*. New York, NY: Springer.

[5] Antweiler, W., Muesgens, F. (2021). On the long-term merit order effect of renewable energies. *Energy Economics*, 99: 105275.

[6] Miettinen, J., Holttinen, H. (2019). Impacts of wind power forecast errors on the real-time balancing need: A Nordic case study. *IET Renewable Power Generation*, 13: 227–233.

[7] Goodarzi, S., Perera, H. N., Bunn, D. (2019). The impact of renewable energy forecast errors on imbalance volumes and electricity spot prices. *Energy Policy*, 134: 110827.

[8] Morthorst, P. E. (2003). Wind power and the conditions at a liberalized power market. *Wind Energy*, 6: 297–308.

[9] Jónsson, T., Pinson, P., Madsen, H. (2010). On the market impact of wind energy forecasts. *Energy Economics*, 32: 313–320.

[10] Cutler, N. J., Boerema, N. D., MacGill, I. F., Outhred, H. R. (2011). High penetration wind generation impacts on spot prices in the Australian national electricity market. *Energy Policy*, 39: 5939–5949.

[11] Forrest, S., MacGill, I. (2013). Assessing the impact of wind generation on wholesale prices and generator dispatch in the Australian National Electricity Market. *Energy Policy*, 59: 120–132.

[12] Bell, W. P., Wild, P., Foster, J., Hewson, M. (2017). Revitalising the wind power induced merit order effect to reduce wholesale and retail electricity prices in Australia. *Energy Economics*, 67: 224–241.

[13] Würzburg, K., Labandeira, X., Linares, P. (2013). Renewable generation and electricity prices: Taking stock and new evidence for Germany and Austria. *Energy Economics*, 40: S159–S171.

[14] Ketterer, J. C. (2014). The impact of wind power generation on the electricity price in Germany. *Energy Economics*, 44: 270–280.

[15] Gürtler, M., Paulsen, T. (2018). The effect of wind and solar power forecasts on day-ahead and intraday electricity prices in Germany. *Energy Economics*, 75: 150–162.

[16] Do, L. P. C., Lyócsa, Š., Molnár, P. (2019). Impact of wind and solar production on electricity prices: Quantile regression approach. *Journal of the Operational Research Society*, 70: 1752–1768.

[17] Keeley, A. R., Matsumoto, K., Tanaka, K., Sugiawan, Y., Managi, S. (2021). The impact of renewable energy generation on the spot market price in Germany: Ex-post analysis using boosting method. *The Energy Journal*, 42: 1–22.

[18] O'Mahoney, A., Denny, E. (2011). The merit order effect of wind generation on the Irish electricity market. MPRA Paper. Available at https://mpra.ub.uni-muenchen.de/56043/

[19] O'Flaherty, M., Riordan, N., O'Neill, N., Ahern, C. (2014). A quantitative analysis of the impact of wind energy penetration on electricity prices in Ireland. *Energy Procedia*, 58: 103–110.

[20] Clò, S., Cataldi, A., Zoppoli, P. (2015). The merit-order effect in the Italian power market: The impact of solar and wind generation on national wholesale electricity prices. *Energy Policy*, 77: 79–88.

[21] Shcherbakova, A., Kleit, A., Blumsack, S., Cho, J., Lee, W. (2014). Effect of increased wind penetration on system prices in Korea's electricity markets. *Wind Energy*, 17: 1469–1482.

[22] Nieuwenhout, F., Brand, A. (2011). The impact of wind power on day-ahead electricity prices in the Netherlands. In: Proceedings of the 2011 8th International Conference on the European Energy Market (EEM), Zagreb, Croatia.

[23] Hu, X., Jaraitė, J., Kažukauskas, A. (2021). The effects of wind power on electricity markets: A case study of the Swedish intraday market. *Energy Economics*, 96: 105159.

[24] Spodniak, P., Ollikka, K., Honkapuro, S. (2021). The impact of wind power and electricity demand on the relevance of different short-term electricity markets: The Nordic case. *Applied Energy*, 283: 116063.

[25] Tselika, K., Tselika, M., Demetriades, E. (2024). Quantifying the short-term asymmetric effects of renewable energy on the electricity merit-order curve. *Energy Economics*, 132: 107471.

[26] Gelabert, L., Labandeira, X., Linares, P. (2011). An ex-post analysis of the effect of renewables and cogeneration on Spanish electricity prices. *Energy Economics*, 33: S59–S65.

[27] Cruz, A., Muñoz, A., Zamora, J. L., Espínola, R. (2011). The effect of wind generation and weekday on Spanish electricity spot price forecasting. *Electric Power Systems Research*, 81: 1924–1935.

[28] Azofra, D., Jiménez, E., Martínez, E., Blanco, J., Saenz-Díez, J. C. (2014). Wind power merit-order and feed-in-tariffs effect: A variability analysis of the Spanish electricity market. *Energy Conversion and Management*, 83: 19–27.

[29] Barthelmie, R. J., Murray, F., Pryor, S. C. (2008). The economic benefit of short-term forecasting for wind energy in the UK electricity market. *Energy Policy*, 36: 1687–1696.

[30] Joos, M., Staffell, I. (2018). Short-term integration costs of variable renewable energy: Wind curtailment and balancing in Britain and Germany. *Renewable and Sustainable Energy Reviews*, 86: 45–65.

[31] Woo, C. K., Horowitz, I., Moore, J., Pacheco, A. (2011). The impact of wind generation on the electricity spot-market price level and variance: The Texas experience. *Energy Policy*, 39: 3939–3944.

[32] Woo, C. K., Moore, J., Schneiderman, B., Ho, T., Olson, A., Alagappan, L., Chawla, K., Toyama, N., Zarnikau, J. (2016). Merit-order effects of renewable energy and price divergence in California's day-ahead and real-time electricity markets. *Energy Policy, 92*: 299–312.

[33] Martinez-Anido, C. B., Brinkman, G., Hodge, B. M. (2016). The impact of wind power on electricity prices. *Renewable Energy*, 94: 474–487.

[34] Prol, J. L., Steininger, K. W., Zilberman, D. (2020). The cannibalization effect of wind and solar in the California wholesale electricity market. *Energy Economics*, 85: 104552.

[35] Stiewe, C., Xu, A. L., Eicke, A., Hirth, L. (2024). Cross-border cannibalization: Spillover effects of wind and solar energy on interconnected European electricity markets. *arXiv preprint*, arXiv: 2405.17166.

[36] Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21: C1–C68.

[37] Fuhr, J., Berens, P., Papies, D. (2024). Estimating causal effects with double machine learning–A method evaluation. *arXiv preprint*, arXiv: 2403.14385.

[38] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA.

[39] Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. Cambridge, UK: Cambridge University Press.

[40] Peters, J., Janzing, D., Schölkopf, B. (2017). *Elements of Causal Inference: Foundations and Learning Algorithms*. Cambridge, MA, USA: MIT Press.