Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO

Article *in* International Journal of Forecasting · March 2019

DOI: 10.1016/j.ijforecast.2019.02.001

3 authors:

Bartosz Uniejewski
Wroclaw University of Science and Technology

Grzegorz Marcjasz
Wroclaw University of Science and Technology

Rafał Weron
Wroclaw University of Science and Technology

ARTICLE IN PRESS

International Journal of Forecasting xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO

Bartosz Uniejewski [a,b], Grzegorz Marcjasz [a,b], Rafał Weron [a,*]

[a] Department of Operations Research, Faculty of Computer Science and Management, Wrocław University of Science and Technology, Wrocław, Poland

[b] Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wrocław, Poland

ARTICLE INFO

Keywords:
Intraday electricity market
Variable selection
Price forecasting
LASSO
ARX model
Diebold-Mariano test
Trading strategy

ABSTRACT

We use a unique set of prices from the German EPEX market and take a closer look at the fine structure of intraday markets for electricity, with their continuous trading for individual load periods up to 30 min before delivery. We apply the *least absolute shrinkage and selection operator* (LASSO) in order to gain statistically sound insights on variable selection and provide recommendations for very short-term electricity price forecasting.

1. Introduction

Since the deregulation of government-controlled power sectors in the 1990s and 2000s and the introduction of competitive markets in many countries worldwide, electricity has been being traded under market rules like any other commodity (Mayer & Trück, 2018). The workhorse of power trading in Europe has been the uniform price auction conducted a day before delivery, and the vast majority of research studies and applications have concerned day-ahead (DA) electricity prices (Weron, 2014). However, the expansion of renewable generation (mostly wind and solar), the modernization of power grids (including an increase in interconnector capacity) and active demand-side management (smart meters, smart appliances) have made the electricity demand/supply and prices more volatile and less predictable than ever before (Hong & Fan, 2016; Kiesel & Kusterman, 2016). This has amplified the importance of intraday markets, which can be used to balance the deviations resulting from differences between positions in day-ahead contracts and the actual demand (Gianfreda, Parisio, & Pelagatti, 2016; Märkle-Huß, Feuerriegel, & Neumann, 2018; Zaleski & Klimczak, 2015). As a result, the last few years have observed a shifting of volume from the DA to intraday markets across Europe (EPEX, 2018).

This article uses a unique set of prices from the German EPEX market and takes a closer look at the fine structure of intraday markets, with their continuous trading for individual load periods up to 30 min before delivery. We apply the *least absolute shrinkage and selection operator* (LASSO) of Tibshirani (1996) in order to gain statistically sound insights on variable selection and provide recommendations for very short-term *electricity price forecasting* (EPF).[1] Given that the literature on the forecasting of intraday prices in European power markets is very scarce — being limited, to the best of our knowledge, to only two papers dealing with Spanish data (Andrade, Filipe, Reis, & Bessa, 2017; Monteiro, Ramirez-Rosado, Fernandez-Jimenez, & Conde, 2016) — our study is a major step towards understanding the intraday price dynamics and developing predictive models that perform

* Corresponding author.
E-mail addresses: uniejewskibartosz@gmail.com (B. Uniejewski), gelusz@hotmail.co.uk (G. Marcjasz), rafal.weron@pwr.edu.pl (R. Weron).

[1] We use EPF as the abbreviation for both *electricity price forecasting* and *electricity price forecast*. The plural form, i.e., forecasts, is abbreviated EPFs.

https://doi.org/10.1016/j.ijforecast.2019.02.001

Please cite this article as: B. Uniejewski, G. Marcjasz and R. Weron, Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO. International Journal of Forecasting (2019), https://doi.org/10.1016/j.ijforecast.2019.02.001.

well for a market that many participants see as the future of electricity trading. The importance of our study is emphasized further by the fact that electricity price forecasts are now fundamental inputs to energy companies' decision-making mechanisms, alongside weather and demand predictions (Nowotarski & Weron, 2018).

The remainder of the paper is structured as follows. Section 2 reviews the literature on intraday electricity markets and on variable selection for EPF. Section 3 begins by introducing the EPEX dataset and the rolling window scheme, then discusses variance stabilization, and finally describes the model structures considered. Section 4 first compares the predictive performance in terms of two commonly-used error measures and the Diebold and Mariano (1995) test, then takes a closer look at the best LASSO-estimated model in order to identify the most important explanatory variables, and thus provide guidelines to the structuring of better-performing models for intraday electricity markets. Finally, Section 5 wraps up the results and concludes.

2. Literature review

2.1. Intraday markets for electricity

There is a small but growing body of literature on intraday electricity markets. Most publications take a fundamental, market behavior or energy policy perspective. For instance, Pape, Hagemann, and Weber (2016) investigate the explanatory power of a fundamental modeling approach in the German power market, explicitly accounting for must-run operations of combined heat and power plants (CHP) and intraday peculiarities such as a shortened intraday supply stack. González-Aparicio and Zucker (2015) analyze the influence of wind power forecasting errors in the time period between the closure of the day-ahead market and the opening of the first intraday session in the Spanish market. Similarly, Ziel (2017) studies the impact of wind and solar generation forecast errors on price formation in the German intraday market.

Kiesel and Paraschiv (2017) investigate bidding behaviors in the intraday market by looking at both last prices and continuous bidding, in the context of a reduced-form econometric analysis. They find that intraday prices adjust asymmetrically to both forecasting errors in renewables and the volume of trades dependent on a threshold variable that reflects the expected non-renewable generation in the DA market. Aïd, Gruet, and Pham (2016) analyze the trading in intraday electricity markets and develop an optimal bidding strategy. They consider as their modeling framework a continuous-time stochastic process for the net position of sales and purchases of electricity. Märkle-Huß et al. (2018) investigate the introduction of 15-min contracts to the German EPEX market and argue that these products are used to balance the intra-hour volatility of renewable energy sources. Finally, Maciejowska, Nitka, and Weron (2019) take the perspective of a small renewable energy generator that trades via a larger company and only has to decide how much electricity it will sell in the DA market; the rest will be sold in either the intraday market (Germany) or the balancing market (Poland). They forecast the price spread between the DA and intraday/balancing markets using autoregressive and probit models and show that statistical measures of the forecast accuracy, such as the percentage of correct sign classifications, do not necessarily coincide with economic benefits.

However, when it comes to the forecasting of intraday electricity prices in European power markets, the literature is very scarce. To the best of our knowledge, only two studies have addressed this important problem, and these only in the context of the Iberian electricity market (MIBEL), which features a very specific design with six intraday sessions of between 9 and 27 delivery hours. Monteiro et al. (2016) utilize neural networks (multilayer perceptron, one hidden layer) with up to 21 input variables selected on an ad-hoc basis: dummies (hour-of-the-day, day-of-the-week), hourly prices on previous days (lags 1 and 7), price values of the daily session, price values of previous intraday sessions, and weather, demand and wind power generation forecasts. They find that the best models for intraday sessions #1 to #5 use only the hourly prices of the daily session, the hourly prices of previous intraday sessions and the seasonal dummies, while the best model for intraday session #6 uses only the hourly prices of previous intraday sessions #3–#5 and the seasonal dummies. Andrade et al. (2017) reach similar conclusions by utilizing a linear quantile regression (LQR) model, namely that high quality point and probabilistic forecasts of intraday prices can be obtained by just exploring the prices from previous sessions (plus deterministic variables for modeling the daily, weekly and annual seasonalities), despite the fact that they consider a large set of fundamental variables.

2.2. Variable selection for electricity price forecasting

The conclusions from these last two studies suggest clearly that variable selection is a very important issue in EPF, and that it may be even more critical for intraday markets than for DA markets because of the vast amounts of data available. In this context, high-dimensional statistical modeling techniques that deal with large amounts of data may come in handy.

The earliest known examples of statistically-sound variable selection in day-ahead EPF include the studies by Karakatsani and Bunn (2008) and Misiorek (2008), who used *stepwise regression* to eliminate statistically insignificant variables from parsimonious regression-type models, and Amjady and Keynia (2009), who introduced a feature selection algorithm based on *mutual information*. We believe that Barnes and Balda (2013) were the first to apply regularization in day-ahead EPF. In a study concerning the profitability of battery storage, they utilized *ridge regression* to compute EPFs for a model with more than 50 regressors. Ludwig, Feuerriegel, and Neumann (2015) used *random forests* and the LASSO to choose which of the 77 available weather stations were relevant, while Keles, Scelle, Paraschiv, and Fichtner (2016) combined the *k-nearest-neighbor* algorithm with *backward elimination* to select the most appropriate inputs out of more than 50 fundamental parameters or lagged versions of these parameters.

A qualitative change came with the papers by Ziel (2016) and Ziel, Steinert, and Husmann (2015), who used LASSO to sparsify very large (100+) sets of model parameters, utilizing B-splines either in a univariate setting or, more efficiently, within a multivariate framework. In the first thorough comparative study, Uniejewski, Nowotarski, and Weron (2016) evaluated six automated selection and shrinkage procedures (single-step elimination, forward and backward stepwise regression, ridge regression, LASSO, and *elastic nets*) applied to a baseline model with 100+ regressors. They concluded that the use of LASSO and elastic nets can achieve significant accuracy gains relative to commonly-used EPF models. In a study on the optimal model structure for day-ahead EPF, Ziel and Weron (2018) considered autoregressive models with 200+ potential explanatory (but not exogenous) variables, and concluded that both uni- and multivariate LASSO-implied structures outperform autoregressive benchmarks significantly, and that combining their forecasts can achieve further improvements in predictive accuracy. Finally, Uniejewski and Weron (2018) show that using a complex regression model with nearly 400 explanatory variables, a well-chosen variance-stabilizing transformation (asinh or N-PIT), and a procedure that recalibrates the LASSO regularization parameter once or twice a day leads to significant accuracy gains compared to the EPF models that are considered typically.

This study follows the approach set forth in the last three articles and considers LASSO models based on hundreds of potential regressors and calibrated to asinh-transformed prices. However, we do not consider recalibrating the LASSO regularization parameter, as this slows down the forecasting procedure considerably.

3. Methodology

3.1. The dataset

The main German 'spot' market is operated by EPEX SPOT SE and allows the trading of power supply contracts with hourly and quarter-hourly delivery. The participants have the option of bidding on hourly products in the day-ahead (DA) auction that is conducted at noon on the day before delivery (i.e., $d-1$), or trading hourly and quarter hourly contracts in the continuous intraday market that opens at 16:00 on day $d-1$ and closes 30 min before the delivery starts (since March 2017, five minutes for transactions within the delivery zone; see EPEX, 2018).

The leading reference price for the intraday market is the recently-introduced ID3 index for contracts with an hourly delivery, which is also an underlying instrument of exchange-traded derivative products (see https://www.eex.com). The index is based exclusively on hourly and 15-min products traded in the German intraday continuous market (i.e., intraday auction data are excluded), and is computed as the volume-weighted average price of all trades performed over the last three hours before the delivery starts. Moreover, cross-trades (i.e., trades with the same entity selling on one side and buying on the other side) are excluded, while cross-border trades with one leg (buy/sell) in Germany are taken into account (EPEX, 2015).

The exchange publishes the index, but the period covered is too short for a proper evaluation of our models. Thus, we have reconstructed an ID3-like time series from the individual transactions. It differs slightly from the actual index: (i) we have not excluded cross-trades (since the data that we have access to are anonymized), and (ii) we have not considered the trades conducted between 30 and 5 min before the delivery starts, because such trades have been allowed only since March 2017. For each hourly product, only the transactions with timestamps between 180 and 30 min were chosen. For products with no transactions in this period, the window was extended to contain transactions conducted from the start of trading to 30 min before the delivery starts. There was no product without transactions in the expanded window.

In addition to the ID3 index (actually its approximation), we are also using the DA prices as external regressors. Both time series are of an hourly resolution and span the 1216 days from 1.01.2015 to 30.04.2018, see Fig. 1. Like many EPF studies, we consider a rolling window scheme and use a 364-day window in order to estimate our models on a sample which is a multiple of the weekly seasonality and covers a full year; for a discussion of calibration window lengths, see Hubicka, Marcjasz, and Weron (2019). Initially, we fit our models to data from 1.01.2015 h 1 to 30.12.2015 h 24, and compute the price forecasts for the first hour of 31.12.2015. Next, the window is rolled forward by one hour, the models are re-estimated, and the predictions for the second hour of 31.12.2015 are generated. This procedure is repeated until forecasts for the last hour in the 852-day out-of-sample test period (i.e., 30.04.2018 h 24) have been made.

3.2. The forecasting framework

We denote the intraday and day-ahead electricity prices at time (hour) $t = 24d+h$ by P_t and S_t respectively, where $d = 0, 1, \ldots, 1215$ is the day in our sample and $h = 1, 2, \ldots, 24$ the hour of the day. For each hour t in our out-of-sample test period we make a prediction at time $t-4$ of the closing value of the ID3 index for that hour, i.e., P_t. This is illustrated in Fig. 2 using actual transaction data for the period from 12.09.2016 16:00 to 13.09.2016 24:00. Observe the four-hour time lag between the moment when the forecast is made and the time when the delivery starts. For instance, at 12:00 on 13.09.2016 we are forecasting the price for 16:00 (denoted by $\longrightarrow$). The most recent intraday price is for 12:00 (i.e., the hourly contract with delivery between 12:00 and 13:00 on 13.09.2016, denoted by $*$), while the most 'forward-looking' (i.e., beyond the target hour) DA price is for hour 24 on 13.09.2016. One hour later, at 13:00 on 13.09.2016, we are forecasting the price for 17:00 and the most recent intraday price is for 13:00. However, since the day-ahead auction results are known a few minutes after 12:00, the most 'forward looking' DA price is for hour 24 on 14.09.2016.

Following the recommendations set forth by Uniejewski, Weron, and Ziel (2018), we calibrate our models (except for the **Naive** benchmark; see Section 3.3) not to raw prices but to transformed; i.e., $X_t = f(P_t)$, where $f(\cdot)$ is

Please cite this article as: B. Uniejewski, G. Marcjasz and R. Weron, Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO. International Journal of Forecasting (2019), https://doi.org/10.1016/j.ijforecast.2019.02.001.

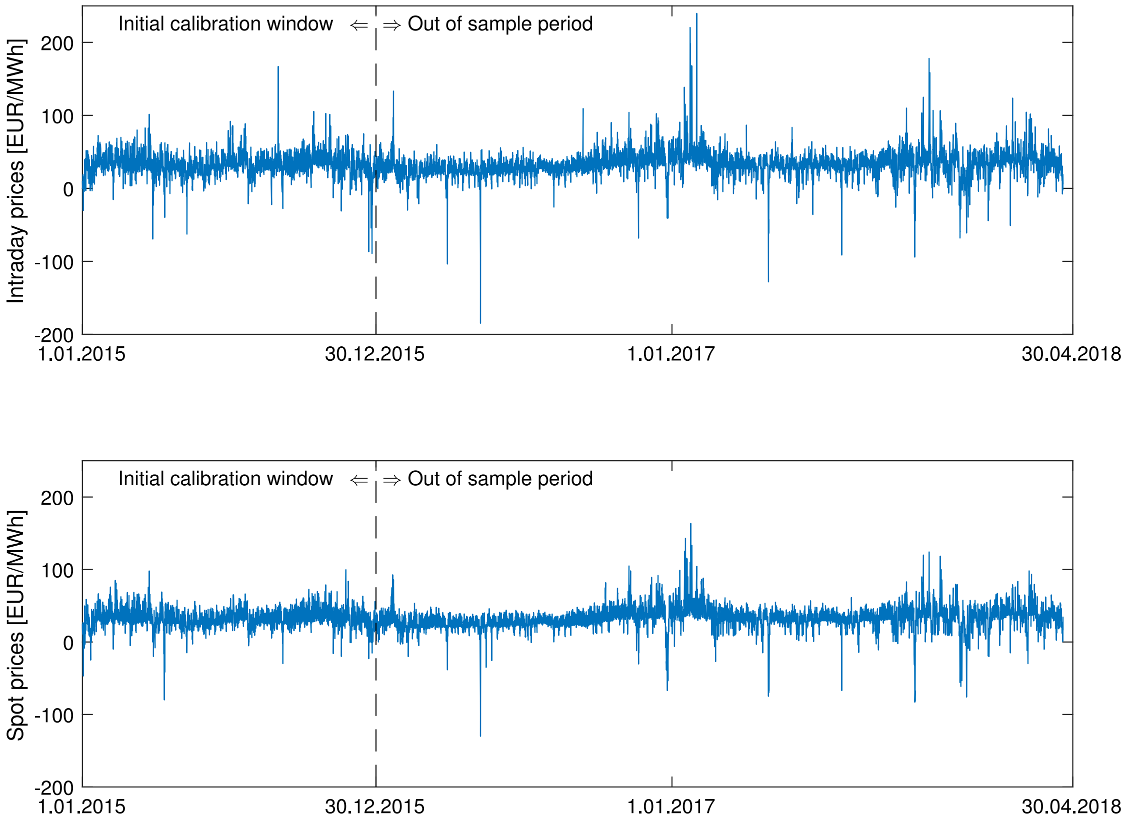

Fig. 1. EPEX hourly intraday (top) and day-ahead (bottom) prices for the period 1.01.2015 to 30.04.2018. The vertical dashed lines mark the beginning of the 852-day out-of-sample test period.

an appropriately chosen *variance stabilizing transformation* (VST). The idea underlying a VST is that of reducing the price variation. As was argued by Janczura, Trück, Weron, and Wolff (2013), a lower variation and/or less spiky behavior of the input data usually allows the forecasting model to yield more accurate predictions.

For electricity markets with only positive prices, the logarithm is the most popular choice for a VST. However, the log-transform is not feasible in our case, since EPEX prices exhibit negative values.[2] Instead, we utilize the *area hyperbolic sine* transformation:

$$X_t = \mathbf{asinh}(p_t) \equiv \log\left(p_t + \sqrt{p_t^2 + 1}\right), \qquad (1)$$

where $p_t = \frac{1}{b}(P_t - a)$ are 'normalized' prices, a is the median of P_t in the calibration window, and b is the sample *median absolute deviation* (MAD) around the median; the latter two parameters are recomputed every day, separately for the ID3 index and the DA prices. Note that the median of all prices after applying the **asinh** transformation is zero, and the variance is close to one. Note also that this transformation can be used for negative data, and its implementation is straightforward. Moreover, it has been found to perform well in a number of EPF studies (Schneider, 2011; Uniejewski & Weron, 2018; Ziel & Weron, 2018). The inverse transformation is the *hyperbolic sine*, i.e., $p_t = \sinh(X_t)$. After computing the forecasts, we apply it to obtain the price predictions:

$$\widehat{P}_t = b\sinh(\widehat{X}_t) + a. \qquad (2)$$

We transform the exogenous series analogously: $Y_t = \mathbf{asinh}(s_t)$, where s_t is the normalized day-ahead price S_t.

3.3. Benchmark models

The first benchmark, denoted by **Naive**, is based on the assumption that the day-ahead and intraday markets are driven by similar data generating processes. It is defined by $\widehat{P}_t = S_t$, where S_t is the DA price for the same day and hour (recall that it is set at noon on day $d-1$). The second benchmark is a parsimonious autoregressive structure inspired by the well-performing *expert*$_{DoW,nl}$ model of Ziel and Weron (2018). In this model, denoted by **ARX**, the VST-transformed price at time t is given by:

$$X_t = \beta_1 X_{t-4} + \beta_2 X_{t-24} + \beta_3 X_{t-48} + \beta_4 X_{t-168} + \beta_5 Y_t + \sum_{i=1}^{7} \beta_{5+i} D_i + \varepsilon_t, \qquad (3)$$

where X_{t-24}, X_{t-48} and X_{t-168} account for the autoregressive effects of the previous days (the same hour yesterday,

[2] Note that negative prices are natural in electricity trading: since plant flexibility is limited (especially for coal-fired power plants) and costly, incurring a negative price for a few hours can actually be economically optimal (Gianfreda, Parisio, & Pelagatti, 2018; Schneider, 2011; Weron, 2006).

Please cite this article as: B. Uniejewski, G. Marcjasz and R. Weron, Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO. International Journal of Forecasting (2019), https://doi.org/10.1016/j.ijforecast.2019.02.001.

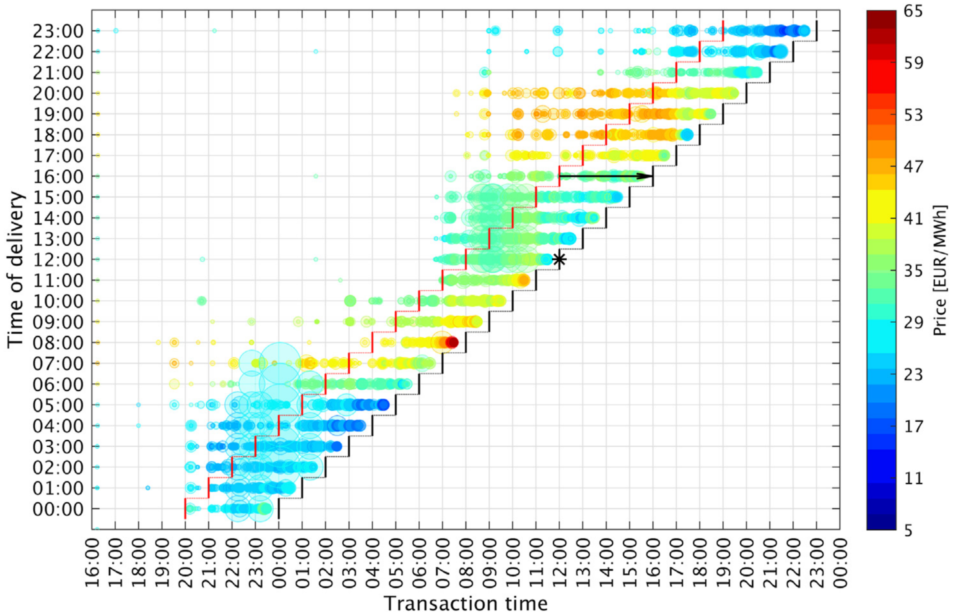

Fig. 2. Illustration of the forecasting framework using actual transaction data for the period from 12.09.2016 16:00 to 13.09.2016 24:00. The black step function indicates the time when the delivery starts (every hour of 13.09.2016), the circles refer to actual trades (the circle size represents the traded volume, from 0.1 to 200 MWh, while the color represents the price, see the colorbar on the right), and the red step function represents the time when the forecasts are made. For instance, when forecasting the price for 16:00 ($\longrightarrow$) at 12:00 on 13.09.2016, the most recent intraday price is for 12:00 ($*$).

two days ago, and one week ago), X_{t-4} is the last observed intraday price, and Y_t is the VST-transformed DA price for the same day and hour. The seven dummy variables $D_1, \ldots, D_7$ account for the weekly seasonality, and are defined as $D_i = 1$ for day of the week i and zero otherwise. Finally, the ε_ts are assumed to be independent and identically distributed normal variables. The **ARX** model is estimated via ordinary least squares (OLS).

3.4. LASSO-estimated models

One advantage of using automatic variable selection is the ability to start out by considering an almost unlimited number of explanatory variables. This study utilizes a baseline model with between 349 (for hour 16) and 372 (for hour 17) potential regressors: seven dummy variables (to account for the weekly seasonality, as in the **ARX** benchmark), 165 last known prices from the intraday market (i.e., nearly the whole week), 169 prices from the DA market (i.e., one week of past prices) and between 8 (for hour 16) and 31 (for hour 17) 'forward-looking' prices from the DA market (i.e., $Y_{t+1}, \ldots, Y_{t+31}$; see also Table 4):

$$X_t = \underbrace{\sum_{k=1}^{7} \beta_k D_k}_{\text{weekday dummies}} + \underbrace{\sum_{i=4}^{168} \beta_{i+4} X_{t-i}}_{\text{past intraday prices}} + \underbrace{\sum_{j=0}^{168} \beta_{173+j} Y_{t-j}}_{\text{current and past DA prices}} + \underbrace{\sum_{j=-31}^{-1} \beta_{373+j} Y_{t-j}}_{\text{'forward-looking' DA prices}} + \varepsilon_t. \quad (4)$$

Note that the price for each hour is predicted with a four-hour time lag (which is why the second sum in the formula above starts with $i = 4$) and using the most recent information. Note also that Eq. (4) does not consider any fundamental regressors. However, the results of Andrade et al. (2017) and Monteiro et al. (2016) suggest that fundamentals (historical and predicted demand, generation and weather) do not have much explanatory power when forecasting intraday electricity prices. Perhaps the DA price for the same day and hour already includes this information.

In order to explain the LASSO scheme, let us rewrite Eq. (4) in a more compact form:

$$X_t = \sum_{i=1}^{n} \beta_i V_t^i, \quad (5)$$

where V_t^i are the regressors and β_i are the corresponding coefficients.

The LASSO of Tibshirani (1996) can be treated as a generalization of a linear regression, where instead of minimizing only the *residual sum of squares* (RSS), we minimize the sum of RSS and a linear penalty function of the β_is:

$$\hat{\boldsymbol{\beta}}^L = \min_{\boldsymbol{\beta}} \{\text{RSS} + \lambda \|\boldsymbol{\beta}\|_1\} = \min_{\boldsymbol{\beta}} \left\{ \text{RSS} + \lambda \sum_{i=1}^{n} |\beta_i| \right\}, \quad (6)$$

where $\lambda \geq 0$ is a *tuning* (or *regularization*) parameter. Note that for $\lambda = 0$ we get the standard OLS estimator; for large λs all β_is become zero; and for intermediate values of λ there is a balance between minimizing the RSS and

Table 1
Mean absolute errors (MAE) and root mean squared errors (RMSE) for the two benchmarks (**Naive**, **ARX**) and the ten **LASSO**($\boldsymbol{\lambda}_i$) models, with $\lambda_i = 10^{-\frac{19-i}{6}}$, $i = 1, \ldots, 10$, over the 852-day out-of-sample test period, see Fig. 1. A heat map is used to indicate better (→ green) and worse (→ red) performing models.

	Naive	ARX	LASSO									
			λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}
MAE	5.0323	4.7234	5.2198	4.9454	4.7323	4.5728	4.4672	4.4135	4.4128	4.4680	4.5704	4.6928
RMSE	8.1098	7.6513	8.0331	7.6670	7.4196	7.2370	7.1122	7.0721	7.1095	7.2356	7.4416	7.6834

shrinking the coefficients towards zero (and each other), and hence performing variable selection.

Selecting a ‘good’ value for λ is critical. Based on the results of Uniejewski and Weron (2018), we have limited our computations to a log-spaced grid of ten values: $\lambda_i = 10^{-\frac{19-i}{6}}$ for $i = 1, \ldots, 10$. Since our dataset is relatively short compared to that used by Uniejewski and Weron (2018), we have not opted to select one optimal λ based on the model's performance in a validation period. Also, we have not chosen the value of λ that maximizes an in-sample information criterion (as per Ziel & Weron, 2018), as this could lead to underperforming models. Since the focus of this study is on variable selection, not on the implementation of the LASSO, we have decided to show the results for all ten λs instead. The models are denoted later in the text by **LASSO**($\boldsymbol{\lambda}_i$), or simply by $\boldsymbol{\lambda}_i$.

4. Empirical results

4.1. Forecast evaluation

We use the *mean absolute error* (MAE) and the *root mean squared error* (RMSE) for the full out-of-sample test period of $D = 852$ days (i.e., 31.12.2015 to 30.04.2018; see Fig. 1) as the main evaluation criteria:

$$\text{MAE} = \frac{1}{24D}\sum_{t=1}^{24D} |\varepsilon_t| \quad \text{and}$$

$$\text{RMSE} = \sqrt{\frac{1}{24D}\sum_{t=1}^{24D} \varepsilon_t^2}, \tag{7}$$

where $\varepsilon_t = P_t - \widehat{P}_t$ is the prediction error at time (hour) t. Recall that the RMSE is the optimal measure for least square problems, but is more sensitive to outliers than the MAE. The MAE and RMSE values obtained can be used to provide a ranking of models, but do not allow us to draw statistically significant conclusions as to the outperformance of the forecasts of one model by those of another. Therefore, we use the (Diebold & Mariano, 1995) test, which is simply an asymptotic z-test of the hypothesis that the mean of the loss differential series:

$$\Delta_{A,B,t} = |\varepsilon_{A,t}|^i - |\varepsilon_{B,t}|^i \tag{8}$$

is zero, where $\varepsilon_{Z,t}$ is the prediction error of model Z at time t and $i = 1, 2$ correspond to the absolute and squared loss, respectively; here, ‘model Z’ is used to refer to one of the benchmarks or the LASSO-estimated models. For each model pair and each dataset we compute the p-values of two one-sided tests: (i) a test with the null hypothesis $H_0 : E(\Delta_{A,B,t}) \leq 0$, i.e., the outperformance of the forecasts of B by those of A, and (ii) the complementary test with the reverse null $H_0^R : E(\Delta_{A,B,t}) \geq 0$, i.e., the outperformance of the forecasts of A by those of B. The loss differential series thus obtained are covariance stationary.

Table 1 reports the MAE and RMSE errors for the benchmarks and the LASSO models, while Fig. 3 depicts the results of the DM tests. We use a heat map to indicate the range of the p-values: the closer they are to zero (→ dark green), the more significant the difference between the forecasts of a set on the X-axis (better) and the forecasts of a set on the Y-axis (worse). For instance, the first row in both panels of Fig. 3 is green except for one black square, indicating that — irrespective of whether we are considering absolute or squared losses — the forecasts of the **Naive** benchmark are outperformed significantly by those of all other models except **LASSO**($\boldsymbol{\lambda}_1$). On the other hand, the column that corresponds to the **LASSO**($\boldsymbol{\lambda}_6$) model is green in the right panel, meaning that this model leads to significantly better forecasts than all others when considering squared losses.

Table 1 clearly shows that the **Naive** benchmark and the **LASSO**($\boldsymbol{\lambda}_1$) model are the worst predictors. Somewhat surprisingly, the **Naive** benchmark even significantly outperforms the worst LASSO model for absolute losses; see the green square in the left panel of Fig. 3 in the **Naive** column. Obviously, the DA price S_t is a good predictor of the intraday price P_t for the same day and hour. However, we can do better than that.

Indeed, even the reasonably parsimonious **ARX** model with only 12 regressors, including seven weekday dummies, significantly outperforms the naive benchmark and one or two of the worst LASSO models, namely $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ (for absolute losses only); see the green squares in the columns labeled **ARX** in Fig. 3. Moreover, it is not outperformed significantly by the naive benchmark or the three worst LASSO models, i.e., $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$ and $\boldsymbol{\lambda}_3$ (for absolute losses) or $\boldsymbol{\lambda}_{10}$ (for squared losses); see the black squares in the rows labeled **ARX**. Now, comparing the LASSO-estimated models among themselves, we can see that the best predictions are obtained for $\boldsymbol{\lambda}_7$ (according to MAE) and $\boldsymbol{\lambda}_6$ (according to RMSE). However, while $\boldsymbol{\lambda}_6$ is a clear winner for squared losses, there are no significant

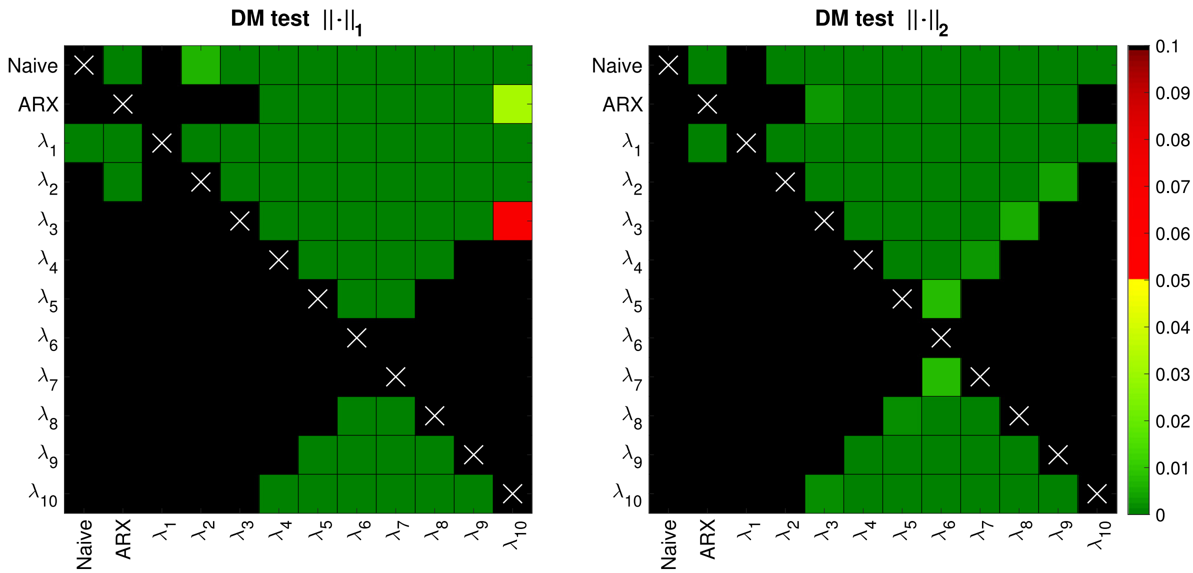

Fig. 3. Results of the Diebold–Mariano (DM) test for the absolute (left) and squared (right) prediction errors, for the same models as in Table 1. A heat map is used to indicate the range of p-values: the closer they are to zero ($\rightarrow$ dark green), the more significant the difference is between the forecasts of a model on the X-axis (better) and those of a model on the Y-axis (worse). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

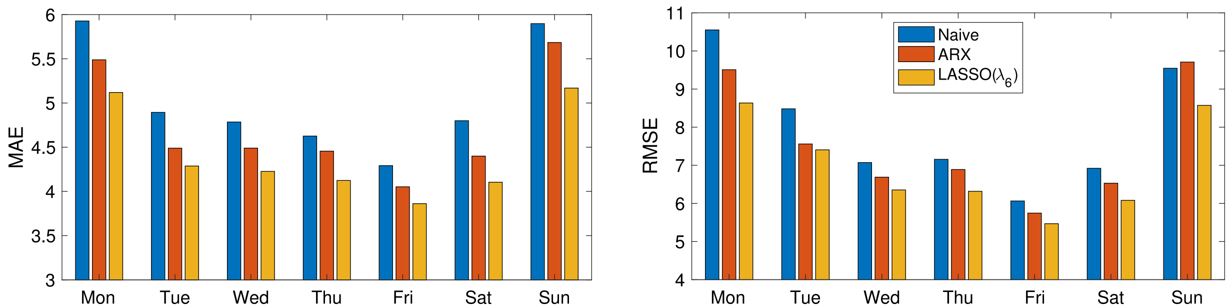

Fig. 4. MAE (left panel) and RMSE (right panel) errors for the two benchmarks (**Naive**, **ARX**) and the **LASSO**($\boldsymbol{\lambda}_6$) model, separately for each day of the week in the 852-day out-of-sample test period, see Fig. 1.

differences in predictive accuracy between $\boldsymbol{\lambda}_6$ and $\boldsymbol{\lambda}_7$ for absolute losses, see Fig. 3. Overall, the picture is consistent with other recent studies that have used the LASSO for day-ahead EPF (Uniejewski et al., 2016; Uniejewski & Weron, 2018; Ziel, 2016; Ziel & Weron, 2018).

4.2. Economic benefits from a simple trading strategy

We now give the error measures above a financial interpretation by considering a simple trading strategy that a participant in the German intraday market can execute. Assume that we want to trade a unit of electricity, say 1 MWh, for each hourly delivery period throughout the 852-day out-of-sample test period; see Fig. 1. We take a long position if the intraday price three hours before the delivery starts (to be precise: the first transaction price after the three hours to delivery time stamp) is lower than our **LASSO**($\boldsymbol{\lambda}_6$) forecast $\widehat{P}_t$ of the ID3 index, and a short position otherwise. We close the position at the first price that exceeds $\widehat{P}_t$; i.e., the first price that is higher than $\widehat{P}_t$ in the case of a long position and the first price that is lower than $\widehat{P}_t$ in the case of a short position. If our ID3 forecast is not breached, then we close the position at the last traded price (to be precise: the last transaction price before the 30 min to delivery time stamp), possibly with a loss. Assuming that there are no transaction costs, this simple strategy leads to a profit of 0.29 EUR/MWh on average across the whole out-of-sample test period. This result clearly shows the usefulness of our approach relative to a profit of 0.03 EUR/MWh from using the **ARX** model and a loss of 0.18 EUR/MWh from using the **Naive** benchmark.

4.3. Performances across days of the week and price regimes

Let us now have a closer look at model performances across days of the week. Fig. 4 plots the MAE (left panel)

Table 2
MAE (top panel) and RMSE (bottom panel) values for the benchmarks (**Naive**, **ARX**) and the ten **LASSO**($\boldsymbol{\lambda}_i$) models with $\lambda_i = 10^{-\frac{19-i}{6}}$, $i = 1, \ldots, 10$, in the 852-day out-of-sample test period, see Fig. 1. As in Table 1, a heat map is used to indicate better ($\rightarrow$ green) and worse ($\rightarrow$ red) performing models.

MAE	Naive	ARX	LASSO									
			λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}
positive spike	29.7105	28.3006	25.8022	26.2783	26.5414	26.7997	27.2811	27.8916	28.2491	28.9377	29.7935	31.1033
normal range	4.6745	4.3522	4.9090	4.6296	4.4127	4.2455	4.1309	4.0652	4.0535	4.0947	4.1805	4.2814
negative spike	30.1848	33.6440	28.2448	27.7597	27.5162	28.0629	28.6342	29.6337	30.9765	32.4805	34.2896	36.3411

RMSE	Naive	ARX	LASSO									
			λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}
positive spike	40.0163	36.9870	34.1606	34.4266	34.7280	35.1616	35.6512	36.1430	36.6081	37.3888	38.5202	40.1076
normal range	6.7899	6.2017	6.9715	6.5417	6.2307	5.9752	5.7921	5.6875	5.6587	5.7086	5.8241	5.9659
negative spike	34.6055	39.6889	34.3290	33.8993	33.9294	34.2932	34.4089	35.1922	36.5425	38.2854	40.4355	42.4274

and RMSE (right panel) values for the two benchmarks (**Naive**, **ARX**) and the best LASSO model, namely $\boldsymbol{\lambda}_6$, separately for each day of the week in the 852-day out-of-sample test period; see Fig. 1. Clearly, the **LASSO**($\boldsymbol{\lambda}_6$) model is better than the benchmarks across all days of the week and for both error measures. As expected, **ARX** outperforms the much simpler **Naive** benchmark, except under the RMSE on Sundays (which is rather surprising).

Following Uniejewski et al. (2018), we obtain a better understanding of model performances across price regimes by considering an evaluation conducted separately for three subsamples defined using the 3σ-rule: (i) a *positive spike* regime: $\mu + 3\sigma < P_t$, (ii) the *normal range*: $\mu - 3\sigma < P_t < \mu + 3\sigma$, and (iii) a *negative spike* regime: $P_t < \mu - 3\sigma$, where μ is the sample mean and σ is the sample standard deviation of P_t in the 852-day out-of-sample test period, see Fig. 1. Overall, there are only 169 (0.83%) positive and 121 (0.59%) negative spikes, and as many as 20,158 (98.58%) 'normal' prices in the test period.

Table 2 reports the MAE and RMSE values for the benchmarks and the LASSO models across the three price regimes. According to both measures, the normal range requires larger λs than the spike regimes, which corresponds to a smaller number of regressors. The best predictions are obtained for $\boldsymbol{\lambda}_7$ (according to both error measures), with $\boldsymbol{\lambda}_6$ following closely behind. On the other hand, positive spikes are captured best by the model with the smallest smoothing parameter, namely $\boldsymbol{\lambda}_1$, while negative spikes are captured best by either $\boldsymbol{\lambda}_2$ or $\boldsymbol{\lambda}_3$. Apparently, extreme prices require much more complex models, utilizing dependencies among many regressors. Finally, as expected, the MAE and RMSE values for the normal range in Table 2 are smaller than those for the full sample in Table 1.

4.4. Variable selection

Let us now comment on variable selection by looking at the results of the model that performed best overall, namely **LASSO**($\boldsymbol{\lambda}_6$). Tables 3 and 4 report the mean occurrences (in %) of model parameters across the 852-day out-of-sample test period. A heat map is used to indicate more ($\rightarrow$ green) and less ($\rightarrow$ red) commonly selected variables. Several interesting conclusions can be drawn:

- The most important variables are the most recent intraday price (i.e., X_{t-4}; see the bottom row in Table 3 with '100' in all columns) and the DA price (i.e., Y_t; see the row labeled '0' in Table 4 with '100' in nearly all columns) that correspond to the predicted hour. Interestingly, the DA prices for the nearby hours (lags -2, -1 and 1) also tend to be selected by the LASSO, which may be an indication that Y_t is not a perfect estimate of X_t and that the prices for the neighboring hours include valuable information.
- Surprisingly, the impact of the previous day's intraday price for the same hour is hardly visible. Hence, there is no reason to have X_{t-24} as an explanatory variable in parsimonious expert models for the intraday market. This is in stark contrast to day-ahead EPF models, where the previous day's price for the same hour is typically one of the most important regressors (Amjady & Keynia, 2009; Karakatsani & Bunn, 2008; Keles et al., 2016; Uniejewski et al., 2016; Ziel & Weron, 2018).
- As was observed by Maciejowska and Nowotarski (2016) and Ziel (2016), the prices for not only hour 24, but also the nearby evening hours, are important predictors. This can be seen by the yellow-green diagonals in Table 3. Note that the first full diagonal (from the bottom, i.e., lags 4 to 27) corresponds to hour 21 of day $d-1$, the second (lags 5 to 28) to hour 20 of day $d-1$, etc. A similar effect can be observed for DA prices; note the yellow-green diagonals starting at the rows labeled '1' (corresponding to hour 24) and '2' (corresponding to hour 23) in Table 4.
- Somewhat surprisingly, the 'forward-looking' DA prices are rarely selected. A notable exception is the price for hour 1 on day $d+1$; see the mostly green

Table 3
Mean occurrence (in %) of model parameters across the 852-day out-of-sample test period. The columns represent the hours ($h = 1, \ldots, 24$) for which the price predictions are made and the rows represent the parameters of the **LASSO**(λ_6) model, see Eq. (4), that correspond to the seven daily dummies and to the past intraday prices X_{t-i}, $i = 168, 167, \ldots, 4$. A heat map is used to indicate more ($\rightarrow$ green) and less ($\rightarrow$ red) commonly-selected variables.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Sun	0	12	12	13	14	18	39	57	63	41	7	5	0	0	17	17	58	31	0	4	1	8	8	3
Mon	1	0	10	19	12	0	0	2	22	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0
Tue	0	0	0	1	2	4	25	32	19	6	27	4	1	5	27	18	13	12	1	0	4	0	0	0
Wed	10	8	19	1	0	0	2	0	0	0	0	0	0	0	0	0	1	3	9	15	0	0	0	0
Thu	11	2	2	2	1	0	0	0	0	0	3	9	11	0	0	0	8	2	3	0	0	0	18	0
Fri	0	0	0	0	0	4	2	0	0	0	0	0	4	0	0	0	0	0	1	0	7	0	0	0
Sat	0	1	0	0	0	0	14	18	0	15	28	16	3	7	4	0	2	34	29	4	10	0	0	0
168	7	0	19	12	8	1	21	47	29	43	48	14	22	38	44	9	48	8	35	20	3	13	9	34
167	1	41	12	15	1	16	4	16	1	14	33	2	19	49	16	45	0	17	2	7	2	6	5	0
166	45	12	2	0	1	3	17	0	0	65	2	4	55	0	34	12	5	0	15	0	66	20	0	0
165	19	2	14	3	12	0	0	0	69	1	0	49	6	8	31	13	12	4	1	60	36	0	28	21
164	0	0	1	1	6	8	4	42	5	6	47	39	5	17	26	35	0	34	61	54	0	0	2	15
163	0	0	32	0	10	2	2	4	0	19	32	9	14	37	37	0	29	50	61	0	19	0	1	0
162	7	22	0	0	7	0	10	9	9	6	2	1	0	15	10	28	34	84	7	30	3	2	10	0
161	25	1	0	2	0	9	1	20	9	13	0	2	9	10	23	37	85	0	34	4	24	25	5	35
160	0	6	1	13	0	16	6	16	23	11	0	22	9	10	13	32	0	0	2	13	21	0	14	63
159	0	6	12	0	0	7	6	35	4	1	29	36	30	1	12	3	0	10	14	13	0	0	42	6
158	9	3	0	1	0	32	19	7	8	2	9	35	3	8	1	1	0	0	21	0	2	47	3	9
157	8	13	1	14	31	5	0	10	2	0	40	9	23	0	19	2	2	0	0	0	4	0	28	0
156	2	0	30	9	61	2	43	10	32	49	9	12	3	6	1	12	12	11	0	24	2	35	28	4
155	7	40	0	38	3	1	1	39	41	52	10	6	7	12	6	10	3	1	6	11	66	19	14	10
154	38	0	16	25	0	0	27	39	25	2	0	17	3	4	11	45	5	21	14	18	5	11	12	1
153	9	2	13	0	0	4	41	8	5	2	0	0	2	0	67	16	18	13	30	2	0	1	12	46
152	0	0	3	0	1	34	25	4	3	1	0	1	1	42	38	29	6	7	0	2	9	5	22	4
151	7	17	3	0	14	51	0	12	10	0	2	1	33	41	9	5	0	8	2	14	3	18	1	24
150	3	9	3	11	32	1	18	0	0	0	12	15	44	0	4	0	26	28	23	0	27	1	40	1
149	38	0	8	34	0	0	2	0	0	12	13	29	0	12	8	3	1	18	0	8	22	54	14	0
148	2	4	32	4	7	0	1	8	0	1	24	3	11	11	12	26	7	0	4	20	48	13	0	18
147	5	6	3	3	1	0	4	0	12	37	0	8	16	25	27	3	9	3	8	5	8	0	22	65
146	0	4	17	4	0	11	0	2	11	18	15	4	6	13	11	21	2	25	1	1	0	13	38	1
145	17	17	8	27	0	0	33	58	36	13	8	1	0	14	10	3	37	26	6	3	25	15	0	2
144	18	5	7	14	9	20	29	29	2	34	7	15	36	2	0	35	29	14	38	41	14	1	6	25
143	1	7	11	0	10	1	30	3	45	0	50	39	2	9	38	9	16	1	5	3	5	24	10	15
142	4	0	5	38	0	4	12	18	0	29	10	0	7	35	5	14	0	0	0	15	56	0	51	9
141	0	19	26	0	47	63	26	27	19	34	0	8	26	0	34	0	0	7	38	36	15	46	21	0
140	10	11	0	57	69	0	7	19	0	15	24	21	0	37	9	10	0	10	22	20	23	0	7	22
139	2	14	54	82	0	3	13	0	19	6	26	1	40	22	13	22	1	40	9	0	4	13	28	3
138	12	72	71	0	0	0	9	4	32	34	0	30	19	16	40	1	42	5	0	0	14	1	2	13
137	63	52	0	0	0	32	7	0	22	0	13	7	8	44	4	12	16	7	0	0	1	3	11	13
136	32	0	0	0	31	7	4	0	0	11	1	3	26	15	13	12	12	0	7	0	2	4	23	28
135	4	13	3	34	13	9	2	2	11	6	0	23	20	3	3	6	5	17	6	27	17	14	15	0
134	25	13	0	14	17	21	41	13	0	16	12	42	3	17	0	8	26	3	36	0	13	8	5	5
133	10	3	11	0	10	36	1	11	31	10	25	6	3	0	0	10	0	30	3	2	8	9	11	1
132	0	4	5	6	27	0	0	15	10	27	40	16	0	2	0	11	12	0	9	18	10	5	15	16
131	12	1	10	0	0	0	26	5	24	25	20	0	0	8	20	0	0	20	21	8	2	3	0	2
130	0	37	0	0	2	3	6	28	29	21	0	1	8	14	6	1	28	18	22	1	0	3	11	26
129	27	0	1	52	24	0	5	14	21	13	0	9	20	5	12	25	14	5	7	10	0	0	3	2
128	7	0	40	2	1	4	23	32	11	4	8	8	14	8	28	30	12	1	15	7	3	3	5	0
127	3	42	2	36	0	29	42	16	9	4	9	27	2	1	38	22	22	25	6	13	26	0	2	12
126	19	0	25	0	19	27	1	0	0	6	21	8	0	25	27	26	27	9	10	6	4	0	19	2
125	0	38	7	34	0	0	0	0	9	25	3	1	28	13	0	13	4	14	19	3	2	29	4	0
124	42	1	49	6	0	52	0	12	11	16	1	48	12	12	17	0	2	13	0	1	24	23	0	9
123	0	36	3	0	36	0	17	0	25	7	16	20	3	0	1	0	15	0	0	46	15	2	0	43
122	0	0	3	1	0	0	0	12	13	12	58	0	0	0	11	11	1	4	50	1	1	9	57	7
121	1	3	0	0	0	3	18	9	8	51	3	1	8	2	0	2	20	15	7	2	42	11	20	9
120	5	2	0	0	0	12	17	40	38	18	9	4	0	3	9	16	4	30	8	77	8	13	0	13
119	0	1	0	1	0	7	37	15	12	19	14	6	22	12	24	13	32	9	4	7	0	2	13	0
118	0	1	0	19	0	35	56	1	29	3	3	15	25	9	1	16	10	1	8	1	4	14	8	2
117	0	0	6	6	19	71	22	15	16	7	9	20	29	0	3	12	0	20	3	10	23	0	0	10
116	3	4	33	43	26	4	11	34	0	15	28	21	7	18	15	8	27	1	0	13	4	0	1	26
115	8	43	45	11	23	0	23	3	2	12	14	10	15	9	21	17	14	1	57	23	0	2	18	13
114	21	56	0	1	0	26	12	10	2	14	25	20	5	12	26	30	5	30	17	0	0	11	15	1
113	22	10	0	11	6	2	24	0	5	4	17	28	11	8	8	11	11	0	0	2	3	1	9	0
112	12	0	13	0	3	1	4	0	9	8	5	5	4	14	9	2	6	0	17	0	1	3	1	5
111	1	0	0	2	8	0	0	20	1	7	10	0	13	10	3	4	0	16	0	0	5	19	6	62
110	5	4	33	3	3	8	0	5	2	25	34	5	1	25	4	0	5	10	0	3	25	5	24	28
109	11	9	2	2	0	17	8	0	33	49	5	0	25	0	0	2	50	2	19	29	54	0	48	17
108	7	19	18	1	10	8	7	44	57	30	0	15	4	0	9	53	26	0	4	47	20	13	2	3
107	4	19	1	29	2	8	46	39	20	0	13	24	3	0	33	0	10	0	51	13	22	5	18	0
106	7	37	3	7	0	2	35	27	0	7	4	7	0	5	0	12	18	5	27	35	5	4	10	27
105	42	9	0	1	5	53	16	0	3	1	0	19	9	12	0	39	2	1	21	1	3	18	12	11
104	11	1	0	34	18	14	7	0	20	0	18	6	8	19	17	27	0	0	1	0	9	0	0	9
103	12	0	19	14	7	1	1	0	0	4	31	7	23	25	40	0	2	5	1	1	0	18	10	0
102	3	1	21	0	0	0	1	0	0	7	12	32	21	43	0	41	43	1	0	0	12	3	20	0
101	1	9	0	11	8	0	2	2	1	17	37	39	62	0	2	25	0	1	4	17	3	7	20	1
100	5	0	2	2	6	5	16	31	5	43	43	13	0	1	19	0	0	1	18	3	39	3	4	4
99	13	4	0	6	5	3	41	0	50	34	19	0	13	3	20	2	18	7	2	0	0	0	7	16
98	25	0	9	1	2	15	5	18	14	1	23	20	17	16	0	3	1	4	1	37	6	0	9	0
97	2	14	0	0	13	32	0	4	1	4	34	13	4	3	2	0	0	4	35	16	9	14	19	54
96	10	1	4	8	15	25	26	20	21	38	4	0	17	5	12	0	0	51	22	69	68	41	20	0
95	18	3	6	3	13	7	22	4	32	4	0	16	3	11	0	0	40	9	46	17	8	9	2	0
94	12	13	8	5	26	28	10	21	14	1	8	0	16	8	3	15	17	24	7	0	0	0	2	24
93	0	0	4	5	32	12	25	20	2	4	3	23	4	15	21	19	25	19	0	3	33	1	43	1
92	0	2	16	19	7	5	9	0	14	5	15	3	21	4	10	18	1	3	6	34	11	6	1	0
91	0	12	15	11	3	31	6	1	27	21	1	22	0	2	0	33	24	10	2	17	12	1	22	7
90	11	1	12	3	17	0	0	29	22	19	2	1	3	1	6	6	31	7	15	9	6	25	24	4

(*continued on next page*)

Table 3 (*continued*).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
89	2	14	12	18	1	12	7	3	1	4	3	4	5	24	7	37	19	1	12	2	30	42	42	0
88	20	7	5	16	1	2	15	2	4	1	1	2	10	1	4	8	2	17	0	28	46	33	8	41
87	17	0	7	12	1	15	3	33	21	0	35	4	29	1	29	23	7	0	17	29	17	12	38	5
86	0	4	11	0	12	4	17	21	1	31	0	0	17	23	3	19	3	7	15	15	4	5	8	0
85	4	25	18	6	0	0	21	0	24	4	12	22	18	28	0	25	0	6	3	1	38	7	1	1
84	6	13	3	2	0	2	8	33	1	6	30	2	2	5	3	0	8	0	0	23	6	21	7	53
83	13	6	6	0	7	4	9	0	25	31	9	2	39	4	7	34	8	7	15	22	44	30	34	10
82	3	8	0	34	3	1	0	2	39	20	0	47	0	4	46	5	2	23	25	25	41	7	18	9
81	13	0	21	3	2	5	12	29	22	1	5	0	7	46	4	11	16	18	25	24	0	19	2	12
80	1	24	33	1	34	10	6	17	2	19	3	3	41	4	4	15	13	0	10	0	2	8	22	0
79	11	35	0	14	24	23	22	0	2	4	0	37	29	29	23	22	8	16	18	3	6	12	0	0
78	0	9	16	26	36	0	0	0	6	20	14	31	28	19	2	6	9	45	0	10	21	0	1	22
77	35	34	27	1	1	0	23	37	13	10	23	35	12	12	43	21	7	0	18	20	26	8	12	35
76	35	11	8	25	0	39	25	14	4	6	42	0	16	46	30	1	7	3	3	19	28	30	33	86
75	0	2	60	8	35	27	25	10	21	22	6	0	18	23	2	8	1	3	0	8	22	37	91	34
74	0	78	33	54	31	36	9	17	2	31	19	2	9	20	1	10	0	0	4	15	0	6	16	9
73	69	11	4	24	24	11	1	23	35	11	1	29	16	0	2	9	0	0	12	1	6	7	30	20
72	22	7	38	27	34	3	26	27	7	6	13	0	3	2	3	8	2	25	12	0	21	23	35	25
71	6	8	20	14	27	40	3	8	2	2	8	0	15	2	11	13	22	26	0	6	17	33	12	1
70	4	4	35	41	19	1	8	39	1	4	11	25	1	9	32	38	20	19	12	1	11	21	5	0
69	2	15	13	5	0	56	45	12	26	10	32	17	0	26	38	24	11	4	6	10	17	2	34	0
68	4	16	3	9	24	7	0	10	2	29	13	0	37	20	38	2	21	3	0	12	2	28	5	2
67	22	7	3	35	23	0	13	7	21	13	6	44	2	29	3	23	0	19	12	44	24	3	9	0
66	22	3	23	0	23	11	7	7	13	9	7	1	41	10	21	0	26	32	45	0	25	21	27	26
65	10	20	0	29	0	0	9	0	16	1	3	19	15	7	10	3	11	33	0	1	23	3	19	30
64	4	0	3	2	2	5	1	6	2	14	12	41	1	25	4	13	1	6	0	0	3	1	43	2
63	5	0	0	32	6	1	3	28	39	2	41	3	20	15	24	20	0	0	0	6	0	33	37	0
62	6	0	0	1	0	2	29	36	8	39	27	16	27	17	0	5	5	0	26	0	37	16	2	7
61	4	0	0	0	12	38	14	13	17	16	12	39	25	0	2	0	0	29	8	53	12	13	15	6
60	4	2	1	6	40	0	27	7	13	5	27	61	5	7	2	0	7	22	37	31	10	36	32	1
59	3	8	4	28	4	6	5	6	0	24	70	5	8	10	2	2	42	48	46	8	39	6	31	3
58	5	0	28	0	9	10	22	0	44	71	4	22	4	0	10	41	34	23	1	33	1	40	41	13
57	2	20	0	24	1	8	13	10	40	5	29	0	1	20	21	19	13	14	5	0	17	42	5	7
56	0	0	10	0	2	0	44	7	13	29	0	0	23	2	37	2	9	0	12	0	41	1	0	3
55	0	13	19	4	0	20	3	1	6	4	0	1	22	21	9	23	0	22	8	45	11	0	19	9
54	0	1	31	15	36	2	6	3	2	1	2	6	10	12	14	1	23	17	38	28	7	0	6	3
53	23	4	47	36	0	3	0	14	0	0	3	6	16	15	0	22	28	55	50	6	30	19	0	13
52	39	38	32	2	3	66	0	8	1	9	20	0	5	0	30	27	35	19	38	44	28	46	28	2
51	5	21	0	5	47	0	14	2	38	33	15	9	6	10	7	14	9	0	14	13	19	49	0	0
50	38	0	22	1	4	10	0	35	25	17	13	6	10	10	0	0	2	3	7	0	42	0	12	7
49	0	28	0	7	28	0	18	0	23	12	5	22	8	3	0	11	6	1	13	11	11	16	23	15
48	42	0	14	0	12	9	0	23	8	4	11	16	10	2	2	13	29	71	11	28	48	51	23	5
47	0	8	1	30	30	15	21	46	0	23	18	7	0	7	23	28	84	24	5	11	12	21	0	22
46	2	0	15	28	22	0	36	2	3	12	3	4	0	24	39	79	6	7	10	0	38	4	3	0
45	0	26	26	13	3	24	0	0	9	2	9	6	17	37	31	0	5	13	0	46	1	5	2	22
44	0	7	23	9	15	0	2	13	0	18	5	6	21	0	0	0	14	1	39	10	0	25	4	21
43	0	24	20	19	29	33	4	5	7	10	0	0	11	21	12	16	24	22	7	5	0	2	43	8
42	55	32	12	26	22	0	25	6	33	5	2	2	9	64	16	20	0	12	5	28	7	11	7	7
41	15	34	37	13	0	21	15	34	9	12	21	4	69	18	0	14	5	7	36	1	7	1	0	27
40	38	31	24	0	0	14	12	4	5	17	15	38	16	1	23	27	0	41	2	23	4	12	59	28
39	0	15	0	7	4	11	2	0	7	14	19	38	1	6	36	0	52	7	4	9	12	30	21	7
38	37	4	12	12	0	0	0	1	10	2	13	2	20	39	1	18	2	2	7	3	22	10	25	0
37	0	0	2	26	0	5	8	29	33	22	2	14	23	2	0	12	6	4	22	25	36	30	0	10
36	0	18	19	18	19	1	17	15	24	2	18	2	1	0	9	0	1	12	11	15	14	41	16	1
35	5	21	5	10	0	44	31	33	6	21	6	0	7	5	9	10	7	6	21	25	26	7	11	0
34	27	3	0	14	2	2	0	13	38	14	0	13	1	9	9	27	21	12	38	3	7	1	0	2
33	0	6	9	0	0	3	2	48	12	2	18	0	3	0	7	7	1	47	2	11	1	0	1	0
32	0	17	2	1	18	1	17	11	1	0	8	0	0	5	33	17	36	3	1	7	0	0	0	9
31	35	8	5	20	7	16	22	19	1	14	5	0	0	41	40	24	41	13	3	16	12	0	5	27
30	0	31	28	12	65	1	7	4	3	2	3	1	17	47	7	37	20	0	2	8	0	18	59	27
29	31	34	3	43	28	14	7	1	13	17	4	7	28	3	1	4	0	15	14	5	7	87	4	0
28	15	3	2	41	16	9	1	0	24	12	5	40	15	3	27	0	33	25	36	5	90	55	2	38
27	0	1	17	3	0	0	0	14	0	7	21	13	19	24	8	20	34	43	7	100	30	0	10	28
26	40	11	17	1	0	2	1	4	13	22	10	5	4	6	17	47	4	4	47	22	0	0	26	18
25	6	21	14	2	1	8	6	28	21	7	0	5	3	0	6	1	2	3	0	0	27	0	22	57
24	0	0	36	25	28	32	23	30	20	0	10	18	0	14	2	1	11	38	2	100	52	71	73	45
23	0	33	28	34	32	36	47	9	0	27	22	0	21	37	21	13	35	9	100	18	91	40	55	33
22	36	4	30	39	19	22	2	10	3	17	1	28	42	27	6	5	15	91	4	48	22	39	22	6
21	28	63	15	10	23	0	11	1	6	0	26	56	15	18	13	7	74	0	1	27	28	62	4	22
20	46	59	3	50	0	14	1	13	13	0	39	7	31	11	25	37	0	4	29	22	60	35	13	0
19	35	17	16	0	3	4	0	11	2	7	9	19	8	26	34	3	13	59	15	51	20	1	3	48
18	6	0	0	8	0	0	1	5	9	11	25	22	25	11	5	8	10	2	47	2	0	0	9	3
17	6	13	21	0	0	20	3	7	12	2	20	21	5	15	2	11	34	48	2	0	2	0	30	10
16	1	15	9	1	43	6	0	11	26	6	30	3	21	16	38	41	17	2	2	7	0	13	27	22
15	17	1	3	70	11	9	32	27	8	10	14	49	28	66	18	17	2	3	8	4	0	42	19	21
14	13	0	48	40	2	13	30	1	0	7	72	31	59	19	1	2	5	0	0	1	58	63	0	12
13	4	8	43	7	0	24	4	0	27	56	37	72	8	10	20	0	3	1	7	34	63	0	8	0
12	14	16	7	0	14	1	6	0	39	16	45	5	6	12	0	9	0	14	35	77	16	3	4	26
11	17	11	0	0	4	0	6	57	21	32	1	15	9	17	21	0	10	37	63	26	44	4	43	0
10	6	0	0	19	1	18	27	23	51	5	8	2	15	5	10	37	44	13	76	68	16	20	10	5
9	5	2	4	0	35	38	36	36	4	5	0	3	23	18	28	47	0	87	74	7	21	0	32	8
8	0	13	40	76	25	2	24	5	10	0	3	5	26	26	45	0	73	33	42	7	4	4	4	25
7	17	20	88	42	0	2	25	15	5	4	0	26	57	32	23	71	12	77	17	0	11	7	31	7
6	19	84	61	0	7	10	82	12	56	16	16	38	59	11	18	8	46	65	0	2	7	4	1	27
5	65	9	29	10	3	10	28	36	39	2	33	75	0	0	0	3	51	0	0	1	3	4	33	37
4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Please cite this article as: B. Uniejewski, G. Marcjasz and R. Weron, Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO. International Journal of Forecasting (2019), https://doi.org/10.1016/j.ijforecast.2019.02.001.

Table 4
Mean occurrence (in %) of model parameters across the 852-day out-of-sample test period. The columns represent the hours ($h = 1, \ldots, 24$) for which the price predictions are made and the rows represent the parameters of the **LASSO**(λ_6) model, see Eq. (4), that correspond to the past day-ahead (DA) prices Y_{t-i}, $i = 168, 167, \ldots, 1$, the DA price for the target hour Y_t, and the 'forward-looking' DA prices Y_{t-i}, $i = -1, -2, \ldots, -31$. A heat map is used to indicate more ($\rightarrow$ green) and less ($\rightarrow$ red) commonly-selected variables.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
168	4	5	2	3	4	9	5	1	34	29	22	11	43	21	12	11	38	54	81	57	43	10	18	14
167	0	4	0	4	8	27	0	72	17	17	12	33	3	15	2	18	30	27	11	34	3	7	0	7
166	0	0	1	10	7	5	15	8	9	0	23	19	2	0	7	1	35	0	5	0	0	0	4	0
165	0	0	10	0	20	8	1	7	0	0	27	1	0	7	1	13	0	7	0	1	28	1	5	12
164	1	6	0	30	29	13	3	0	19	30	1	0	14	9	0	0	0	0	1	2	0	13	0	0
163	1	2	43	21	0	1	0	3	19	4	0	5	9	0	3	0	0	27	0	0	3	0	0	0
162	0	38	14	2	0	0	1	6	13	0	0	18	7	0	0	3	18	1	0	0	6	0	12	18
161	64	29	8	0	39	0	3	47	0	0	15	8	9	1	9	14	12	5	0	0	0	6	12	5
160	18	16	1	53	6	0	1	5	0	23	20	0	0	9	6	36	4	0	0	4	11	7	8	24
159	20	7	0	1	0	6	0	7	21	3	0	0	0	11	26	24	0	0	2	34	4	7	1	6
158	7	4	0	0	24	0	17	7	21	16	0	0	0	36	11	0	0	1	23	0	17	0	0	21
157	2	0	0	14	0	9	5	14	18	0	4	0	37	14	0	22	30	40	0	1	5	0	0	0
156	0	0	35	0	8	2	6	14	2	5	0	0	7	0	35	16	8	0	3	3	0	0	0	8
155	0	32	0	0	2	16	24	10	7	10	0	0	9	11	0	4	0	2	26	1	0	5	33	34
154	11	0	6	0	3	24	3	0	15	0	0	1	4	0	3	1	0	31	3	5	30	21	10	6
153	0	0	1	10	15	0	0	7	12	4	1	3	1	0	0	0	31	1	1	36	16	14	3	2
152	0	0	0	0	6	1	0	2	0	0	4	6	0	0	17	25	7	2	20	24	5	3	10	6
151	0	0	13	4	6	2	18	8	0	41	21	0	1	8	6	0	9	0	22	0	2	9	1	0
150	1	17	8	10	32	0	0	0	49	23	0	0	0	15	0	0	1	4	1	0	3	0	1	2
149	37	11	0	7	0	0	0	23	11	1	0	6	32	8	0	18	2	5	5	23	0	2	3	3
148	1	0	13	0	0	0	0	0	0	0	6	33	9	0	19	15	21	5	9	0	48	8	2	27
147	1	8	0	1	0	0	0	0	0	2	36	4	7	15	13	47	0	0	0	34	14	0	0	0
146	0	0	22	8	0	16	0	0	1	11	4	7	0	13	17	0	0	0	2	0	0	2	4	3
145	2	0	7	4	3	2	21	4	3	0	0	0	3	2	3	4	0	0	0	10	14	32	29	3
144	8	5	6	4	9	6	5	0	0	0	1	10	1	6	17	1	23	0	29	77	75	41	29	7
143	1	2	3	22	0	1	0	1	0	0	5	10	1	11	1	27	6	0	2	31	7	0	0	7
142	0	0	1	0	8	4	19	23	6	1	5	4	0	0	0	13	4	1	0	6	5	0	0	0
141	0	0	2	0	15	35	31	3	21	3	0	0	0	1	7	41	0	0	9	8	2	5	0	4
140	0	1	0	7	30	0	17	0	0	5	0	13	0	4	28	6	1	0	3	10	8	0	17	0
139	0	13	36	29	0	37	0	1	8	1	11	4	13	8	2	9	0	11	4	0	0	0	4	6
138	2	7	7	3	36	0	3	19	1	24	1	17	1	0	11	0	9	10	21	2	11	0	31	0
137	9	18	11	21	2	22	9	0	38	8	0	3	0	30	0	6	1	62	14	1	5	21	1	3
136	19	13	6	22	3	19	0	31	1	0	0	0	6	1	0	1	50	10	7	7	7	0	17	42
135	14	0	0	9	23	0	29	0	0	0	13	0	12	7	0	35	0	1	7	35	0	0	9	0
134	0	0	0	21	0	3	11	0	2	13	0	0	15	10	12	0	12	0	0	5	0	3	4	0
133	2	0	23	3	0	5	13	2	0	0	2	13	9	2	0	0	0	0	0	31	0	0	4	0
132	0	5	0	3	0	2	14	0	2	4	8	3	0	0	3	0	0	0	8	10	6	0	0	52
131	0	0	0	0	9	0	0	20	11	1	1	1	2	5	0	0	1	6	7	2	0	11	12	2
130	1	2	0	0	11	0	38	21	0	13	1	0	2	0	0	6	5	0	16	2	8	13	0	36
129	0	0	24	4	2	7	53	2	5	0	2	0	6	0	0	2	10	0	0	13	2	0	0	0
128	0	7	8	12	5	0	0	2	0	0	0	0	29	0	0	25	1	2	19	1	7	0	0	0
127	19	15	0	2	0	6	1	0	2	10	5	43	0	0	2	5	0	23	7	16	1	0	6	0
126	4	0	0	3	0	8	0	7	6	2	41	0	2	8	0	0	0	12	14	10	0	14	5	1
125	13	0	5	2	17	0	0	0	0	45	0	9	23	0	3	0	8	11	10	1	9	63	4	37
124	0	6	17	41	0	0	0	1	16	3	17	35	2	0	0	0	6	8	4	0	27	1	50	23
123	2	19	18	0	0	0	0	0	32	4	30	0	7	0	0	1	3	0	12	8	5	25	35	0
122	30	19	0	0	0	17	0	23	1	12	4	10	12	0	11	7	0	12	1	0	19	13	0	27
121	0	0	0	0	22	8	0	7	7	5	7	3	4	7	0	0	0	9	4	5	1	2	37	38
120	0	0	2	12	13	0	7	21	9	9	3	0	7	6	3	7	8	2	11	18	7	16	45	41
119	0	2	10	8	0	2	27	3	13	1	3	0	17	24	2	11	0	5	5	2	5	3	33	0
118	0	3	4	0	0	21	21	25	12	0	0	0	17	0	18	14	6	11	22	0	8	16	0	8
117	0	9	0	1	15	5	19	0	1	20	0	28	0	10	12	33	0	10	0	25	2	0	0	0
116	11	0	1	6	21	5	0	7	25	23	10	0	7	18	49	3	6	0	6	6	0	0	0	10
115	0	15	19	27	4	0	19	0	25	2	0	1	14	28	0	1	5	0	1	8	9	1	0	1
114	4	0	1	0	0	1	8	0	7	4	1	11	22	8	1	0	7	0	0	0	1	0	0	12
113	5	22	5	0	8	2	0	1	0	1	10	0	0	0	0	1	7	0	0	1	0	0	0	14
112	1	7	0	1	4	10	0	5	4	3	0	0	0	1	0	36	8	0	2	1	3	0	1	4
111	0	0	12	3	2	0	0	8	0	0	2	0	0	0	19	12	0	0	2	3	1	17	6	0
110	1	0	3	18	2	1	0	2	3	11	5	13	14	16	0	3	0	0	0	0	43	7	11	7
109	0	5	4	4	0	4	4	1	12	8	12	14	12	15	0	2	0	0	0	32	0	4	0	4
108	6	5	4	1	3	0	4	12	15	8	18	5	21	2	0	0	0	2	0	12	2	15	0	3
107	0	3	0	0	6	10	19	10	8	47	0	14	1	3	0	0	0	12	11	6	10	1	3	9
106	19	0	0	6	50	23	16	0	37	0	8	22	7	0	0	4	2	52	5	0	13	5	5	1
105	0	0	16	29	25	0	0	12	0	15	6	6	1	0	5	1	38	3	13	12	22	0	2	0
104	0	0	7	17	0	1	3	6	17	0	1	1	1	9	6	57	0	1	3	27	0	0	0	0
103	0	1	5	0	2	10	1	0	0	0	0	7	21	4	29	0	37	0	10	2	0	0	0	2
102	1	2	2	7	16	2	0	0	0	0	5	27	4	17	5	40	1	0	4	0	4	0	6	0

(*continued on next page*)

'half-diagonal' in Table 4 starting at the row labeled '−8' for the column (i.e., target hour) labeled '17'. A plausible explanation is that the DA price for the first hour of the next day includes important information about the price evolution for the late night intraday prices; it is known a few minutes after noon, and hence is available for forecasting intraday prices for deliveries starting at hour 17.

- Finally, which also comes as a surprise, the dummies are usually removed by the LASSO (except for two morning hours and one afternoon hour on Sunday), see the top seven rows in Table 3. Again, this is in contrast to day-ahead EPF models, where the weekday dummies are typically selected (Uniejewski et al., 2016; Ziel & Weron, 2018).

Table 4 (*continued*).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
101	34	11	16	37	0	0	3	0	0	8	19	10	5	9	10	12	0	2	2	0	6	37	3	4
100	17	12	10	0	0	2	0	0	7	19	14	0	3	16	31	8	0	0	12	0	23	30	4	25
99	0	17	1	0	1	1	0	0	16	5	0	0	12	11	39	0	0	0	12	0	1	1	14	0
98	0	2	0	0	2	0	0	1	2	36	0	4	9	29	3	0	0	0	1	0	5	2	4	1
97	1	0	0	1	2	4	2	13	26	0	1	1	0	1	0	1	15	20	48	10	4	1	5	25
96	22	1	9	0	6	0	7	12	3	19	0	23	9	2	0	17	40	75	31	42	25	25	38	19
95	3	13	0	0	0	6	7	8	23	7	35	4	12	0	6	30	46	6	1	18	16	22	0	31
94	7	0	0	0	7	0	1	6	1	20	30	31	6	8	28	2	0	0	17	0	0	1	0	0
93	0	0	0	25	14	0	0	2	9	19	25	6	43	17	2	4	0	4	0	0	15	0	0	0
92	0	0	9	27	1	0	7	0	4	5	3	52	20	9	17	16	0	1	2	33	0	3	0	0
91	0	0	33	10	0	1	0	0	0	9	15	13	17	20	3	0	0	1	42	0	2	0	5	0
90	0	8	18	0	8	0	0	0	0	15	1	16	21	0	0	1	2	16	3	16	0	1	21	0
89	0	49	1	0	0	0	14	2	0	0	0	8	0	11	15	0	4	7	21	0	0	14	2	0
88	21	3	0	1	2	4	56	15	15	2	0	14	8	10	22	8	0	0	2	0	12	0	1	0
87	0	0	0	1	2	7	9	14	0	0	17	15	10	26	2	5	0	0	2	16	4	1	0	9
86	0	5	0	13	19	17	23	0	0	16	4	12	19	4	0	0	8	2	15	2	11	1	7	8
85	1	0	10	0	0	13	0	0	14	0	15	1	2	0	0	6	0	3	2	14	8	6	0	5
84	3	15	0	0	0	3	0	14	31	11	2	19	0	1	7	0	0	2	15	3	2	18	0	0
83	23	0	0	0	10	20	13	31	7	1	5	0	0	6	0	0	0	8	1	0	12	0	0	0
82	2	0	0	15	31	10	5	17	7	11	2	0	2	0	0	0	5	2	4	5	0	0	0	24
81	4	0	1	35	8	52	12	7	4	0	0	0	0	7	2	27	12	2	23	0	8	1	1	5
80	3	2	33	6	32	15	0	0	0	0	0	1	1	0	36	0	0	10	1	0	1	2	10	1
79	12	24	6	39	18	23	0	9	0	0	0	0	0	36	13	5	17	3	6	0	0	10	2	0
78	8	1	46	25	21	0	0	0	0	4	0	0	36	7	2	18	26	14	0	10	27	9	24	0
77	1	7	1	16	2	0	0	5	0	0	0	28	8	8	0	32	6	1	42	38	12	44	0	17
76	18	0	0	7	0	0	15	1	0	0	2	30	7	3	5	24	2	51	20	0	28	16	19	0
75	0	0	0	0	0	25	0	0	0	2	18	26	0	14	20	0	0	5	0	30	5	15	0	1
74	22	2	17	5	34	36	6	6	8	4	33	0	13	3	4	0	0	0	1	3	0	0	1	17
73	0	1	0	29	18	13	33	0	6	0	8	1	0	0	1	0	0	0	1	0	40	3	22	47
72	0	7	24	39	0	32	0	15	0	21	5	0	0	0	15	0	1	0	34	39	4	38	48	27
71	4	2	21	15	0	0	18	7	0	0	0	0	0	22	8	0	0	16	11	21	16	12	5	3
70	0	12	8	0	0	4	29	0	1	1	0	13	19	7	0	0	0	13	33	0	0	0	0	0
69	0	18	0	0	14	13	8	4	17	0	4	26	13	0	4	3	0	27	14	3	3	3	0	2
68	0	10	0	13	39	0	0	1	2	3	26	1	0	0	4	1	12	6	10	4	0	0	12	0
67	4	0	3	9	0	0	3	0	2	22	0	0	0	0	0	39	5	15	1	0	0	0	1	1
66	8	0	12	0	8	3	0	0	3	16	2	0	0	8	32	1	32	0	9	16	0	1	2	0
65	1	0	0	9	1	0	0	0	14	4	0	0	12	37	16	12	4	14	5	21	0	0	7	20
64	1	0	12	9	0	0	0	19	0	2	0	0	36	3	5	13	7	1	48	5	1	0	20	4
63	2	9	2	0	0	20	18	0	16	0	0	21	14	0	17	2	6	31	8	2	0	13	5	9
62	0	0	1	1	0	19	6	20	6	1	4	13	4	7	5	0	24	0	5	0	19	8	0	2
61	0	0	0	0	0	4	1	11	0	6	0	11	4	6	0	3	0	0	0	17	3	2	2	0
60	9	0	0	0	0	5	5	0	11	0	0	1	2	0	0	1	0	0	7	20	0	2	0	13
59	0	1	0	1	0	17	0	8	6	0	0	0	6	4	16	16	0	9	21	10	11	3	16	12
58	6	0	2	0	0	7	34	1	0	1	6	4	1	16	39	0	41	0	3	1	8	8	6	2
57	2	23	0	0	0	14	11	0	2	2	7	0	6	42	5	20	6	6	0	1	11	6	0	3
56	15	0	0	3	0	0	0	8	13	7	13	10	3	0	11	20	42	0	1	4	1	0	0	0
55	32	0	1	1	1	0	4	0	0	0	0	1	3	26	18	46	4	36	7	7	14	1	5	37
54	0	0	4	10	17	0	2	0	2	0	0	0	34	13	31	3	21	6	7	0	0	0	11	0
53	0	0	11	5	0	0	0	0	6	0	0	37	0	15	12	17	25	3	1	0	0	4	4	7
52	8	0	0	0	15	1	3	0	2	0	34	0	0	9	43	46	1	1	20	0	0	19	6	1
51	5	0	0	8	3	6	0	0	0	36	13	0	2	20	44	0	0	0	1	0	1	0	2	0
50	0	4	2	0	12	18	0	0	39	10	0	1	3	22	2	0	0	0	1	9	0	17	17	0
49	17	12	0	13	24	3	0	49	8	0	1	3	0	16	12	0	0	3	31	29	17	10	15	7
48	0	0	0	9	8	24	23	14	0	1	7	0	14	10	0	0	8	45	30	80	41	10	36	15
47	0	6	11	4	36	21	25	2	3	0	10	6	9	4	3	1	1	19	45	12	4	7	20	5
46	0	6	2	47	7	16	0	0	8	16	5	0	7	2	0	0	1	0	0	0	3	2	16	2
45	0	7	53	3	1	0	16	2	3	2	0	3	2	4	0	0	5	0	14	2	1	0	1	5
44	0	48	2	4	10	0	13	0	0	3	4	3	8	0	4	12	1	0	7	0	2	0	2	4
43	21	13	4	8	3	5	9	4	0	8	4	29	4	7	19	1	0	0	5	6	0	7	19	26
42	15	2	9	9	17	0	3	6	11	6	19	18	0	2	12	0	2	0	8	0	27	14	0	0
41	20	0	16	0	2	0	10	19	0	4	16	7	0	26	0	0	2	3	2	15	27	2	3	1
40	0	9	1	2	0	4	1	1	0	1	11	4	27	1	0	5	5	6	29	1	0	6	30	16
39	10	1	0	1	1	8	2	11	1	4	2	24	0	0	6	0	10	2	4	0	6	17	5	10
38	5	0	3	0	4	0	5	4	5	0	10	0	0	11	0	5	0	25	0	0	14	0	0	4
37	2	3	0	0	11	0	18	2	0	5	0	8	8	4	0	4	1	0	0	9	10	0	0	0
36	0	3	7	25	1	3	0	7	6	1	10	8	7	0	27	2	0	0	12	17	2	0	3	40
35	16	4	45	27	0	0	5	9	0	64	26	22	0	42	0	3	0	37	1	4	0	0	28	0

(*continued on next page*)

4.5. Using the LASSO to build parsimonious ARX-type models

Now that we know which variables are selected by the LASSO, we can ask ourselves whether this information can be used to build well-performing, parsimonious ARX-type models. We address this question by computing the frequency of selecting each of the 349 (for hour 16) to 372 (for hour 17) regressors of the **LASSO**(λ_6) model, see Eq. (4), in a 364-day rolling 'selection' window that follows the 364-day calibration window directly. Then, we forecast prices in the remaining 488-day out-of-sample test period by building ARX-type models that include as regressors only those variables that have been selected at least $x\% = 50\%, 60\%, \ldots, 90\%$ of the time. The resulting models are denoted by $\mathbf{ARX}_{x\%}$.

The MAE and RMSE values are reported in Table 5. Note that they differ from those reported for the benchmarks and the **LASSO**(λ_6) model in Table 1 because of the much shorter out-of-sample test period (488 vs. 852 days). Surprisingly, some of the $\mathbf{ARX}_{x\%}$ models perform

Table 4 (*continued*).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
34	6	35	16	16	0	14	11	0	68	13	16	0	44	0	4	0	14	1	9	0	0	0	0	5
33	50	17	12	0	19	9	3	4	5	13	0	41	17	3	0	17	25	4	0	0	0	3	0	19
32	1	26	0	10	13	8	3	0	0	0	0	0	5	0	0	28	21	0	0	29	0	12	0	1
31	20	4	14	1	0	25	0	0	0	0	0	2	3	4	3	17	0	0	0	0	36	1	0	0
30	0	2	0	4	14	21	20	0	15	0	17	0	0	0	1	0	0	8	0	21	3	4	3	0
29	7	0	0	5	2	2	0	0	0	7	0	0	0	0	11	0	1	1	19	15	3	3	4	0
28	2	5	7	18	0	0	0	0	0	1	4	0	0	8	6	1	0	24	26	17	11	35	0	0
27	0	29	28	0	0	0	0	4	11	18	5	0	9	4	22	0	14	2	0	0	34	0	0	6
26	34	4	1	0	0	0	0	6	10	4	4	16	5	20	2	20	2	0	0	2	0	0	17	0
25	8	0	5	0	2	0	1	11	9	1	33	5	4	5	1	2	0	0	12	0	29	24	12	35
24	4	0	0	4	13	6	2	34	14	55	6	7	12	0	0	0	0	10	1	19	29	46	42	0
23	0	0	0	40	14	7	43	6	12	3	0	12	0	0	0	0	7	0	44	1	1	2	0	2
22	0	0	13	9	0	62	19	6	32	4	10	0	0	7	0	7	0	15	0	0	0	9	6	2
21	0	0	5	4	96	0	0	29	10	11	0	0	30	0	10	20	5	0	0	16	8	0	0	0
20	0	1	1	53	0	0	17	12	15	0	4	23	0	14	16	4	1	5	14	26	0	0	0	0
19	0	32	5	13	0	0	0	8	0	2	5	3	15	11	8	22	12	46	38	8	12	0	0	13
18	41	0	0	0	0	40	16	0	0	23	14	8	40	0	38	17	37	44	31	31	0	0	3	17
17	0	14	0	4	21	30	0	0	8	32	32	39	0	23	16	33	80	33	13	27	3	4	1	51
16	16	16	0	5	51	8	0	25	29	36	39	0	0	27	30	59	19	9	42	2	4	1	35	2
15	18	7	4	20	16	0	36	5	13	49	0	1	30	40	78	8	0	26	3	6	5	45	4	36
14	2	11	25	17	0	0	17	11	71	32	0	33	36	76	1	1	14	0	45	6	40	22	27	15
13	1	18	10	2	1	14	10	74	36	8	40	20	85	6	0	7	0	52	9	40	3	7	19	3
12	13	28	17	0	0	2	19	32	15	25	48	90	4	0	2	0	52	5	44	3	21	2	0	8
11	1	7	0	0	0	25	36	0	13	42	93	5	0	0	0	27	16	30	4	23	16	5	15	0
10	24	0	9	0	0	0	3	2	35	56	19	0	0	1	12	35	38	5	28	23	0	17	14	10
9	0	3	0	4	0	0	0	38	79	51	7	2	6	13	9	23	37	30	33	33	0	54	19	34
8	6	1	3	29	1	0	97	86	38	24	21	7	7	6	23	2	22	21	51	16	31	3	14	3
7	7	12	45	22	38	100	62	38	24	48	27	6	0	24	4	32	36	64	26	53	9	54	43	17
6	18	67	55	12	96	0	53	18	57	51	2	0	28	26	37	71	19	85	90	37	81	51	4	42
5	51	31	0	66	0	28	12	87	82	51	29	25	55	90	71	38	61	66	48	76	68	7	33	34
4	15	0	6	0	20	0	96	59	60	34	83	100	100	54	50	69	88	42	18	37	32	73	93	97
3	0	3	0	0	0	0	7	30	13	3	11	51	0	4	0	0	0	0	0	8	3	9	0	35
2	38	14	0	0	0	0	1	0	0	80	10	0	0	0	0	0	0	0	3	17	0	2	0	0
1	42	0	1	27	12	9	0	1	100	78	0	8	57	36	20	43	61	44	54	15	69	4	14	89
0	100	100	65	100	100	100	99	100	100	92	100	100	100	100	100	100	100	100	100	100	100	100	100	100
-1	96	54	70	74	36	43	94	100	50	64	92	66	83	86	60	68	55	27	24	37	53	30	2	80
-2	56	69	27	58	9	93	64	2	35	84	23	44	68	44	34	67	0	6	16	0	0	1	100	29
-3	47	11	36	2	26	51	2	10	45	0	30	72	4	62	42	10	6	0	2	4	17	100	21	35
-4	6	27	10	0	5	1	16	48	0	4	33	35	51	38	3	4	7	6	5	10	100	27	0	6
-5	34	39	1	0	38	0	0	12	9	53	14	51	15	16	22	9	8	3	9	100	25	2	0	0
-6	39	0	0	2	13	5	4	10	12	31	41	11	33	13	1	21	11	10	80	21	8	0	0	16
-7	5	0	0	10	0	0	1	0	44	17	19	38	17	0	28	20	3	76	23	45	3	0	0	7
-8	9	0	17	4	0	1	0	8	16	9	41	9	10	37	29	79	57	5	38	9	1	0	23	28
-9	0	7	1	0	0	0	34	21	15	64	8	35	37	19	65		14	56	0	0	0	47	26	0
-10	9	0	1	0	3	11	23	25	55	31	80	11	45	77			62	0	0	0	63	31	10	54
-11	1	11	0	1	25	42	6	14	25	93	1	42	80				5	0	0	39	61	19	56	7
-12	10	1	9	22	44	9	36	55	57	8	22	16					0	0	5	48	25	64	4	24
-13	0	30	15	69	21	12	2	20	19	18	11						11	29	0	11	40	8	5	5
-14	0	5	80	9	6	16	28	12	33	9							32	2	12	9	6	0	29	7
-15	0	77	12	15	0	8	14	18	13								14	5	18	47	4	0	13	16
-16	79	7	11	14	23	19	0	1									2	3	88	10	0	31	12	0
-17	3	0	19	6	11	0	16										0	19	5	11	18	22	2	10
-18	0	31	6	43	27	1											48	56	23	1	41	4	9	0
-19	13	0	32	0	37												50	23	0	24	2	0	18	7
-20	0	24	12	36													19	3	3	0	3	8	5	12
-21	7	8	20														9	2	1	5	4	4	1	1
-22	92	39															6	0	28	0	0	6	0	5
-23	8																4	47	7	25	50	9	0	68
-24																	41	8	30	20	36	34	62	7
-25																	2	21	0	13	12	40	5	
-26																	23	9	2	0	0	6		
-27																	25	3	2	8	21			
-28																	1	8	16	25				
-29																	2	7	27					
-30																	3	40						
-31																	8							

exceptionally well. In particular, **ARX**$_{70\%}$ is slightly better in terms of RMSE than **LASSO**(λ_6), despite utilizing only 13.4 explanatory variables on average, i.e., about 27 times fewer than the baseline LASSO model in Eq. (4) and nearly 3.5 times fewer than **LASSO**(λ_6), which selects an average of 46 regressors. However, the difference is not statistically significant, see the DM plots in Fig. 5. On the other hand, the **LASSO**(λ_6) model is the best in terms of MAE, being significantly better than all of its competitors except **ARX**$_{70\%}$.

5. Conclusions

We have used a unique set of electricity prices from the German EPEX market to address the problem of the optimal choice of explanatory variables for forecasting intraday prices. Given that the literature on this topic is very scarce, our study is a major step towards understanding the intraday price dynamics and developing well-performing predictive models for a market that many participants see as the future of electricity trading.

Table 5
MAE and RMSE values for the two benchmarks (**Naive**, **ARX**), the **LASSO**(λ_6) model and five **ARX**$_{x\%}$ models built on the latter with cutoffs of $x\% = 50\%, 60\%, \ldots, 90\%$, over the 488-day out-of-sample test period (see Section 4.5 for details).

	Naive	ARX	LASSO	ARX				
			λ_6	50%	60%	70%	80%	90%
MAE	5.7051	5.3095	4.9330	5.0119	4.9711	4.9539	5.0021	5.0936
RMSE	9.2912	8.5738	7.8343	7.8883	7.8552	7.8325	7.9419	8.1578

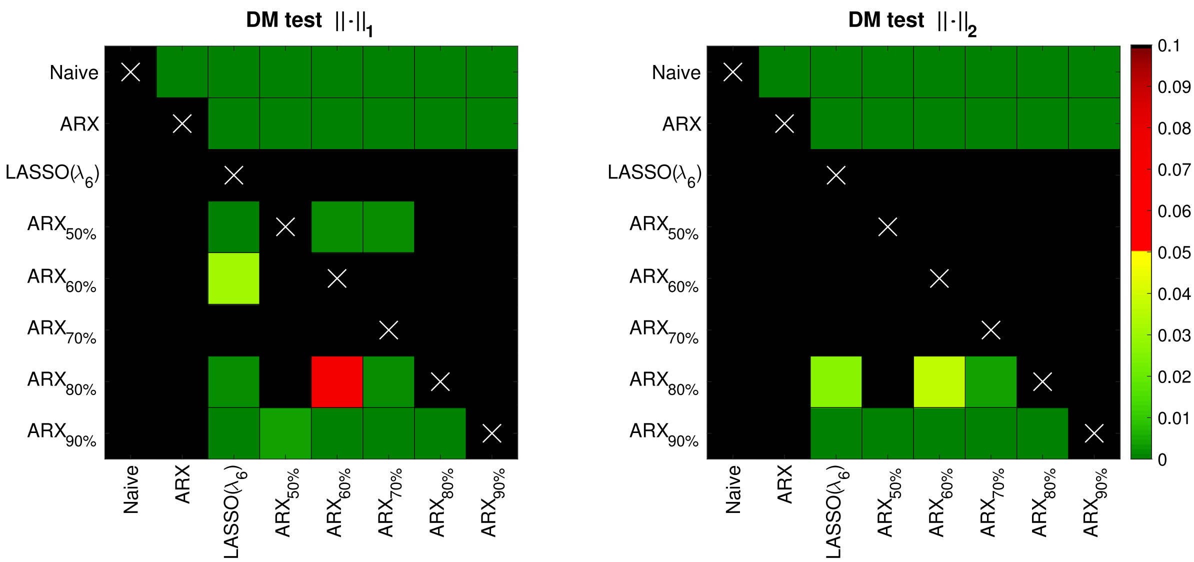

Fig. 5. Results of the Diebold–Mariano (DM) test for the absolute (left) and squared (right) prediction errors, using the same models as in Table 5. A heat map is used to indicate the range of the p-values: the closer they are to zero (→ dark green), the more significant the difference is between the forecasts of a model on the X-axis (better) and the forecasts of a model on the Y-axis (worse). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

To this end, we have considered 12 models, namely a naive benchmark, a parsimonious autoregressive structure inspired by the well-performing $expert_{DoW,nl}$ model of Ziel and Weron (2018), and a LASSO-estimated model with 349 (for hour 16) to 372 (for hour 17) potential regressors and ten different values of the *tuning* parameter. We have found that, for an appropriately chosen value of λ, the LASSO model significantly outperforms its competitors, as measured by the Diebold–Mariano test.

The most important explanatory variables turned out to be the most recent intraday price and the day-ahead (DA) price that corresponds to the same hour. The intraday and — to a lesser extent — DA prices for late evening hours could also be considered as regressors. On the other hand, in contrast to day-ahead EPF models, neither the previous day's price for the same hour nor weekday dummies were found to be important predictors.

Finally, we have shown that the LASSO can be used to build well-performing, parsimonious ARX-type models. In particular, the performance of an OLS-estimated model with regressors that have been selected by the LASSO at least 70% of the time in a 364-day rolling 'selection' window is comparable to that of the best LASSO model. At the same time, it utilizes an average of only 13.4 explanatory variables, i.e., about 27 times fewer than the baseline model and nearly 3.5 times fewer than the best LASSO model, **LASSO**($\boldsymbol{\lambda}_6$).

Acknowledgments

This work was partially supported by the German Research Foundation (DFG, Germany) and the National Science Center (NCN, Poland) through BEETHOVEN grant no. 2016/23/G/HS4/01005. The authors would like to thank the participants of the IMMORTAL kick-off meeting (9-11.05.2018, Wrocław, Poland), the 7th International Ruhr Energy Conference INREC2018 (24-25.09.2018, Essen, Germany) and the Energy Finance Christmas Workshop EFC18 (12-13.12.2018, Bolzano, Italy) for fruitful discussions and suggestions.

References

Aïd, R., Gruet, P., & Pham, H. (2016). An optimal trading problem in intraday electricity markets. *Mathematics and Financial Economics*, *10*(1), 49–85.

Amjady, N., & Keynia, F. (2009). Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm. *IEEE Transactions on Power Systems*, *24*(1), 306–318.
Andrade, J., Filipe, J., Reis, M., & Bessa, R. (2017). Probabilistic price forecasting for day-ahead and intraday markets: beyond the statistical model. *Sustainability*, *9*(11), 1990.
Barnes, A. K., & Balda, J. C. (2013). Sizing and economic assessment of energy storage with real-time pricing and ancillary services. In *PEDG 2013 Conference Proceedings*. http://dx.doi.org/10.1109/PEDG.2013.6785651.
Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*, 253–263.
EPEX (2015). New ID3-price index on German intraday continuous market. Report no. 2015-21, ver. July 2015, http://www.epexspot.com/en/extras/download-center.
EPEX (2018). Annual Report 2017, http://www.epexspot.com/en/extras/download-center.
Gianfreda, A., Parisio, L., & Pelagatti, M. (2016). The impact of RES in the Italian day-ahead and balancing markets. *Energy Journal*, *37*, 161–184.
Gianfreda, A., Parisio, L., & Pelagatti, M. (2018). A review of balancing costs in Italy before and after RES introduction. *Renewable & Sustainable Energy Reviews*, *91*, 549–563.
González-Aparicio, I., & Zucker, A. (2015). Impact of wind power uncertainty forecasting on the market integration of wind energy in Spain. *Applied Energy*, *159*, 334–349.
Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, *32*, 914–938.
Hubicka, K., Marcjasz, G., & Weron, R. (2019). A note on averaging day-ahead electricity price forecasts across calibration windows. *IEEE Transactions on Sustainable Energy*, *10*(1), 321–323.
Janczura, J., Trück, S., Weron, R., & Wolff, R. (2013). Identifying spikes and seasonal components in electricity spot price data: a guide to robust modeling. *Energy Economics*, *38*, 96–110.
Karakatsani, N., & Bunn, D. (2008). Forecasting electricity prices: the impact of fundamentals and time-varying coefficients. *International Journal of Forecasting*, *24*, 764–785.
Keles, D., Scelle, J., Paraschiv, F., & Fichtner, W. (2016). Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. *Applied Energy*, *162*, 218–230.
Kiesel, R., & Kusterman, M. (2016). Structural models for coupled electricity markets. *Journal of Commodity Markets*, *3*(1), 16–38.
Kiesel, R., & Paraschiv, F. (2017). Econometric analysis of 15-minute intraday electricity prices. *Energy Economics*, *64*, 77–90.
Ludwig, N., Feuerriegel, S., & Neumann, D. (2015). Putting big data analytics to work: Feature selection for forecasting electricity prices using the LASSO and random forests. *Journal of Decision Systems*, *24*(1), 19–36.
Maciejowska, K., Nitka, W., & Weron, T. (2019). Day-ahead vs. intraday – forecasting the price spread to maximize economic benefits. *Energies*, *12*(4), 631.
Maciejowska, K., & Nowotarski, J. (2016). A hybrid model for GEFCom2014 probabilistic electricity price forecasting. *International Journal of Forecasting*, *32*(3), 1051–1056.
Märkle-Huß, J., Feuerriegel, S., & Neumann, D. (2018). Contract durations in the electricity market: causal impact of 15 min trading on the EPEX SPOT market. *Energy Economics*, *69*, 367–378.
Mayer, K., & Trück, S. (2018). Electricity markets around the world. *Journal of Commodity Markets*, *9*, 77–100.
Misiorek, A. (2008). Short-term forecasting of electricity prices: Do we need a different model for each hour? *Medium Econometrisch Toepassingen*, *16*(2), 8–13.
Monteiro, C., Ramirez-Rosado, I., Fernandez-Jimenez, L., & Conde, P. (2016). Short-term price forecasting models based on artificial neural networks for intraday sessions in the Iberian electricity market. *Energies*, *9*(9), 721.
Nowotarski, J., & Weron, R. (2018). Recent advances in electricity price forecasting: a review of probabilistic forecasting. *Renewable & Sustainable Energy Reviews*, *81*, 1548–1568.
Pape, C., Hagemann, S., & Weber, C. (2016). Are fundamentals enough? Explaining price variations in the German day-ahead and intraday power market. *Energy Economics*, *54*, 376–387.
Schneider, S. (2011). Power spot price models with negative prices. *Journal of Energy Markets*, *4*(4), 77–102.
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B.*, *58*, 267–288.
Uniejewski, B., Nowotarski, J., & Weron, R. (2016). Automated variable selection and shrinkage for day-ahead electricity price forecasting. *Energies*, *9*, 621.
Uniejewski, B., & Weron, R. (2018). Efficient forecasting of electricity spot prices with expert and lasso models. *Energies*, *11*, 2039.
Uniejewski, B., Weron, R., & Ziel, F. (2018). Variance stabilizing transformations for electricity spot price forecasting. *IEEE Transactions on Power Systems*, *33*, 2219–2229.
Weron, R. (2006). *Modeling and forecasting electricity loads and prices: A statistical approach*. John Wiley & Sons, Chichester.
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.
Zaleski, P., & Klimczak, D. (2015). Prospects for the rise of renewable sources of energy in Poland. balancing renewables on the intraday market. In K. Zamasz (Ed.), *Capacity market in contemporary economic policy* (pp. 124–138). Difin.
Ziel, F. (2016). Forecasting electricity spot prices using LASSO: on capturing the autoregressive intraday structure. *IEEE Transactions on Power Systems*, *31*(6), 4977–4987.
Ziel, F. (2017). Modeling the impact of wind and solar power forecasting errors on intraday electricity prices. In *IEEE conference proceedings — EEM17*. http://dx.doi.org/10.1109/EEM.2017.7981900.
Ziel, F., Steinert, R., & Husmann, S. (2015). Efficient modeling and forecasting of electricity spot prices. *Energy Economics*, *47*, 89–111.
Ziel, F., & Weron, R. (2018). Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks. *Energy Economics*, *70*, 396–420.

Bartosz Uniejewski is a Master student of Applied Mathematics at the Wrocław University of Science and Technology, Poland. Despite being a (very) early stage researcher, he has given talks on electricity price forecasting at conferences and workshops in Essen (Germany), Kraków (Poland), Milan (Italy), Oxford (UK), Padova (Italy) and Trondheim (Norway). He is the recipient of the prestigious scholarship for outstanding students of the Ministry of Higher Education in Poland (2016, 2017, 2018).

Grzegorz Marcjasz is a Master student of Applied Mathematics at the Wrocław University of Science and Technology, Poland. He is a proficient programmer in Matlab and Python. His research interests include energy forecasting and machine learning. He is the recipient of the prestigious scholarship for outstanding students of the Ministry of Higher Education in Poland (2017, 2018) and the Best Student Paper prize (runner-up) at the Computational Management Science conference in Trondheim (2018).

Rafał Weron is Professor of Economics in the Department of Operations Research, Wrocław University of Science and Technology, Poland. His research focuses on developing risk management and forecasting tools for the energy industry and computational statistics as applied to finance and insurance. He is the author of the widely acclaimed Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach (Wiley, 2006) and co-author of over 80 peer-reviewed book chapters and journal articles. His review paper on electricity price forecasting was selected for the 2017 edition of the Emerald Citations of Excellence and received the 2017 IIF Tao Hong Award for the best International Journal of Forecasting paper on energy forecasting. With a Ph.D. (1999) in Financial Mathematics and a habilitation (2009) and professor title (2015) in Economics, he is periodically engaged as a consultant to financial, energy and software engineering companies.

Please cite this article as: B. Uniejewski, G. Marcjasz and R. Weron, Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO. International Journal of Forecasting (2019), https://doi.org/10.1016/j.ijforecast.2019.02.001.