Impact of wind and solar production on electricity prices: quantile regression approach

Forthcoming in Journal of the Operational Research Society

Linh Phuong Catherine Do[1], Štefan Lyócsa[2] and Peter Molnár[3]

Abstract

We study the impact of fuel prices, emission allowances, demand, past prices, wind and solar production on hourly day-ahead electricity prices in Germany over the period from January 2015 until June 2018. Working within a linear regression, ARX-EGARCH and quantile regression framework we compare how different pricing factors influence the mean and quantiles of the electricity prices. Contrary to the existing literature we find that short-term price fluctuations on the fuel markets and emission allowances have little effect on the electricity prices. We also find that day-of-the-week as well as monthly effects have significant impact on the electricity prices in Germany and should not be ignored in model specifications. Three main factors are found to drive extreme prices: price persistence, expected demand and expected wind production. Our findings contribute to understanding of extreme price movements, which can be used in pricing models and hedging strategies.

Keywords: price modelling, wind, solar, renewables, quantile regression

[1] Norwegian University of Science and Technology, Trondheim, Norway

[2] Faculty of Management, University of Prešov, Slovakia and Faculty of Economics and Administration, Masaryk University, Brno, Czech Republic

[3] UiS Business School, University of Stavanger, Stavanger, Norway; Norwegian University of Science and Technology, Trondheim, Norway; and Faculty of Finance and Accounting, University of Economics, Prague, Czech Republic and Faculty of Economic Sciences and Management, Nicolaus Copernicus University in Toruń, Toruń, Poland. Peter Molnár acknowledges the support provided by the Czech Grant Agency under Grant No. 19-26812X. Corresponding author, email: peto.molnar@gmail.com.

1. Introduction

The liberalization of the power market in Germany has introduced competition and increased market participants' exposure to risk. In the new market structure, the extreme electricity price volatility can be higher than for commodities or stocks. Coupling to the price risk, the market participants face risks associated with unexpected outage, fluctuation in demand, fuel price and emission allowances. The expansion of renewable energy in Germany and its volatility in production has increased the day-ahead electricity price variance even further (Jacobsen and Zvingilaite 2010; Green and Vasilakos 2010). Hence, an understanding how the fundamental drivers of electricity price affect the electricity price is necessary in order to manage the risks involved in the market.

Electricity pricing exhibits several intrinsic features, which are unique in comparison to the prices of commodities such as gas and oil. Since electricity is not storable, and is also influenced by transmission constraints, electricity market is organized in a more complicated way than market for other commodities. For a typical commodity, there are two main markets: a forward market and a spot market. The main difference between market for electricity and market for other commodities lies in the spot market. The electricity markets which are somehow equivalent to the spot market of other commodities are the day-ahead market and the intraday/balancing market. Focus of this paper is the day-ahead market. For reader interested in intraday market we recommend an excellent article by Kiesel and Paraschiv (2017).

Furthermore, the electricity price exhibits both volatility clustering and large spikes. The possibility of extreme price movements increases the risk for the market participants. Hence, modelling the probability of extreme prices can be more important than the expected values (Bunn et al, 2016, Hagfors et al, 2016b). In this paper we analyze how electricity price reacts to fundamental variables, with focus on wind and solar production, using both an ordinary and a quantile regression.

The response of electricity price to wind and solar production has been studied extensively for various markets, for example Australia (Worthington and Higgs, 2017), Denmark (Jónsson, Pinson and Madsen, 2010), Great Britain (Green and Vasilakos, 2010), Germany (Ketterer, 2014; Paraschiv et al, 2014; Ziel, Steinert and Husmann, 2015), Italy (Clò, Cataldi and Zoppoli, 2015; Sapio, 2019), New England (Martinez-Anido, Brinkman and Hodge, 2016), and many others, for a review see Würzburg, Labandeira and Linares (2013) and Dillig, Jung a Karl (2016).

These studies are interested in an average impact of wind and solar production on price. However, this is not sufficient for the purpose of risk management and related applications. Jónsson, Pinson and Madsen (2010), who conclude that "It is quite obvious that the spot prices are not Gaussian distributed and therefore it must be deemed highly unlikely that models constructed with least squares techniques will have Gaussian residuals. Prediction intervals for such models should therefore be estimated using other techniques. In fact the distributions are so far from parameterized distributions that it seems reasonable to conclude that non-parametric approaches, like for instance quantile regression, will return the most reliable prediction intervals." In accordance with this recommendation, we complement the linear regression with an ARX-EGARCH model that allows for innovations to follow a very flexible distribution that accounts for potential heavy-tails and we also utilize quantile regression that allows us to study the impact of wind production, solar production, and fundamental factors on the whole distribution of electricity price in a non-parametric way.

Quantile regressions have been applied in financial risk management and recently in energy market studies: household energy consumption (Kaza, 2010), electricity demand (Do et al, 2016a; He et al, 2019), oil prices (Lee and Zeng, 2011) and CO_2 emission allowance price (Hammoudeh et al, 2014). Quantile regression has been successfully applied also to electricity price forecasting, see Jónsson et al. (2014), Nowotarski and Weron (2014), Nowotarski and Weron (2015), Juban et al. (2016), Maciejowska and Nowotarski (2016), Moreira, Bessa and Gama (2016), Maciejowska et al (2016), Bello et al. (2017), Liu et al (2017), Mosquera-López et al. (2017), Uniejewski, Marcjasz and Weron (2018). Since the focus of our paper is not on forecasting, reader interested in electricity price forecasting can see Weron (2014), Uniejewski, Nowotarski and Weron (2016), Hong and Fan (2016), Nowotarski and Weron (2018) and Ziel and Steinert (2018).

Papers most closely related to our work are Bunn et al (2016) and Hagfors et al (2016).[1] They both study impact of several fundamental factors, mainly fuel prices, on UK electricity price. However, neither of these two studies include wind and solar production in their analysis. Our in-sample approach, allows us to study in detail impact of various variables, including wind and solar production, on the distribution of hourly electricity price in Germany over period 2015 – 2018. We make several contributions to the understanding on the drivers of electricity prices across quantiles. We show that short-term price fluctuations on the coal, Brent oil, natural gas and emission allowances (EUA) markets does not have impact on the German day-ahead electricity prices. The role of the past production of electricity from renewable resources of wind and solar had not been research is such depth before. We find that extreme (low and high alike) electricity prices are not only persistent but that they are at least partly driven by wind

[1] A brief conference paper of Hagfors et al. (2016) is more closely related to our study, but our analysis is much more detailed in terms of models specifications, models employed, discussion and we utilize new data which have no overlap with their dataset.

production, and particularly wind production drives low prices during night sessions. Lastly, we also control for the weekday effects, weekly seasonality in prices, holiday effects, day-light seasonality, and long-term seasonality using monthly dummies and it turns out that all seasonal effects are important drivers of electricity price and should not be neglected in pricing models.

Rest of this paper is organized as follows: Section 2 provides an overview of the German electricity market; Section 3 discusses the relationship between renewables and extreme prices; Section 4 describes the models while Section 5 describes data used in this study. In Section 6 we present our key results and the final section concludes.

2. The German electricity market

The German electricity market was fully liberalized in 1998. In the liberalization process eight utilities merged into four utilities: RWE, E.ON, Vattenfall and EnBw Energie. These four vertically integrated utilities were responsible for the supply, transmission and balancing of electricity. Since the liberalization was considered to be progressing too slowly, a directive was issued by the European commission establishing an unbundling policy, and the four utilities then sold a majority stake in their transmission share to third parties. Today, there are still four large electricity generators and four transmission companies, but they act independently. The market is liberalized both for the supply and retail electricity markets. The German market is considered a competitive environment although there is some degree of market power (Janssen and Wobben, 2008).

The supply and demand curves are important components for understanding the electricity market. Figure 1 illustrates the Merit Order curve as a sorted short-term marginal cost curve of electricity production; the renewables have the lowest marginal cost, followed by nuclear energy, lignite, hard coal, natural gas and oil power plants. As depicted in Figure 1 the short-term marginal costs consist mainly of fuel and CO_2 costs. This suggests that an increase in the marginal cost of the input variables will lead to an increase in the electricity price. The demand curve is inelastic, meaning that demand remains almost unchanged with changes in electricity price (Sensfuß et al, 2008).

The intersection between the supply and demand curves determines the clearing price for electricity. Every day, a day-ahead auction for each of the 24 hours takes place at 12 p.m.. Each hour is dominated by a different type of power plant (Murray, 2009); conventional power plants remain the price setting utilities in the German market. Normally, nuclear energy, lignite and coal power plants cover the base load, while gas power plants cover the peak load (Sensfuß et al, 2008).

Renewable production was given priority access to the grid and has nearly zero marginal costs. As a result, renewable production enters at the base of the Merit Order curve and shifts the curve to the right, so that cheaper conventional power plants set the price (Zachman, 2013). This means that additional renewable infeed to the grid will reduce the electricity prices.

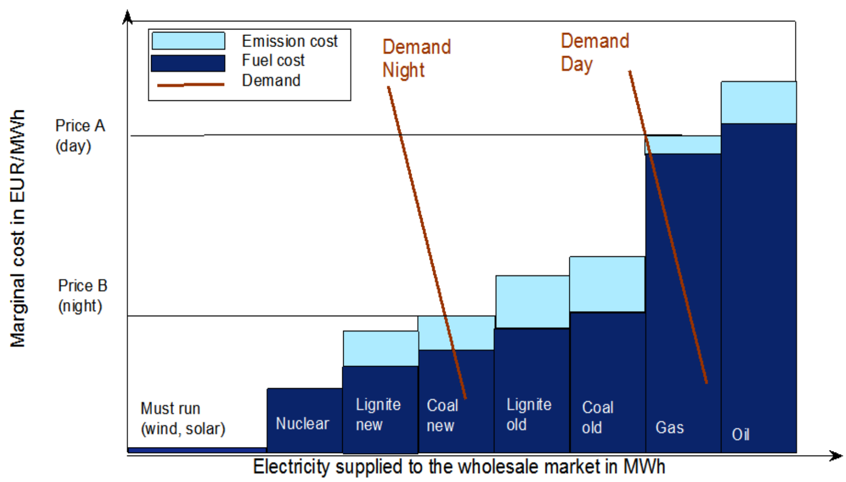

Figure 1 A stylized example of the stepwise marginal cost function and demand function for day and night.

3. Renewables and extreme prices

Lindstrom and Regland (2012) study electricity prices on six European electricity markets and find that the frequency of extreme events is positively correlated with the amount of renewable sources installed on the grid. On the electricity price market, extreme prices are not only extremely positive prices, but also negative prices, which are not uncommon. Compared to financial and other commodity markets, this makes the electricity price market unique. Extreme prices are not uncommon on the German electricity market studied in this paper as well. In Figure 2 the upper panel shows continuous electricity price series (24 prices for each hour of each day of a weak), while the lower panels show electricity prices for selected hours (1 price for each day corresponding to a given hour). Sudden price upsurge and drops are not uncommon, but it is also evident that the price series differ with respect to a given hour. For example, looking at the middle panel of Figure 2, the price series is more volatile than the price series in for the bottom panel of Figure 2 (also see Table 2 for comparison of summary statistics). Detailed examination of our data shows that the negative prices occur more often during the night than during the day, while price spikes appear during the day. Furthermore Figure 3 shows how extreme prices are distributed between various hours. The upper panel show the distribution of negative prices, while the middle and bottom panels show the distribution for prices below the 5^{th} and above the 95^{th} quantiles. A close inspection of our data reveals that similarly as in Lindstrom and Regland (2012) increasing share of electricity production from renewable sources is associated with higher occurrence of extremely low prices (below the 5^{th} quantile). At the same time, higher share of production from renewable sources is associated with lower occurrence of extremely high prices (above the 95^{th} quantile)[2].

[2] The pairwise correlation between exceedance of prices below the 5^{th} percentile (1 if yes, 0 otherwise) and the share of energy produced from renewable sources is at 0.40. The corresponding correlation for the exceedance of prices above the 95^{th} percentile (1 if yes, 0 otherwise) is -0.21.

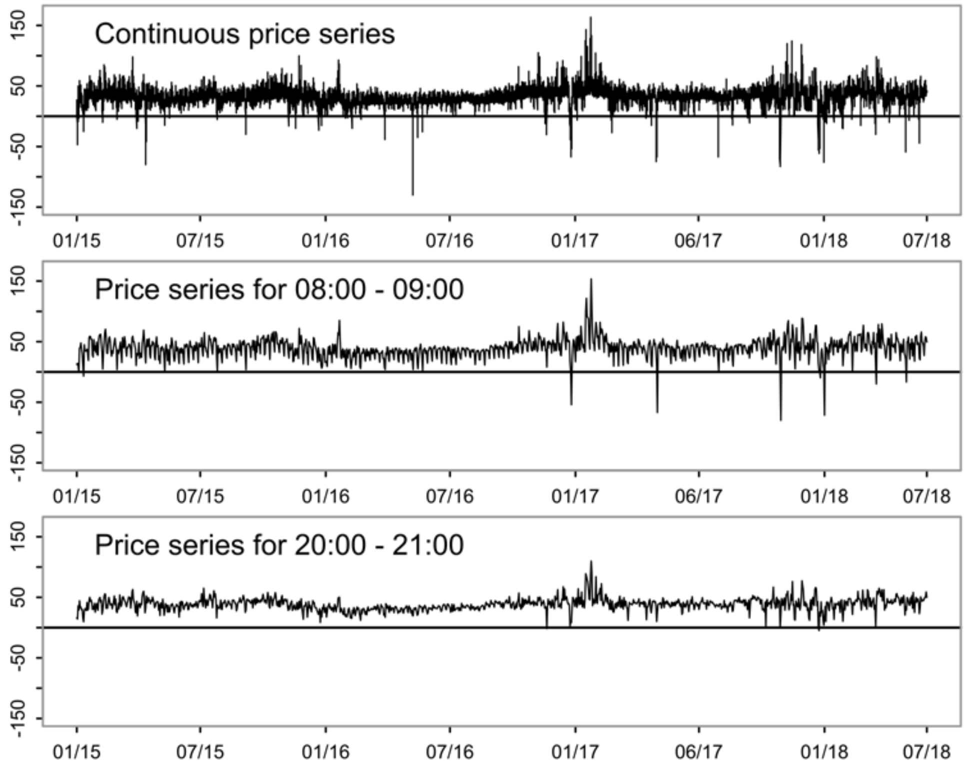

Figure 2 Electricity price in the German market

The reason behind negative electricity prices is that must-run inflexible utilities, like nuclear power plants, are willing to pay the consumer because the cost of shutting down exceeds the costs of accepting the negative price (Keles et al. 2012). Additionally, a high level of solar and wind generation with essentially zero marginal costs leads to negative electricity prices when coupled with lower demand.

On the other hand, electricity price spikes can occur for many different reasons, for instance unpredicted generation outage or transmission failures. Another reason is high demand coupled with low renewable production, which results in additional firing of power plants higher on the merit order curve, pushing the prices up.

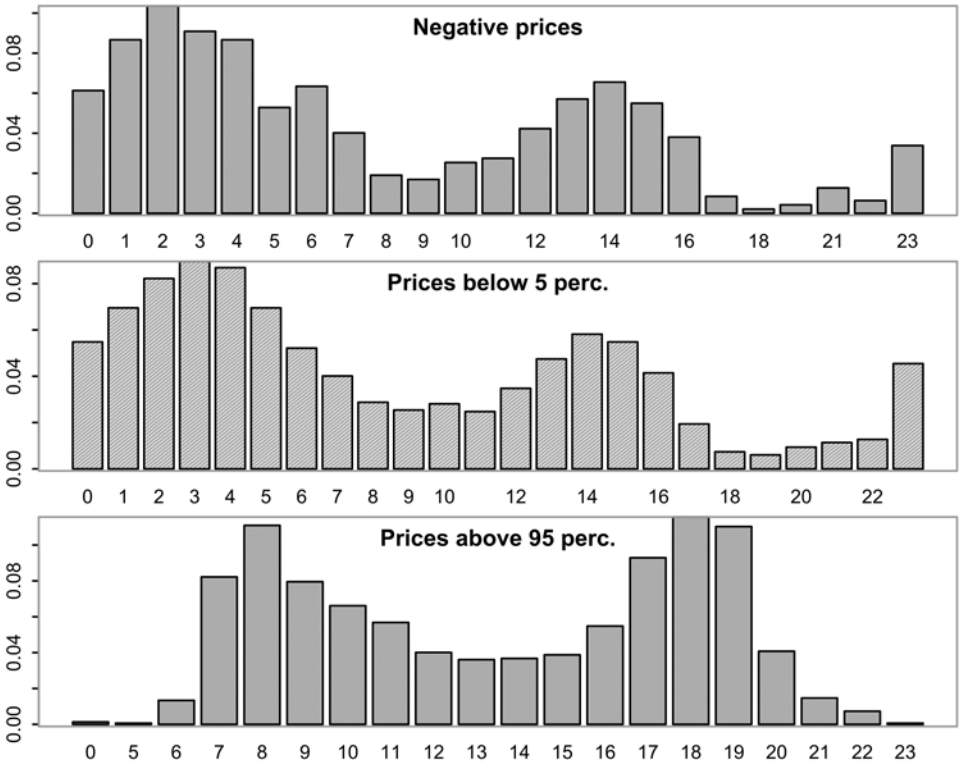

Figure 3 Distribution of extreme prices over different time of the day

4. Modeling framework

We model electricity prices on the German market using a linear regression model to establish baseline results and subsequently we use a quantile regression framework because it allows us to examine the tails of the distribution of electricity prices, i.e. extreme upward and downward prices that can be regularly found for the German electricity market. In the previous Section 3 as well as from the next Section 5 it is evident, that electricity prices on the German market show distinct features for each of the hours during the day. Therefore each model is estimated for each hour of the day.

The baseline results are established within the linear regression model estimated via ordinary least squares. More specifically, we estimate the following model specification:

$$\begin{aligned} P_t^h = {} & \beta_0 + \beta_1 P_{t-1}^h + \beta_2 P_{t-7}^h + \beta_3 DF_t^h + \beta_4 DA_{t-1}^h + \beta_5 \Delta C_{t-1} + \beta_6 \Delta G_{t-1} + \beta_7 \Delta B_{t-1} + \beta_8 \Delta E_{t-1} + \\ & \beta_9 WF_t^h + \beta_{10} SF_t^h + \beta_{11} SR_{t-1} + \sum_{j=1} \beta_{j+11} Day_{t,j} + \sum_{k=1} \beta_{k+17} Month_{t,k} + \beta_{29} Hol_t + \varepsilon_t^h \end{aligned} \quad (1)$$

The electricity price P_t^h at day t and for hour h is explained via the electricity price for the same hour of the previous day, $P_{t\text{-}1}^h$ and of the previous weak $P_{t\text{-}7}^h$. This lag-structure should capture most of the systematic short-term seasonal patterns. The DF_t^h is the forecasted demand for day t (it is therefore exogenous) and hour h, while $DA_{t\text{-}1}^h$ is the actual consumption from the previous day. The changes on the energy markets are captured by $\Delta C_{t\text{-}1}$ for coal, $\Delta G_{t\text{-}1}$ for Natural gas, $\Delta B_{t\text{-}1}$ for Brent oil and $\Delta E_{t\text{-}1}$ for the market of the European Union Emission Allowances. Expected production of renewables for day t is captured by WF_t^h and SF_t^h, while the actual share of energy production using renewables resources (wind and solar) is incorporated into the $SR_{t\text{-}1}$ variable. From evening 22:00 until morning 05:00, the electricity production from solar energy sources is minimal, therefore SF_t^h is excluded from specifications for these hours. Finally, we also use a set of control variables, namely day of the week dummies $Day_{t,j}$ (Wednesday excluded), monthly dummies $Month_{t,j}$ (July excluded) and a variable controlling for the upcoming national holidays Hol_t.

We complement the regression analysis by employing an ARX(7,0)-EGARCH(1,1) model which is defined in Appendix A. The motivation for considering an ARX(7,0)-EGARCH(1,1) model is twofold. First, in the ARX(7,0)-EGARCH(1,1) model the innovations are allowed to follow a distribution that accounts for possible skewness and kurtosis making it potentially more useful when extreme observations are of concern. This might be indeed the case with electricity prices, that due to demand surge or supply shocks can suddenly increase or increase by considerable magnitude. Second, within an ARX(7,0)-EGARCH(1,1) model we account for the serial dependence of the innovations directly and also take into account possible conditional heteroscedasticity. Interestingly, as results in Appendix A show, this modeling choice is well supported by the data[3], but as qualitatively the results do not differ much, in the

[3] The residuals from ARX(7,0)-EGARCH(1,1) show no traces of serial dependence and all the parameters of this model are significant.

main body of the manuscript we present results from the linear regression only, while the interested reader finds ARX(7,0)-EGARCH(1,1) results in Appendix A.

To study the drivers behind extreme price movements we work within a quantile regression framework. We assume that the price P_t^h is related to a set of exogenous explanatory variables. Let $\mathbf{P}^h$ denote a ($T \times 1$) vector of prices for hour h, with T denoting the number of observations (t = 1, 2, …, T). $k - 1$ exogenous variables are stacked in a ($T \times k$) matrix $\mathbf{X}$ that also includes a constant, while $\boldsymbol{\beta}(\tau)$ is a ($k \times 1$) vector of unknown parameters, $\boldsymbol{\varepsilon}(\tau)$ the ($T \times 1$) vector of disturbances and τ a quantile (0, 1). Quantile regression model can be formulated as:

$$\mathbf{P}^h = \mathbf{X}^{\mathrm{T}}\boldsymbol{\beta}(\tau) + \boldsymbol{\varepsilon}(\tau) \tag{2}$$

while assuming that τ-th quantile error term conditional on $\mathbf{X}$, $\boldsymbol{\beta}(\tau)$ is equal to 0. Coefficients are estimated by minimizing the weighted sum of absolute deviations between the electricity price P_t^h and a linear combination of variables:

$$\widehat{\boldsymbol{\beta}}(\tau) = \underset{\beta(\tau)\in \mathrm{R}^{\mathrm{k}}}{\arg\min} \left\{ \sum_{t:P_t^h \geq \mathbf{X}_t^{\mathrm{T}}\boldsymbol{\beta}(\tau)} \tau \left| P_t^h - \mathbf{X}_t^{\mathrm{T}}\boldsymbol{\beta}(\tau) \right| + \sum_{t:P_t^h < \mathbf{X}_t^{\mathrm{T}}\boldsymbol{\beta}(\tau)} (1-\tau) \left| P_t^h - \mathbf{X}_t^{\mathrm{T}}\boldsymbol{\beta}(\tau) \right| \right\} \tag{3}$$

The vector $\mathbf{X}_t^{\mathrm{T}}$ contains all our explanatory variables defined in (1). We estimate equation (3) for τ = 0.05, 0.25, 0.50, 0.75, 0.95, while the optimization of equation (3) is performed via the Frisch–Newton interior point algorithm (see p. 289 in Portnoy and Koenker 1997; for details), while to address the quantile crossing issue, (3) is estimated simultaneously across all quantiles, given the non-crossing quantile restriction along the lines of eq. 2 in Bondell et al., (2010). The significance of the quantile regression coefficients is calculated using a fixed block length bootstrap procedure, where given the potential weakly seasonality, the block length was set to 7 and 1000 bootstrap samples were reproduced.

5. Data

5.1 The electricity price

This paper uses hourly day-ahead German electricity prices P_t^h (the Physical Electricity Index) provided by the European Energy Exchange market (EEX), where t is a time index which denotes a given day (t = 1, 2, …) and h is an index that denotes a given hour, h = 0, 1, …, 23. We choose to study the day-ahead prices for a given hour rather than intraday prices because they represent a larger share of the trading volume. The electricity price dataset covers the period from January 1st 2015 to May 31st 2018. Given that we prices are settled for each hour within a day, this lead to 29929 price observations on 24 time series[4].

Some earlier studies applied logarithmic transformation to the electricity price series in order to achieve variance stabilization (Conejo et al, 2005; Bunn et al 2016; Hagfors et al 2016a). Karakatsani and Bunn (2010), on the other hand, argue that logarithmic transformation is not relevant in the electricity market, because it conceals detailed statistical properties and induces error effects. Further, logarithm can be applied only to positive numbers and the electricity price data also contains negative prices. Also, in this study we focus also on modeling extreme prices that would be largely mitigated using the logarithmic transformation. Finally, the logarithmic transformation is justified if extreme prices would be higher than the average or median of the price in order of several magnitudes and this is not the case of electricity prices in Germany. We therefore use the price data directly rather than using the logarithm of prices.

The full and sample price series are visualized in Figure 2, while in Table 1 we present descriptive data on price series. In each row, summary statistics present the price data for given hour. As already noted in previous Section 3, the price series exhibit quite distinct price paths

[4] The overall number of observations in not even due to the day-light saving times and occasional missing observation.

that are further amplified in Table 1. The price is peaks twice, during morning from 08:00 to 09:00 and evening 19:00 to 20:00. During the evening session, negative prices are less likely as indicated by lower price skewness.

A high level of persistence of the price series would indicate a possible presence of a unit root, as we latter work with the 24 price series within an OLS and quantile regression framework, such issues would need to be addressed properly. In Table 1 we indicate the autocorrelation at first, fifth and seventh lag. The autocorrelation coefficients range from 0.47 to 0.68 at the 1st lag, they tend to decline to a range from 0.04 to 0.38 at the 5th lag, and are up again at the 7th lag at a range from 0.11 to 0.57. These results suggest two important features of our series that are distinct from price data of financial assets: i) the time series are less persistent, ii) the time series show weakly price seasonality. The present of a seasonal-unit root was tested using three standard tests of Osborn et al., (1988), Hylleberg et al., (1990), and Canova and Hansen (1995). They each indicated an absence of a seasonal-unit root. We therefore study each price series in their level.

Table 1 Descriptive statistics of the electricity price and demand for given hours

	Electricity price, P_t^h [EUR]								Total demand, DA_t^h [MWh]							
Hour	Mean	SD	Min	Max	Skew.	$\rho^A(1)$	$\rho^A(5)$	$\rho^A(7)$	Mean	SD	Min	Max	Skew.	$\rho^A(1)$	$\rho^A(5)$	$\rho^A(7)$
00:00 - 01:00	25.86	10.09	-79.94	57.01	-2.20	0.47	0.11	0.18	46371	4460	33830	61699	0.20	0.78	0.45	0.81
01:00 - 02:00	23.98	10.57	-83.00	51.04	-2.35	0.52	0.08	0.17	44437	4434	32638	59909	0.19	0.77	0.44	0.79
02:00 - 03:00	22.65	11.60	-83.03	53.05	-2.95	0.54	0.07	0.11	43480†	4456	31615	58463	0.19	0.78	0.44	0.79
03:00 - 04:00	22.07	10.92	-83.03	48.82	-2.46	0.52	0.09	0.17	43511	5038	0	57893	-1.43	0.64	0.32	0.69
04:00 - 05:00	22.49	10.74	-83.04	51.52	-2.49	0.54	0.11	0.20	44543	4810	31455	58668	-0.02	0.75	0.36	0.82
05:00 - 06:00	24.49	10.44	-83.02	56.06	-2.61	0.54	0.12	0.26	46830	5754	31854	60979	-0.30	0.63	0.16	0.83
06:00 - 07:00	30.31	13.44	-83.01	91.21	-2.13	0.48	0.06	0.44	52265	8528	32758	68258	-0.56	0.48	-0.07	0.83
07:00 - 08:00	36.88	16.39	-80.00	163.52	-0.46	0.47	0.04	0.53	57136†	10056	33014	73738	-0.62	0.43	-0.12	0.83
08:00 - 09:00	38.99	16.50	-79.96	153.67	-0.40	0.47	0.06	0.50	60094†	9655	34323	76304	-0.70	0.41	-0.12	0.82
09:00 - 10:00	37.50	15.26	-76.02	150.10	-0.09	0.48	0.11	0.46	61582	8548	37430	76706	-0.72	0.41	-0.10	0.81
10:00 - 11:00	35.49	15.11	-67.08	151.07	0.08	0.50	0.14	0.43	63073	8086	40168	77791	-0.73	0.42	-0.10	0.79
11:00 - 12:00	34.70	14.86	-81.95	135.00	-0.02	0.52	0.16	0.40	64325	7723	41945	78207	-0.72	0.43	-0.10	0.79
12:00 - 13:00	32.18	14.84	-76.09	121.58	-0.41	0.51	0.16	0.37	63822	7700	41065	78327	-0.70	0.43	-0.11	0.79
13:00 - 14:00	30.49	16.23	-100.06	117.68	-1.13	0.50	0.14	0.38	62695	8152	39664	77079	-0.69	0.44	-0.12	0.79
14:00 - 15:00	29.75	16.70	-130.09	112.21	-1.36	0.50	0.16	0.43	61357	8250	38922	76720	-0.67	0.44	-0.12	0.79
15:00 - 16:00	30.92	15.84	-82.06	117.18	-0.49	0.55	0.20	0.49	60466	8163	38974	76839	-0.64	0.45	-0.09	0.80
16:00 - 17:00	32.83	15.34	-76.00	120.00	0.35	0.61	0.29	0.54	59896	7879	38515	78483	-0.54	0.49	-0.02	0.81
17:00 - 18:00	38.07	16.11	-6.00	142.78	1.48	0.68	0.38	0.57	60744	7948	33261	79063	-0.36	0.57	0.15	0.82
18:00 - 19:00	41.43	15.02	-1.21	143.09	1.75	0.63	0.33	0.51	61277	7633	39044	77741	-0.28	0.64	0.24	0.83
19:00 - 20:00	41.89	13.04	1.80	124.94	1.17	0.60	0.26	0.43	60764	7296	42362	76186	-0.27	0.64	0.22	0.84
20:00 - 21:00	38.29	10.57	-4.94	109.92	0.48	0.54	0.21	0.38	58257	6430	41725	72594	-0.28	0.63	0.15	0.83
21:00 - 22:00	34.62	9.24	-38.19	78.98	-0.71	0.52	0.19	0.33	55744	5589	39670	69863	-0.25	0.63	0.14	0.82
22:00 - 23:00	32.84	8.78	-49.98	66.17	-1.10	0.51	0.22	0.26	53412	4970	38079	68267	-0.06	0.71	0.30	0.81
23:00 - 24:00	27.84	9.90	-70.09	56.61	-3.02	0.48	0.12	0.15	49520	4624	35776	64729	0.10	0.76	0.41	0.80

Note: SD denotes sample standard deviation, $\rho^A(.)$ is the value of the autocorrelation coefficient at a given lag. Using the procedures developed in Hyndman et al., (2018), we used the Osborn et al., (1988), Hylleberg et al., (1990), and Canova and Hansen (1995) tests of weakly seasonal unit-root for each of the series. We omit the explicit reporting of results of these tests as they all indicated no seasonal unit-root in hourly electricity prices, except for series denoted with symbol † where the Canova and Hansen (1995) test indicated that one seasonal differencing is necessary to remove the seasonal unit-root.

5.2 Actual and forecasted energy demand

As previously discussed, demand and supply are important in the electricity price formation process, and their components should be included in our model. The supply side is determined by several factors: fuel for the power plants, emission allowances and production of renewable energy (Paraschiv et al, 2014). Electricity consumption represents the demand side. As the demand is almost inelastic, the day-ahead electricity prices are strongly affected by unscheduled plant outages (Bunn et al, 2013). A clear understanding of the underlying factors is important in developing insights into the electricity price. The demand variable will be represented by the aggregated electricity consumption in Germany, DA_t^h. As with the price time series, the demand is positively autocorrelated with even more profound seasonal pattern.

However, in almost all price series the seasonal unit-root tests indicated no presence of the unit-root[5]. To proxy for the expected demand, we use the day-ahead total load forecast obtained from the Transparency Platform operated by the European Network of Transmission System Operators for Electricity (ENTSO-E, 2018), DF_t^h. For markets where system operator does not provide load forecast, a model such as Do et al. (2016b) could be used.

5.3 Renewable energy

The biggest share of renewable production in Germany consists of wind and solar power. Moreover, production of these two renewables is price inelastic. We therefore focus only on these two renewables in our paper and use the term renewables interchangeably with wind and solar power. Woo et al. (2011) and Keles et al. (2013), employ econometric techniques to investigate the impact of wind on the electricity price in the Netherlands and German markets. Both papers find that wind production has reduced the electricity prices.

Installed capacity and production from renewable sources have increased in the recent years. We therefore expect the production of energy from renewables to have a negative impact on the electricity price in Germany. For illustration purposes, the following Figure 4 shows the time evolution of the share of wind and solar energy production on the overall electricity production. The average value over our sample period is at 0.22, but occasionally it reached values over 0.50, thus making it a potentially important price factor.

[5] An exception was found for three price series, where the Canova and Hansen's (1995) test indicated a seasonal-unit root, but the remaining tests no.

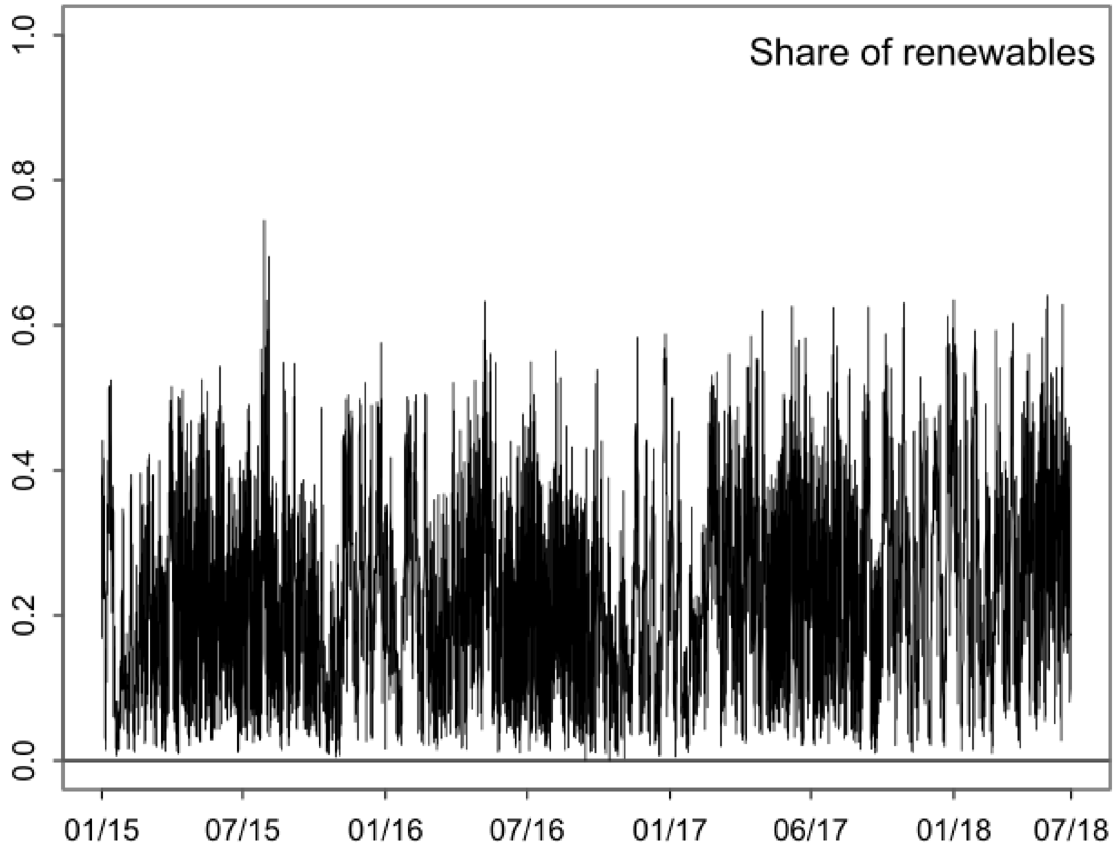

Figure 4 Share of the wind and solar energy production

The summary statistics presented in Table 2 cover production forecasts for electricity generated from wind and solar sources (ENTSO-E, 2018). Using forecasts is advantageous in a similar fashion as using demand forecasts: the forecasts are exogenous with respect to the day-ahead price data, the forecasts are publicly available and manifest market participant's expectations that are not captured in the actual lagged production. The descriptive statistics reported in Table 2 confirm some stylized facts about energy production from wind and solar renewable sources. For the wind energy, the hourly price series are most of the time quite similar, while for the solar energy the price path show 'typical' features of no production during night, while the production peaks around noon. Compared to the consumption (demand), in production, there is no weakly seasonal effect to be expected, nor visible in the autocorrelation structure. Using the unit-root tests and the Sul et al., (2005) version of the Kwiatkowski et al., (1992) stationarity test the null of no unit-root was also not rejected.

Table 2 Descriptive statistics for wind and solar forecasts

	Wind production forecast, WF_t^h [MWh]								Solar production forecast, SF_t^h [MWh]							
Hour	Mean	SD	Min	Max	Skew.	$\rho^A(1)$	$\rho^A(5)$	$\rho^A(7)$	Mean	SD	Min	Max	Skew.	$\rho^A(1)$	$\rho^A(5)$	$\rho^A(7)$
00:00 - 01:00	10642	7769	738	40938	1.14	0.59	0.20	0.17	--	--	--	--	--	--	--	--
01:00 - 02:00	10661	7802	576	39861	1.14	0.59	0.20	0.18	--	--	--	--	--	--	--	--
02:00 - 03:00	10613	7809	574	38827	1.13	0.59	0.20	0.18	--	--	--	--	--	--	--	--
03:00 - 04:00	10466	7774	0	38842	1.12	0.59	0.20	0.18	--	--	--	--	--	--	--	--
04:00 - 05:00	10424	7754	730	38383	1.11	0.59	0.20	0.18	--	--	--	--	--	--	--	--
05:00 - 06:00	10336	7760	846	38305	1.12	0.60	0.20	0.18	41	94	0	1663	5.51	0.73	0.70	0.68
06:00 - 07:00	10232	7802	781	38139	1.13	0.60	0.20	0.19	381	578	0	2132	1.39	0.95	0.92	0.90
07:00 - 08:00	10092	7892	618	38046	1.15	0.61	0.21	0.19	1626	1763	0	6245	0.80	0.93	0.87	0.85
08:00 - 09:00	9929	8031	412	37995	1.16	0.61	0.21	0.19	4108	3338	4	12149	0.44	0.89	0.79	0.77
09:00 - 10:00	9873	8234	314	38126	1.15	0.61	0.20	0.18	7269	4750	283	18356	0.28	0.85	0.71	0.69
10:00 - 11:00	9954	8408	292	38602	1.14	0.61	0.18	0.16	10141	5800	733	23490	0.19	0.82	0.66	0.63
11:00 - 12:00	10101	8533	293	38990	1.12	0.61	0.16	0.14	12162	6503	1116	26908	0.14	0.82	0.64	0.61
12:00 - 13:00	10271	8595	334	39191	1.10	0.60	0.15	0.13	13091	6946	1190	28651	0.11	0.83	0.65	0.62
13:00 - 14:00	10401	8584	366	38933	1.08	0.60	0.14	0.11	12897	7153	982	28817	0.12	0.84	0.68	0.66
14:00 - 15:00	10445	8486	436	38778	1.07	0.59	0.13	0.11	11697	7149	620	27511	0.14	0.87	0.73	0.71
15:00 - 16:00	10414	8347	381	38879	1.07	0.57	0.13	0.11	9629	6851	178	24753	0.17	0.90	0.79	0.77
16:00 - 17:00	10385	8201	398	39164	1.08	0.56	0.14	0.11	7109	6016	7	20402	0.26	0.92	0.84	0.82
17:00 - 18:00	10360	8060	463	39619	1.09	0.56	0.16	0.13	4569	4499	0	14735	0.40	0.94	0.87	0.86
18:00 - 19:00	10329	7938	545	40062	1.10	0.57	0.18	0.16	2362	2665	0	8873	0.60	0.95	0.90	0.90
19:00 - 20:00	10319	7855	610	40423	1.12	0.58	0.21	0.19	880	1176	0	3921	1.01	0.97	0.93	0.93
20:00 - 21:00	10410	7818	651	40605	1.13	0.58	0.23	0.20	197	344	0	2130	1.76	0.95	0.92	0.91
21:00 - 22:00	10536	7757	672	40514	1.13	0.59	0.23	0.20	12	56	0	1731	23.74	0.22	0.21	0.20
22:00 - 23:00	10667	7706	833	40292	1.13	0.59	0.22	0.19	--	--	--	--	--	--	--	--
23:00 - 24:00	10703	7659	855	39895	1.13	0.59	0.21	0.18	--	--	--	--	--	--	--	--

Note: SD denotes sample standard deviation, $\rho^A(.)$ is the value of the autocorrelation coefficient at a given lag. Using the procedures developed in Hyndman et al., (2018) we used the Osborn et al., (1988), Hylleberg et al., (1990), and Canova and Hansen (1995) tests of weakly seasonal unit-root for each of the series. We have also used the Sul et al. (2005) version of the Kwiatkovski et al. (1992) test indicated rejection of the null hypothesis of a unit-root. We do not report results from these tests explicitly as for all series the results indicated no presence of seasonal unit-root.

5.4 Market prices: Coal, Brent, Natural Gas, EUA

According to Sensfuß et al. (2008) different means of electricity generation have distinct fuel price dependencies; for example, coal power plants are dependent on coal prices and gas power plants are dependent on gas prices. Mjelde and Bessler (2009) study two US electricity markets and include uranium prices along with other fuel prices. Ferkingstad et al (2011) use a cointegration approach on prices in Northern Europe and find that electricity prices are closely connected with gas prices, while coal and oil prices are less important. However, Parashiv et al

(2014) studies the German electricity price and finds that coal, gas and oil are all important fundamental variables driving the electricity price, with coal price more important during off peak hours and gas and oil prices important during peak hours. This finding is in line with Murray (2009), who found that the relationship between fuel prices and electricity price is depended on the marginal electricity price setting technology used in each specific hour.

Usually, nuclear power plants run at almost constant power for economic reasons, even when the load is low (International Atomic Energy Agency, 1999). Nuclear power plants have low marginal costs on the merit order curve. They have therefore very small impact on the variation of electricity price. Coal and lignite are the primary fuels (45% in 2013) used to cover the base load for the electricity market (AG Energiebilanzen, 2015). Unlike hard coal, lignite is based on local distribution and there is currently no single market price for lignite. The coal price is represented by a futures contract on the price of coal imported to northwestern Europe via Amsterdam, Rotterdam and Antwerp. Unlike coal power plants, gas power plants are mainly used to cover the peak load due to their greater flexibility to ramp up and down. In Germany, gas is traded under contracts from Gaspool and NetConnect Germany (NCG). We choose to use the NCG contracts because there is higher liquidity in this market. Production from oil power plants forms a small fraction (1%) of Germany's total electricity production (AG Energiebilanzen, 2015). Oil price has therefore a low impact on the merit order curve (Sensfuß et al., 2008). Further, oil consumption is dominated by the transport sector and industry. However, the oil price might serve as a proxy for economic activity and transport fuel for coal fuel. In this paper, the European Brent spot price is used to represent the oil fuel cost. Based on the arguments above, we include coal, gas and oil prices in our model.

The CO_2 markets are an attempt to increase investment in cleaner technology by fuel switching to less carbon intensive power plants or reducing the use of carbon intensive power plants. Both Fell (2010) and Parashiv et al (2014) find that the short-term influence of CO_2 price on the electricity price is higher in off peak hours than in peak hours. This is because coal emits twice as much CO_2 as natural gas. We therefore expect the CO_2 price to have a higher effect on the electricity price during periods when coal power plants are the price setting technology.

The following Table 3 presents the summary statistics of the market prices of the Brent oil, coal, natural gas and EUA. In line with the existing literature, the unit-root test rejected the null of no unit-root in the level of the series. Therefore in the subsequent analysis, we use the 1st differences as they the test was unable to reject the null of no-unit root. Note, that using these market prices in level form is not advised as they are integrated of order one, while electricity prices (dependent variable) is integrated of order zero. A linear combination would potentially lead to a variable (within a linear regression model to residuals) that are integrated of order one.

Table 3 Descriptive statistics for the prices of Brent oil, Coal, Natural gas and EUA

		Mean	SD	Min	Max	Skew.	$\rho^A(1)$	$\rho^A(5)$	$\rho^A(7)$
Panel A: Level series									
Brent	B_t	53.53†	9.85	27.88	79.80	0.17	0.993	0.967	0.955
Coal	C_t	69.18†	16.49	43.40	96.65	0.08	0.998	0.989	0.984
Gas	G_t	42.33†	8.23	26.38	66.31	0.02	0.993	0.965	0.953
EUA	E_t	6.94†	2.29	3.93	16.28	1.66	0.992	0.960	0.942
Panel B: Differenced series									
Brent	ΔB_t	0.02	0.98	-4.59	5.46	0.26	-0.081	-0.044	-0.003
Coal	ΔC_t	0.02	0.77	-6.20	11.95	2.63	0.062	0.014	-0.025
Gas	ΔG_t	0.00	0.87	-6.50	6.50	-0.14	0.036	-0.001	-0.023
EUA	ΔE_t	0.01	0.16	-0.93	1.10	0.47	-0.042	0.019	-0.061

Note: SD denotes sample standard deviation, $\rho^A(.)$ is the value of the autocorrelation coefficient at a given lag. Symbol † denotes series where the Sul et al. (2005) version of the Kwiatkovski et al. (1992) test indicated rejection of the null hypothesis of a unit-root.

6. Empirical results

6.1 Baseline results from linear models

Similar to the demand model, we generate a multiple regression model for the price. The reason for using a separate equation is that each hour displays a rather distinct price profile, reflecting the daily variation in demand, fuel costs and operational constraints (Chen and Bunn, 2010). Furthermore, extensive research on price forecasting has generally favored the multi-model specification for short-term predictions (Chen and Bunn, 2010; Paraschiv et al., 2014).

Based on the description of the electricity market in Germany given in Section 2 and on data availability, we estimated 24 separate linear regression models to estimate electricity prices in Germany. To save space, Table 4 we reports estimated coefficients from Eq. (1) for four selected hours of the day, results for the remaining 20 hours are available upon request. In order to give a more comprehensive picture of results for all hours, we provide a graphic representation on the estimated coefficient from the ordinary regression in Figure 5 and 6.

Table 4 Baseline OLS results explaining electricity prices for selected hours

		08:00 - 09:00	12:00 - 13:00	19:00 - 20:00	00:00 - 01:00
Panel A: Variables of interest					
Constant		-1.632	3.968	-14.889*	-15.964***
Lagged price	P_{t-1}	0.261***	0.218***	0.332***	0.175***
Previous week's price	P_{t-7}	0.165***	0.106***	0.167***	0.085***
Forecasted demand/1000	DF_t	0.150***	0.100***	0.050*	0.020*
Lagged demand/1000	Da_t	0.001	0.070**	0.130***	0.230***
Lagged coal price return	ΔC_{t-1}	-0.099	0.011	-0.06	-0.219
Lagged natural gas price return	ΔG_{t-2}	0.164	0.098	0.013	0.292
Lagged Brent oil price return	ΔB_{t-3}	-0.348*	-0.081	-0.018	0.076
Lagged EUA price return	ΔE_{t-4}	-0.330	1.067	1.904	0.466
Wind production forecast/1000	WF_t	-0.240***	-0.240***	-0.220***	-0.230***
Solar production forecast/1000	SF_t	-0.130***	-0.230***	0.060	
Panel B: Control variables					
Monday	Mon_t	4.597	6.517***	9.933***	3.076***
Tuesday	Tue_t	0.673	0.373	0.621	4.097***
Thursday	Thu_t	-0.514	-1.100**	-1.251*	-0.286
Friday	Fri_t	-1.651***	-1.34**	-2.906***	-0.627*
Saturday	Sat_t	-5.239**	-4.784***	-4.333***	-0.393
Sunday	Sun_t	-5.647*	-4.019**	0.771	-3.546***
January	Jan_t	6.226**	-1.089	4.851*	-2.473
February	Feb_t	1.903	-4.771**	1.959	-3.839**
March	Mar_t	1.035	-3.948**	2.674	-3.770**
April	Apr_t	2.747*	-0.544	1.433	-0.302
May	May_t	1.550	-0.058	0.799	-0.51
Jun	Jun_t	-0.293	-1.360	-1.087	-1.621
August	Aug_t	-1.089	-1.347	0.362	-1.188
September	Sep_t	1.732	-1.946	2.804	-1.050
October	Oct_t	3.810**	-2.106	5.120**	-0.779
November	Nov_t	6.043***	-1.356	4.563**	-0.004
December	Dec_t	3.681**	-2.520	4.322**	-0.971
Holiday	Hol_t	-8.335**	-6.687**	-2.046*	-1.251
R^2		0.747	0.721	0.711	0.678
1st order autocorrelation of residuals		0.026	0.138	0.126	0.158
7th order autocorrelation of residuals		0.074	0.111	0.162	0.116

*Note: *, **, *** denotes significance of coefficients at the 10%, 5%, and 1% level. The significance of coefficients is based on the estimation of standard errors of regression coefficients that are consistent even in the presence of autocorrelated and heteroscedastic residuals, using the Newey and West (1994) automatic bandwidth procedure with Quadratic spectral weighting scheme. Autocorrelation of residuals was tested using the Peña and Rodríguez (2002) test using the Monte Carlo version of Lin and McLeod (2006). R^2 denotes the coefficient of determination.*

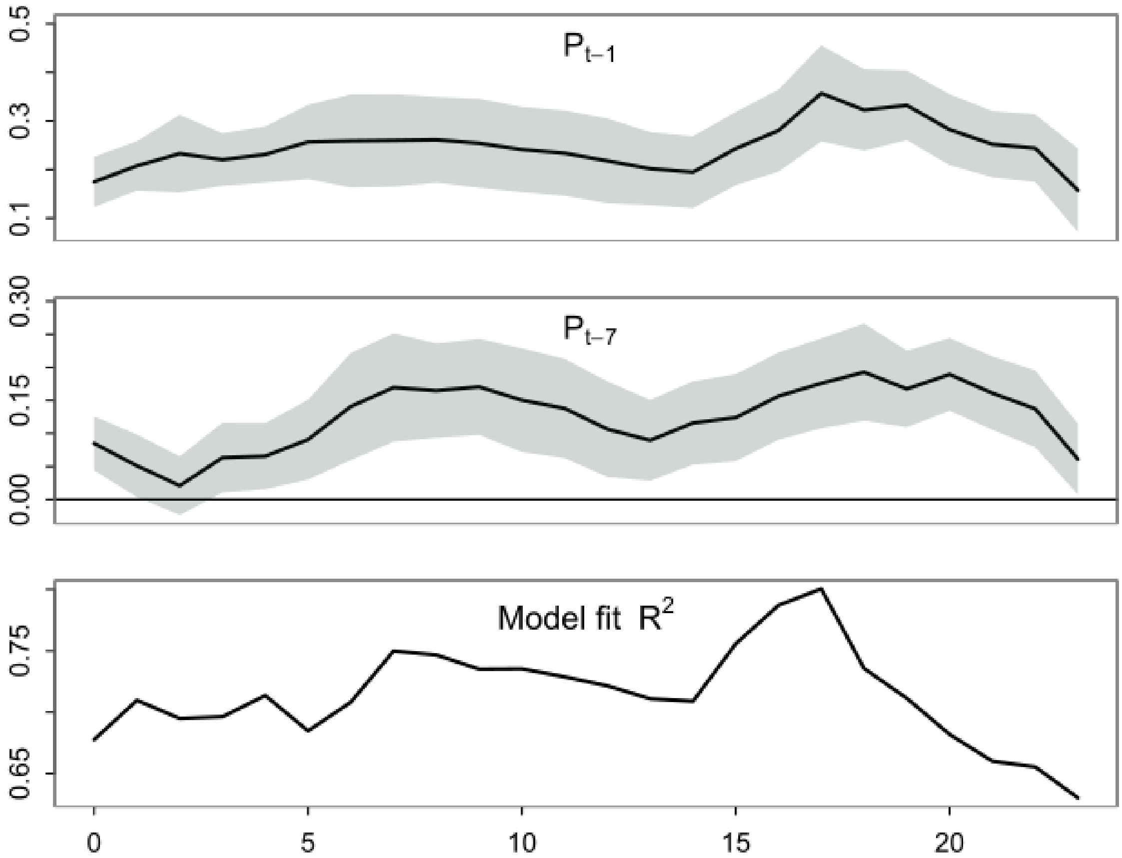

Figure 5 Estimated OLS coefficients of price models across different daily hours: Lagged prices and model fit

Note: The shaded area represents the 95% confidence interval around the estimated coefficients.

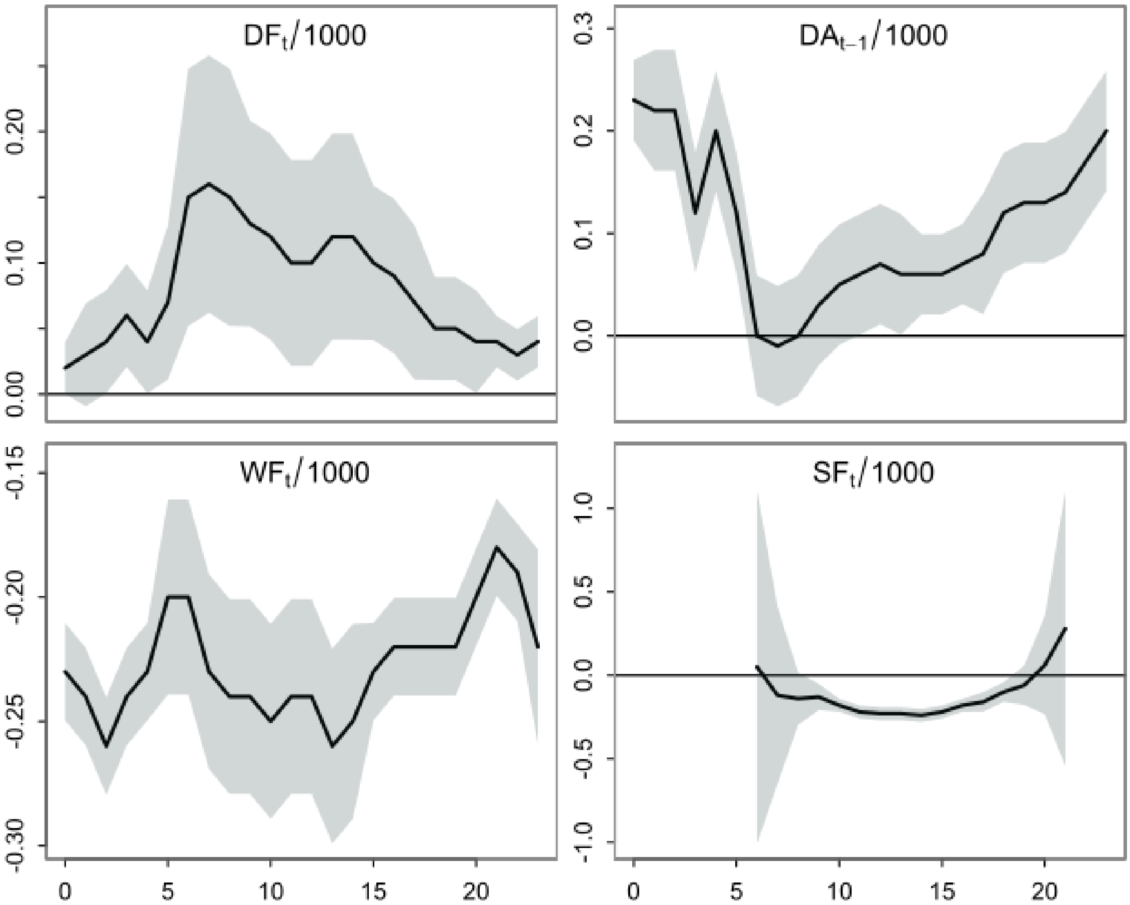

Figure 6 Estimated OLS coefficients of price models across different daily hours: Demand and production from renewable energy sources
Note: The shaded area represents the 95% confidence interval around the estimated coefficients.

Our results (see Table 4 and Figure 5) confirm the existence of persistence in the electricity prices in Germany. The coefficient (at $P_{t\text{-}1}$) ranges from 0.157 (for the 23:00 to 00:00 period) to 0.357 (18:00 to 19:00). We have also introduced the lagged weakly price variable to account for weakly seasonality. The coefficients (at $P_{t\text{-}7}$) was positive and significant for almost all daily prices all models. Although price persistence of electricity prices is present across all pricing equations, the size of the persistence differs across the day, as higher persistence (daily and weekly) is found late afternoon (from 15:00 to 20:00). These results suggest that prices late afternoon, when the demand peaks are more predictable. Supply side explanation for this effect would be that later afternoon production runs closer to the overall available production capacity.

Therefore sudden changes in demand have a relatively low effect on the supply side and thus the electricity price.

Even though the demand is considered to be mostly inelastic (Sensfuß et al, 2008), we find, that the variation in the actual and forecasted demand is positively associated with next day's prices. However, the magnitude of the effect that actual and forecasted demand exert on the prices over the day differs quite considerably. The forecasted demand is most important for electricity prices in the morning (from 06:00 to 11:00), while actual one-day lagged demand in the midnight (from 23:00 to 02:00).

The effects of renewables on German electricity prices are estimated using exogenous terms for the total expected (forecasted) electricity production from wind and solar resources. Coefficients on the forecasted electricity production from renewable resources are negative, thus confirming earlier finding that increased production through renewables decreases electricity prices. This effect can be attributed to the minimal marginal costs associated with higher production.

The infeed from solar production has a lower impact on the electricity price than wind production, which can be explained by higher total installed wind capacity compared to that of solar sources (AG Energiebilanzen, 2015). We further observe that the solar production coefficient is smaller (higher effect) during the day and almost zero as during the night, in accordance with the solar production level.

The role of the wind production in setting prices declines during production peaks mornings and late afternoon up until midnight. The possible explanation is that when the demand for electricity is high, flexible gas power plants are required to cover the demand, which on the other hand have high marginal costs. Using the additional capacity of such high cost power plants decreases the price lowering effect from renewables.

Price of coal, natural gas and Brent oil in their differenced form are statistically unrelated to the electricity prices (see Table 4, results for remaining hours lead to the same conclusion). This might seem to be surprising given the fact that energy prices are driving costs for non-renewable energy sources (as well as opportunity costs for renewable energy sources). Also European Union emission allowances have no relevant effect on the day-ahead electricity prices. However, as we use price differences (due to the inability to reject the null of stationarity of the price series) in the analysis, the reported result only suggest that short-term (daily) changes in market prices on the right hand side of our model, does not lead electricity prices (in level form) on the German market. This is in line with the fact how commodities are purchased – contracts (and prices) are settled several months in advance, therefore daily changes should in fact only limited effect on day-ahead electricity prices, which is what our results suggest.

We also included a set of control variables in a form of a dummy variables (see Table 4). The results indicate significant and relevant day-of-the-week effects and monthly effects. For example, compared to Wednesdays, the average price on Monday during the peak hours (19:00 - 20:00) is 9.93 EUR higher. While compared to July, prices for the same hour in October are on average up 5.12 EUR. We can also observe that day-of-the-week and monthly effects are more relevant during peak hours.

6.2 Results from ARX-EGARCH model

We complement the linear regression results with those from the ARX(7,0)-EGARCH(1,1) model. Fundamental variables show similar magnitude and significance as those reported via the linear regression. Detailed results for selected hours can be found in Appendix A. However, few new results are worth to mention. First, the inclusion of one to seven daily price lags (not just one day and one week price lags) removed traces of serial correlation in residuals, and also led to less pronounced day-of-the-week effects. Still, prices during the weekend are lower compared to the benchmark of Wednesday's prices. Second, the residuals show signs of conditional heteroscedasticity and the presence of the sign effects that vary with respect to the hourly price being modelled. For example, during the night, when negative price shocks are more likely we observe a leverage effect, where negative price shocks lead to an increase of volatility of higher magnitude than comparable positive price shocks. On the other hand, during the production peak hours, when positive price shocks are more likely, we observe the opposite effect (positive sign coefficient).

Overall, our models are able to explain the electricity prices better during the day, particularly when demand is high (see Figure 5). We attribute this to the: i) stronger persistence of prices during peak hours, ii) the stronger (and positive) effect of the demand, and iii) more relevant seasonal effects. The impact of the renewable energy sources on the electricity price is negative, i.e. higher expected production from wind and solar renewable sources leads lower electricity prices. One explanation for lower model fit and smaller effects of fundamentals during the night and off-peak hours is that prices are subject to extreme swings during the night (see Figure 3), because of low demand and excess production from wind sources (see Paraschiv et al., 2014). In the next section, we address this issue by studying how key fundamental drivers effect the tails of electricity prices.

6.3 Modeling extreme prices

Quantile regression enables us to estimate a set of regression lines, corresponding to selected quantiles of the electricity price distribution. Using the quantile regression framework allows us to compare results across different quantiles (and compared to the OLS estimators) and thus observe potential asymmetric and non-linear effects on the electricity price. In the following models, we report results on the 5^{th}, 25^{th}, 50^{th}, 75^{th} and 95^{th} percentiles of the electricity prices. Focusing on tails of the price distribution, 5^{th} and 95^{th}, percentiles is useful for our understanding of drivers of extreme prices that also tend to have the largest effect on all market participants: suppliers, consumers and regulators alike.

Since we estimate coefficients for five quantiles for each variable for 24 hours of the day, that lead to over 3600 coefficients, thus presenting all the coefficients in tables would be cumbersome. We therefore summarize results for key variables in following Figures 7 – 8.

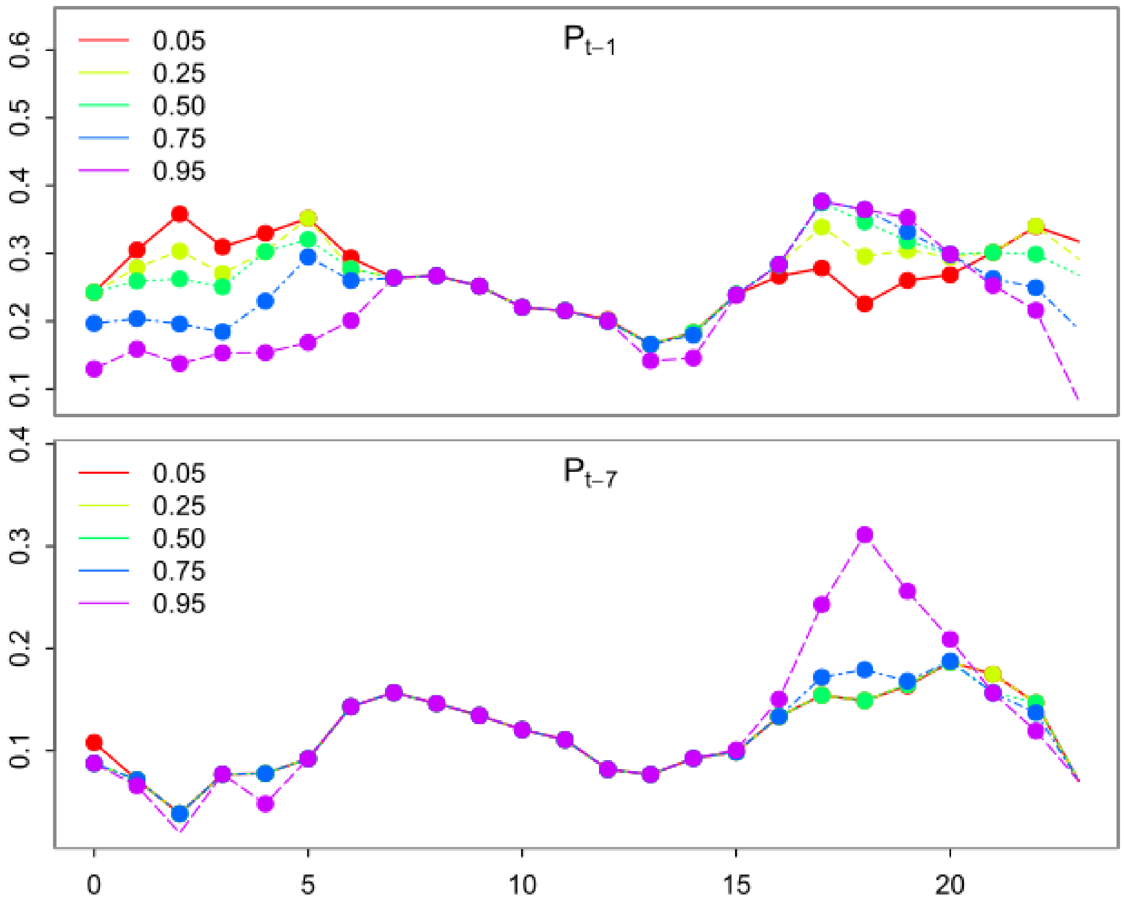

Figure 7 Estimated quantile regression coefficients across different daily prices and quantiles: Lagged prices

Note: The dot in the line corresponds to a statistically significant coefficient at the 5% significance level.

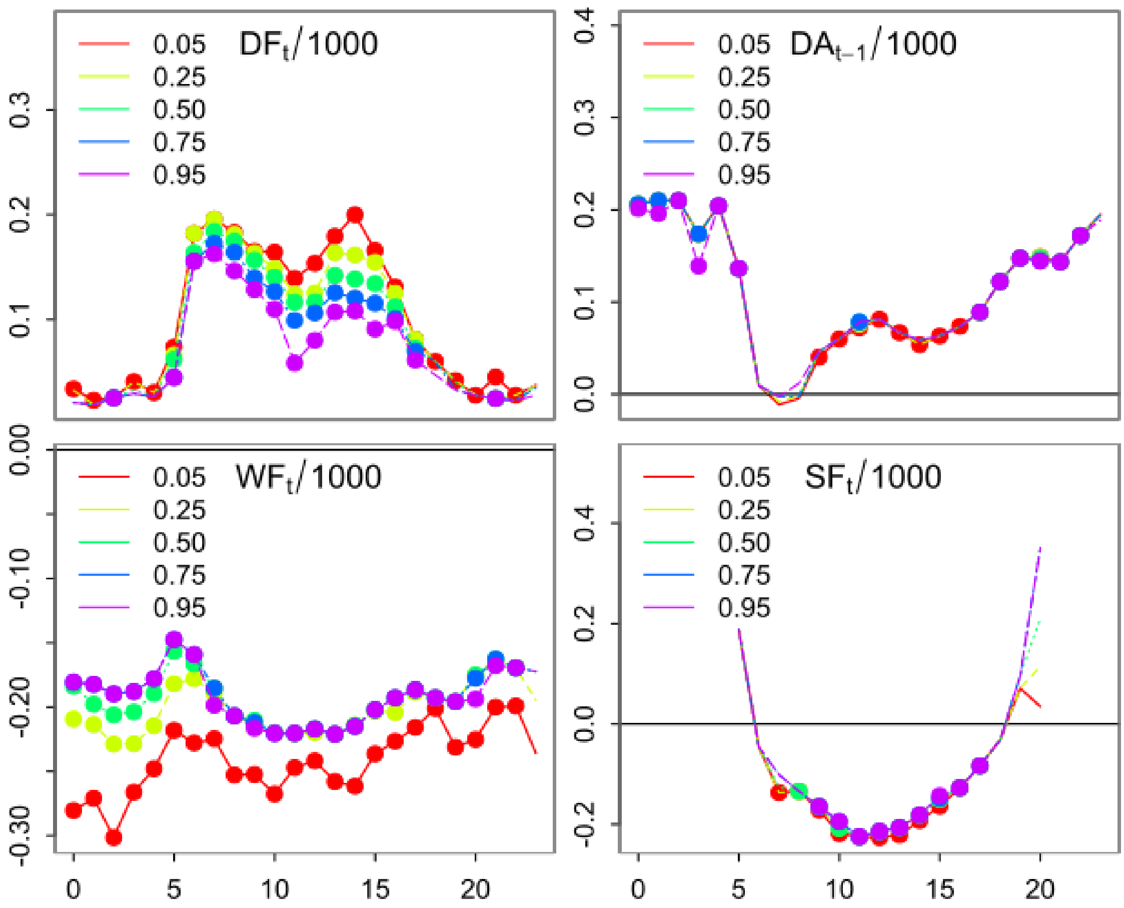

Figure 8 Estimated quantile regression coefficients across different daily prices and quantiles: Demand and production from renewable energy sources

Note: The dot in the line corresponds to a statistically significant coefficient at the 5% significance level.

In Figures 7 and 8 we report results only from fundamental variables that systematically influenced prices. Similarly, as for the OLS model, market prices in their differenced form had no relevant effect on the percentiles of the day-ahead electricity prices in Germany and were therefore excluded from the Figure 7 and 8.

In Figure 7 (upper panel), we can observe several notable asymmetric effects of the persistence coefficients (at P_{t-1}). For example, during the night (from 02:00 until 06:00) the persistence of extremely low prices (below 5th percentile) is much stronger than persistence of high prices (above the 95th percentile). On the other hand, during the late afternoon production peak (from 17:00 until 21:00) the persistence of extremely low prices is much smaller. This shows, that at least part of the variation in the extremely negative prices can be attributed to different persistence of prices across quantiles and hours.

The weekly seasonality price coefficient (at P_{t-7}) tend to be of lower magnitude than the daily lagged coefficient (lower panel of Figure 6), but is still significant across most quantiles and hours. The effect across quantiles of the electricity prices shows similar patterns. An exception is found for the production peak period (17:00 until 21:00), where weekly lagged price coefficient reaches values above 0.20 for the 95th percentile of the electricity price.

The importance of the estimated coefficients of forecasted demand differs across percentiles. The positive relationship between forecasted demand and electricity price is greater in magnitude at the lower tail (5th and 25th percentiles) of the distribution than at the upper tail (75th and 95th quantiles). The differences are amplified during production peak period in the morning and early afternoon. For example, with respect to the price during the period from 14:00 to 15:00 the effect of the forecasted demand at the 95th percentile of the price is 0.108 (after multiplying by 1000) and at the 5th percentile it is 0.200, while both significant.

The effect of the actual lagged demand on the prices differs particularly with respect to given hour, while it is similar across quantiles. During the night, when production/demand is low the lagged demand shows strong effect across all quantiles of the price distribution, while during the day, the effect of lagged demand is much smaller and before noon even insignificant.

The effect of expected (forecasted) renewable electricity production resources show several interesting features. First note that the effect of the solar power sources over the relevant daily hours is very similar across quantiles. Increasing level of production leads to lower prices across the whole distribution. Therefore solar power sources does not seem to amplify extreme prices. On the other hand, wind production shows that it indeed has the ability to shift extreme prices in both direction. Note, that during the whole day coefficients related to the 5th price percentile, that explain extremely low electricity prices, are much lower (larger in magnitude) compared to the coefficients on remaining percentiles. At the same time, coefficients for the upper tail of the price distribution, although smaller in magnitude, are also still significant. This means that although wind production shifts left-tail price distribution in a more pronounced way, wind production in general drives both, the right- and left-tail of the price distribution. During the night session (from 22:00 to 06:00) the effect of the wind power production is amplified with respect to extremely low prices.

7. Conclusions

This paper studies the main drivers of the electricity prices in Germany over the period from January 2015 until Jun 2018. We analyze hourly day-ahead prices using standard linear regression, ARX-EGARCH, and a (non-quantile crossing) quantile regression models. Our contribution to the existing literature is five-fold.

First, we find that short-term price fluctuations of the coal, Brent oil, natural gas and emission allowances (EUA) do not impact the German day-ahead electricity prices. Second, we show that controlling day-of-the-week as well as monthly effects is necessary for a correct specification as these effects are not only statistically significant but also economically relevant price drivers. Third, we identify that extremely low and high price movements on the market are driven partly by price persistence. The extremely low prices (below the 5^{th} percentile) tend to persist in the prices for the night session. At the same time, extremely high prices (above the 95^{th} percentile) tend to persist in the prices settled during the late afternoon production peak period. Fourth, extremely low prices are driven by expected demand. When the demand drops during the day, the impact on the left-tail of the price distribution is much higher compared to the impact on the right-tail of the price distribution. Fifth, we identify that forecasted wind production effects both, extremely high and low price, although low prices are more affected. The difference in the effect of the wind production on electricity prices is amplified during the night session, where low prices are smaller with larger increase of wind production.

The results presented in this study could be used by market participants to understand what drive extreme price movements, i.e. that extreme prices are at least partly driven by market fundaments like demand and production from renewable energy sources. For example, risk managers could design better trading strategies to mitigate effects of extreme negative price movements on their portfolios.

References

AG Energiebilanzen, (2015) Bruttostromerzeugung in Deutschland von 1990 bis 2014 nach Energieträgern. Downloaded on 20/04/2015 from ⟨http://www. ag-energiebilanzen.de/⟩.

Bello, A., Bunn, D. W., Reneses, J., & Muñoz, A. (2017). Medium-term probabilistic forecasting of electricity prices: A hybrid approach. *IEEE Transactions on Power Systems*, *32*(1), 334-343.

Bondell, H. D., Reich, B. J., & Wang, H. (2010). Noncrossing quantile regression curve estimation. Biometrika, 97(4), 825-838.

Bunn, D. W., & Chen, D. (2013). The forward premium in electricity futures. *Journal of Empirical Finance*, *23*, 173-186.

Bunn, D., Andresen, A., Chen, D., & Westgaard, S. (2016). Analysis and Forecasting of Electricty Price Risks with Quantile Factor Models. *The Energy Journal*, *37*(1), 101-122.

Canova, F., & Hansen, B. E. (1995). Are seasonal patterns constant over time? A test for seasonal stability. *Journal of Business & Economic Statistics*, *13*(3), 237-252.

Chen, D., & Bunn, D. W. (2010). Analysis of the nonlinear response of electricity prices to fundamental and strategic factors. *IEEE Transactions on Power Systems*, *25*(2), 595-606.

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.

Conejo, A. J., Contreras, J., Espinola, R., & Plazas, M. A. (2005). Forecasting electricity prices for a day-ahead pool-based electric energy market. *International Journal of Forecasting*, *21*(3), 435-462.

Dillig, M., Jung, M., & Karl, J. (2016). The impact of renewables on electricity prices in Germany–An estimation based on historic spot prices in the years 2011–2013. *Renewable and Sustainable Energy Reviews*, *57*, 7-15.

Do, L. P. C., Hagfors, L. I., Lin, K. H., & Molnár, P. (2016a). Demand and residual demand modelling using quantile regression. *Renewable Energy and Environmental Sustainability*, *1*, 41.

Do, L. P. C., Lin, K. H., & Molnár, P. (2016b). Electricity consumption modelling: A case of Germany. *Economic Modelling*, *55*, 92-101.

ENTSO-E (2018). Transparency Platform.<https://transparency.entsoe.eu>.

Fell, H. (2010). EU-ETS and Nordic electricity: a CVAR analysis. *The Energy Journal*, *31*(2), 1-25.

Ferkingstad, E., Løland, A., & Wilhelmsen, M. (2011). Causal modeling and inference for electricity markets. *Energy Economics*, *33*(3), 404-412.

Green, R., & Vasilakos, N. (2010). Market behaviour with large amounts of intermittent generation. *Energy Policy*, *38*(7), 3211-3220.

Hagfors, L. I., Bunn, D., Kristoffersen, E., Staver, T. T., & Westgaard, S. (2016a). Modeling the UK electricity price distributions using quantile regression. *Energy*, *102*, 231-243.

Hagfors, L. I., Kamperud, H. H., Paraschiv, F., Prokopczuk, M., Sator, A., & Westgaard, S. (2016b). Prediction of extreme price occurrences in the German day-ahead electricity market. *Quantitative Finance*, *16*(12), 1929-1948.

Hagfors, L. I., Paraschiv, F., Molnár, P., & Westgaard, S. (2016c). Using quantile regression to analyze the effect of renewables on EEX price formation. *Renewable Energy and Environmental Sustainability*, *1*, 32.

Hammoudeh, S., Nguyen, D. K., & Sousa, R. M. (2014). Energy prices and CO2 emission allowance prices: A quantile regression approach. *Energy Policy*, *70*, 201-206.

He, Y., Qin, Y., Wang, S., Wang, X., & Wang, C. (2019). Electricity consumption probability density forecasting method based on LASSO-Quantile Regression Neural Network. *Applied Energy*, 233, 565-575.

Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, *32*(3), 914-938.

Hylleberg, S., Engle, R. F., Granger, C. W., & Yoo, B. S. (1990). Seasonal integration and cointegration. *Journal of econometrics*, *44*(1-2), 215-238.

Hyndman R, Athanasopoulos G, Bergmeir C, Caceres G, Chhay L, O'Hara-Wild M, Petropoulos F, Razbash S, Wang E, Yasmeen F (2018). *forecast: Forecasting functions for time series and linear models*. R package version 8.4, http://pkg.robjhyndman.com/forecast.

International Atomic Energy Agency (1999), Modern Instrumentation and Control for Nuclear Power Plants: A Guidebook, Technical Report Series N°387.

Jacobsen, H. K., & Zvingilaite, E. (2010). Reducing the market impact of large shares of intermittent energy in Denmark. *Energy Policy*, *38*(7), 3403-3413.

Janssen, M., & Wobben, M. (2009). Electricity pricing and market power—evidence from Germany. *European Transactions on Electrical Power*, *19*(4), 591-611.

Johnson, N.L. (1949a). Systems of Frequency Curves Generated by Method of Translation. *Biometrika* 36(1/2), 149–176.

Johnson, N.L. (1949b). Bivariate Distributions Based on Simple Translation Systems. *Biometrika* 36(3/4), 297–304.

Jónsson, T., Pinson, P., & Madsen, H. (2010). On the market impact of wind energy forecasts. *Energy Economics*, *32*(2), 313-320.

Jónsson, T., Pinson, P., Madsen, H., & Nielsen, H. A. (2014). Predictive densities for day-ahead electricity prices using time-adaptive quantile regression. *Energies*, *7*(9), 5523-5547.

Juban, R., Ohlsson, H., Maasoumy, M., Poirier, L., & Kolter, J. Z. (2016). A multiple quantile regression approach to the wind, solar, and price tracks of GEFCom2014. *International Journal of Forecasting*, *32*(3), 1094-1102.

Karakatsani, N. V., & Bunn, D. W. (2010). Fundamental and behavioural drivers of electricity price volatility. *Studies in Nonlinear Dynamics & Econometrics*, 14(4).

Kaza, N. (2010). Understanding the spectrum of residential energy consumption: a quantile regression approach. *Energy policy*, *38*(11), 6574-6585.

Keles, D., Genoese, M., Möst, D., & Fichtner, W. (2012). Comparison of extended mean-reversion and time series models for electricity spot price simulation considering negative prices. *Energy Economics*, 34(4), 1012-1032.

Keles, D., Genoese, M., Möst, D., Ortlieb, S., & Fichtner, W. (2013). A combined modeling approach for wind power feed-in and electricity spot prices. *Energy Policy*, *59*, 213-225.

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

Kiesel, R., & Paraschiv, F. (2017). Econometric analysis of 15-minute intraday electricity prices. *Energy Economics*, *64*, 77-90.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. *Journal of econometrics*, *54*(1-3), 159-178.

Lee, C. C., & Zeng, J. H. (2011). The impact of oil price shocks on stock market activities: Asymmetric effect with quantile regression. *Mathematics and Computers in Simulation*, *81*(9), 1910-1920.

Lin, J. W., & McLeod, A. I. (2006). Improved Peňa–Rodriguez portmanteau test. *Computational Statistics & Data Analysis*, *51*(3), 1731-1738.

Lindström, E., & Regland, F. (2012). Modeling extreme dependence between European electricity markets. *Energy Economics*, 34(4), 899-904.
Liu, B., Nowotarski, J., Hong, T., & Weron, R. (2017). Probabilistic load forecasting via quantile regression averaging on sister forecasts. *IEEE Transactions on Smart Grid*, *8*(2), 730-737.
Maciejowska, K., & Nowotarski, J. (2016). A hybrid model for GEFCom2014 probabilistic electricity price forecasting. *International Journal of Forecasting*, *32*(3), 1051-1056.
Maciejowska, K., Nowotarski, J., & Weron, R. (2016). Probabilistic forecasting of electricity spot prices using Factor Quantile Regression Averaging. *International Journal of Forecasting*, *32*(3), 957-965.
Martinez-Anido, C. B., Brinkman, G., & Hodge, B. M. (2016). The impact of wind power on electricity prices. *Renewable Energy*, *94*, 474-487.
Mjelde, J. W., & Bessler, D. A. (2009). Market integration among electricity markets and their major fuel source markets. *Energy Economics*, *31*(3), 482-491.
Moreira, R., Bessa, R., & Gama, J. (2016). Probabilistic forecasting of day-ahead electricity prices for the Iberian electricity market. In *2016 13th International Conference on the European Energy Market (EEM)*, Porto, pp. 1-5. doi: 10.1109/EEM.2016.7521226
Mosquera-López, S., Uribe, J. M., & Manotas-Duque, D. F. (2017). Nonlinear empirical pricing in electricity markets using fundamental weather factors. *Energy*, *139*, 594-605.
Murray, B. (2009) Power markets and economics: energy costs, trading, emissions. Chichester, U.K., Wiley.
Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.
Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, *61*(4), 631-653.
Nowotarski, J., & Weron, R. (2014). Merging quantile regression with forecast averaging to obtain more accurate interval forecasts of Nord Pool spot prices. In *11th International Conference on the European Energy Market (EEM14)*, Krakow, pp. 1-5. doi: 10.1109/EEM.2014.6861285
Nowotarski, J., & Weron, R. (2015). Computing electricity spot price prediction intervals using quantile regression and forecast averaging. *Computational Statistics*, *30*(3), 791-803.
Nowotarski, J., & Weron, R. (2018). Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable and Sustainable Energy Reviews*, *81*, 1548-1568.
Osborn, D. R., Chui, A. P., Smith, J. P., & Birchenhall, C. R. (1988). Seasonality and the order of integration for consumption. *Oxford Bulletin of Economics and Statistics*, *50*(4), 361-377.
Paraschiv, F., Erni, D., & Pietsch, R. (2014). The impact of renewable energies on EEX day-ahead electricity prices. *Energy Policy*, *73*, 196-210.
Peña, D., & Rodríguez, J. (2002). A powerful portmanteau test of lack of fit for time series. *Journal of the American Statistical Association*, *97*(458), 601-610.
Portnoy, S., & Koenker, R. (1997). The Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute-error estimators. *Statistical Science*, *12*(4), 279-300.
Sapio, A. (2019). Greener, more integrated, and less volatile? A quantile regression analysis of Italian wholesale electricity prices. *Energy Policy*, 126, 452-469.
Sensfuß, F., Ragwitz, M., & Genoese, M. (2008). The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany. Energy policy, 36(8), 3086-3094.

Sul, D., Phillips, P. C., & Choi, C. Y. (2005). Prewhitening bias in HAC estimation. *Oxford Bulletin of Economics and Statistics*, *67*(4), 517-546.
Uniejewski, B., Marcjasz, G., & Weron, R. (2018). On the importance of the long-term seasonal component in day-ahead electricity price forecasting: Part II—Probabilistic forecasting. *Energy Economics*, *79*, 171-182.
Uniejewski, B., Nowotarski, J., & Weron, R. (2016). Automated variable selection and shrinkage for day-ahead electricity price forecasting. *Energies*, *9*(8), 621.
Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, *30*(4), 1030-1081.
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*(7), 3939-3944.
Worthington, A. C., & Higgs, H. (2017). The impact of generation mix on Australian wholesale electricity prices. *Energy Sources, Part B: Economics, Planning, and Policy*, 12(3), 223-230.
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.
Zachmann, G. (2013). A stochastic fuel switching model for electricity prices. Energy Economics, 35, 5-13.
Ziel, F., Steinert, R., & Husmann, S. (2015). Efficient modeling and forecasting of electricity spot prices. *Energy Economics*, *47*, 98-111.
Ziel, F., & Steinert, R. (2018). Probabilistic mid-and long-term electricity price forecasting. *Renewable and Sustainable Energy Reviews*, 94, 251-266.

Appendix The ARX-EGARCH model: specification and estimation

Alternatively to the standard OLS model, we consider an **A**uto**R**egressive model with e**X**ogenous regressors with **E**xponential **G**eneralized **A**uto**R**egressive **C**onditional **H**eteroscedasticity model ARX-EGARCH. The model is defined as:

$$\begin{aligned} P_t^h &= \beta_0 + \beta_1 DF_t^h + \beta_2 DA_{t-1}^h + \beta_3 \Delta C_{t-1} + \beta_4 \Delta G_{t-1} + \beta_5 \Delta B_{t-1} + \beta_6 \Delta E_{t-1} + \\ &\quad \beta_7 WF_t^h + \beta_8 SF_t^h + \beta_9 SR_{t-1} + \sum_{j=1} \beta_{j+9} Day_{t,j} + \sum_{k=1} \beta_{k+15} Month_{t,k} + \beta_{27} Hol_t + z_t^h \end{aligned} \tag{A.1}$$

$$z_t^h = \sum_{p=1}^{7} \rho_p z_{t-p}^h + \varepsilon_t^h \tag{A.2}$$

$$\varepsilon_t^h = \sigma_t^h \eta_t^h \tag{A.3}$$

$$\ln\left(\sigma_t^h\right)^2 = \omega + \alpha s_{t-1}^h + \gamma\left(\left|s_{t-1}^h\right| - E\left|s_{t-1}^h\right|\right) + \beta \ln\left(\sigma_{t-1}^h\right)^2 \tag{A.4}$$

$$\begin{aligned} &\eta_t^h \sim f\left(\eta_t^h; \varsigma^h, \lambda^h\right) = \left(2\pi\right)^{-1/2} J e^{-\frac{\left(r^h\right)^2}{2}} \\ &r^h = \varsigma^{-1}\left(\sinh^{-1}\left(\eta_t^h\right) - \lambda^h\right) \\ &J^h = \varsigma^{-1}\left(\left(\eta_t^h\right)^2 + 1\right)^{1/2} \end{aligned} \tag{A.5}$$

The model given by (A.1)-(A.5) is estimated by maximum likelihood. Compared to Eq. (1) in the main text, Eq. (A.1) excludes the autoregressive terms in the mean equation. Instead, the persistence is modelled separately in Eq. (A.7). We use seven lags (one week) of the error terms to account for the serial correlation. The advantage of this approach is that we handle the serial correlation of residuals directly within the model, on the other hand, the interpretation of persistence parameters is unclear. However, other fundamental variables in Eq. (A.1) can be interpreted in the same way as in the OLS regression Eq. (1).

The EGARCH model of Nelson (1991) models the volatility, where s_t^h denotes standardized innovations, and α and γ control for the sign and size effects, respectively, while β for the persistence of the latent volatility process. We choose this specification, for two reasons. First, it allows to address potential asymmetric volatility effects, while it models the log of the volatility, which makes the estimation of the model easier, a property we desire given that we have around 1250 observations for each model. Finally, we allow η_t^h to follow the Johnson's distribution (1949a, 1949b) that is capable of capturing skewness and heavy tail properties of the underlying random variable (if present).

Table A1 ARX-EGARCH model explaining electricity prices for selected hours

	08:00 - 09:00	12:00 - 13:00	19:00 - 20:00	00:00 - 01:00
Panel A: Variables of interest				
Constant	4.722***	23.727***	24.106***	17.672***
Lagged price – one day	0.370***	0.422***	0.353***	0.346***
Lagged price – two day	0.089***	0.156***	0.170***	0.113***
Lagged price – three day	0.029*	0.036**	0.077***	0.120***
Lagged price – four day	0.042*	0.042***	0.066**	0.078***
Lagged price – five day	0.064**	0.027***	0.031	0.093***
Lagged price – six day	0.091***	0.041***	0.098***	0.063***
Previous week's price	0.168***	0.086***	0.120***	0.107***
Forecasted demand/1000	0.189***	0.110***	0.065***	-0.001
Lagged demand/1000	0.006	0.019***	0.044***	0.067***
Lagged coal price return	-0.117***	-0.016	-0.108	-0.251
Lagged natural gas price return	0.164	-0.013	0.076	0.086
Lagged Brent oil price return	0.084	-0.027	-0.025	0.086
Lagged EUA price return	-1.847***	-0.506	-0.752	0.019
Wind production forecast/1000	-0.216***	-0.222***	-0.211***	-0.192***
Solar production forecast/1000	-0.175***	-0.245***	0.142***	
Panel B: Control variables				
Monday	-1.739	-3.910**	-0.903	-0.581
Tuesday	0.262	-2.387	-0.078	2.103
Thursday	-1.524***	-2.656*	-0.021	0.605
Friday	0.599**	-1.737	-1.217	1.490
Saturday	-0.236	-0.005	-1.235	-0.631
Sunday	-0.632	-0.858	-0.587	-1.972*
January	-0.017	-0.675	0.735	-0.843
February	0.360	-0.386**	1.570	-0.841
March	1.198	0.063	1.883	0.259
April	-2.181	0.508	2.202	0.839
May	-4.832**	-3.479*	0.544	2.144
Jun	-1.260	-3.107***	-3.273***	-0.636
August	-1.739	-3.910**	-0.903	-0.581
September	0.262	-2.387	-0.078	2.103
October	-1.524***	-2.656*	-0.021	0.605
November	0.599**	-1.737	-1.217	1.490
December	-0.236	-0.005	-1.235	-0.631
Holiday	-0.632	-0.858	-0.587	-1.972*
Panel C: Variance equation				
Constant	0.295*	1.688***	0.107***	0.325***
Sign effect	0.029	0.011	0.128***	-0.107**
Persistence	0.917***	0.522***	0.967***	0.898***
Size effect	0.304***	0.514***	0.127***	0.263***
Panel D: Distribution parameters				
Skewness	0.151*	-0.129*	0.348***	-0.462***
Shape	1.201***	1.183***	1.451***	1.157***
Panel E: Model fit				
R^2	0.763	0.743	0.772	0.686
1st order autocorrelation of residuals	-0.016	-0.106	0.084	-0.005
7th order autocorrelation of residuals	-0.036	-0.019	-0.016	-0.067

*Note: *, **, *** denotes significance of coefficients at the 10%, 5%, and 1% significance level.*