HSC Research Report

HSC/18/07

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

Bartosz Uniejewski^{1,2}
Grzegorz Marcjasz^{1,2}
Rafał Weron¹

¹ Department of Operations Research, Faculty of Computer Science and Management, Wrocław University of Science and Technology, Poland

² Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Poland

Hugo Steinhaus Center
Wrocław University of Science and Technology
Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland
<http://www.im.pwr.wroc.pl/~hugo/>

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*

Abstract

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

Keywords: Intraday electricity market, variable selection, price forecasting, LASSO

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 is traded under market rules as any other commodity (Mayer and Trück, 2018). In Europe, the workhorse of power trading has been the uniform price auction conducted a day before delivery, and a vast majority of research and applications has concerned day-ahead electricity prices (Weron, 2014). However, the expansion of renewable generation (mostly wind and solar), power grid modernization (including increase of 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 and Fan, 2016; Kiesel and Kusterman, 2016). This has amplified the importance of intraday markets, which can be used to balance deviations resulting from positions in day-ahead contracts and the actual demand (Gianfreda et al., 2016; Märkle-Huß et al., 2018; Zaleski and Klimczak, 2015). As a result, during the last few years we have observed a shifting of volume from the day-ahead to intraday markets across Europe (EPEX, 2018).

In this article we use a unique set of prices from the German EPEX market and take a closer look at the fine structure of intraday markets with its continuous trading for individual load periods up to 30 minutes before delivery. We apply the *least absolute shrinkage and selection operator* (LASSO) of Tibshirani (1996) to gain statistically sound insights on variable selection and provide

Email addresses: uniejewskibartosz@gmail.com (Bartosz Uniejewski), gelusz@hotmail.co.uk (Grzegorz Marcjasz), rafal.weron@pwr.edu.pl (Rafał Weron)

recommendations for very short-term *electricity price forecasting* (EPF).¹ Given that the literature on forecasting intraday prices in European power markets is very scarce and, to our best knowledge, limited to only two papers dealing with Spanish data (Andrade et al., 2017; Monteiro et al., 2016), our study is a major step forward 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. The importance of our study is further emphasized by the fact that electricity price forecasts – alongside weather and demand predictions – are nowadays fundamental inputs to energy companies’ decision-making mechanisms (Nowotarski and Weron, 2018).

The remainder of the paper is structured as follows. In Section 2 we review the literature on intraday electricity markets and on variable selection for EPF. In Section 3, we first introduce the EPEX dataset and the rolling window scheme, then discuss variance stabilization and, finally, describe the considered model structures. In Section 4, we first compare the predictive performance in terms of two commonly used error measures and the Diebold and Mariano (1995) test, then take a closer look at the best LASSO-estimated model to identify the most important explanatory variables and thus provide guidelines to structuring better performing models for intraday electricity markets. Finally, in Section 5, we wrap up the results and conclude.

2. Literature Review

2.1. Intraday Markets for Electricity

There is a small, but increasing in numbers literature on intraday electricity markets. Most publications take a fundamental, market behavior or energy policy perspective. For instance, Pape et al. (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 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 the bidding behavior 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 to the volume of trades dependent on a threshold variable, which reflects the expected non-renewable generation in the day-ahead market. Aïd et al. (2016) analyze trading in intraday electricity markets and develop an optimal bidding strategy. As a modeling framework they consider a continuous-time stochastic process for the net position of sales and purchases of electricity. Finally, Märkle-Huß et al. (2018) investigate the introduction of 15 minute contracts in the German EPEX market and argue that these products are used to balance the intra-hour volatility of renewable energy sources.

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

However, when it comes to forecasting of intraday electricity prices in European power markets the literature is very scarce. To our best knowledge, only two studies address this important problem and only in the context of the Iberian electricity market (MIBEL), which features a very specific design with six intraday sessions spanning from 9 to 27 delivery hours. Monteiro et al. (2016) utilize neural networks (multilayer perceptron, one hidden layer) with up to 21 ad-hoc selected input variables: 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, 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 utilizing a Linear Quantile Regression (LQR) model, i.e., that high quality point and probabilistic forecasts of intraday prices can be obtained by just exploring the prices from previous sessions (plus deterministic variables to model the daily, weekly and annual seasonalities). And this despite considering a large set of fundamental variables.

2.2. Variable Selection for Electricity Price Forecasting

Conclusions from the latter two studies clearly suggest that variable selection is a very important issue in EPF, and because of the vast amounts of available data, it may be even more critical for intraday than for day-ahead markets. 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 Karakatsani and Bunn (2008) and Misiorek (2008) who use *stepwise regression* to eliminate statistically insignificant variables in parsimonious regression-type models and Amjady and Keynia (2009) who introduce a feature selection algorithm based on *mutual information*. To our best knowledge, Barnes and Balda (2013) were the first to apply regularization in day-ahead EPF. In a study concerning the profitability of battery storage, they utilize *ridge regression* to compute EPFs for a model with more than 50 regressors. Ludwig et al. (2015) use *random forests* and the *least absolute shrinkage and selection operator* (LASSO) to choose which of the 77 available weather stations were relevant, while Keles et al. (2016) combine 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 of Ziel et al. (2015) and Ziel (2016), who used LASSO to sparsify very large (100+) sets of model parameters, either utilizing B-splines in a univariate setting or, more efficiently, within a multivariate framework. In the first thorough comparative study, Uniejewski et al. (2016) evaluate 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 conclude that using LASSO and elastic nets can bring significant accuracy gains compared to commonly used EPF models. In a study on the optimal model structure for day-ahead EPF, Ziel and Weron (2018) consider autoregressive models with 200+ potential explanatory variables (but not exogenous) and conclude that both uni- and multivariate LASSO-implied structures significantly outperform autoregressive benchmarks and that combining their forecasts can bring 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 typically considered EPF models.

In this study we follow the approach set forth in the latter three articles and consider LASSO models based on hundreds of potential regressors and calibrated to asinh-transformed prices. We do not, however, consider recalibrating the LASSO regularization parameter, as this considerably slows down the forecasting procedure.

3. Methodology

3.1. The Dataset

The main German ‘spot’ market is operated by EPEX SPOT SE and admits trading of power supply contracts with hourly and quarter hourly delivery. Participants have the option of bidding hourly products in the day-ahead auction conducted at noon on the day before delivery (i.e., $d - 1$) or trading the hourly and quarter hourly contracts in the continuous intraday market which opens at 16:00 on day $d - 1$ and closes 30 minutes before the delivery starts (EPEX, 2018).

The leading reference price for the intraday market is the recently introduced ID3 index for contracts with hourly delivery. It takes into account only the most recent trades, i.e., transactions that took place no earlier than 3 hours before delivery. The exchange publishes the index, however, the covered period is too short for a proper evaluation of our models. Therefore we have reconstructed an ID3-like time-series from the individual transactions. It slightly differs from the actual index, because (i) we have disregarded block contracts which constitute only a small percentage of the volume traded (especially when the length of the block contract vs. occurrence of the trade in the last 3 hours before delivery is taken into consideration), and (ii) we have not excluded the so-called cross-trades, i.e., trades with the same entity buying and selling (the data we have access to is anonymized).

Apart from the ID3 index (actually its approximation), we are also using the day-ahead prices as external regressors. Both time series are of hourly resolution and span 1216 days ranging from 1.01.2015 to 30.04.2018, see Figure 1. Like the majority of EPF studies, we consider a rolling window scheme; we use a 364-day window. Initially, we fit our models to data from 1.01.2015 hour 1 to 30.12.2015 hour 24, and compute the price forecasts for the first hour of 31.12.2015. Next, the window is rolled forward by 1 hour 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 long out-of-sample test period (i.e., 30.04.2018 hour 24) are made.

3.2. The Forecasting Framework

We denote by P_t the intraday and by S_t the day-ahead electricity price at time (hour) $t = 24d + h$, where d is the day and h 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 Figure 2 using actual transaction data for the period from 12.09.2016 16:00 to 13.09.2016 24:00. Observe the 4-hour time lag between the moment the forecast is made and the time the delivery starts. For instance, at 12:00 on 13.09.2016 we are forecasting the price for 16:00

Figure 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 long out-of-sample test period.

(denoted by: $\rightarrow$). The most recent intraday price (ID3 index value) is for 12:00 (denoted by: *), i.e., the hourly contract with delivery between 12:00 and 13:00 on 13.09.2016.

Following the recommendations set forth in Uniejewski et al. (2018), we do not calibrate our models (except for the three **naïve** benchmarks; see Section 3.3) to raw prices but to transformed, i.e., $X_t = f(P_t)$, where $f(\cdot)$ is an appropriately chosen *variance stabilizing transformation* (VST). The idea underlying a VST is that of reducing the price variation. As argued by Janczura et al. (2013), 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, since EPEX prices exhibit negative values², the log-transform is not feasible in our case. Instead, we utilize the *area hyperbolic sine* transformation:

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

²Note, 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 et al., 2018; Schneider, 2011; Weron, 2006).

Figure 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 the delivery starts (every hour of 13.09.2016), the circles refer to actual trades (circle size represents the traded volume – from 0.1 to 200 MWh, color represents the price – see the colorbar on the *right*) and the red step function represents the time the forecasts are made. For instance, at 12:00 on 13.09.2016 when forecasting the price for 16:00 ($\rightarrow$), the most recent ID3 value is for 12:00 (*).

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 **asinh** 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 and Weron, 2018; Ziel and 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:

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

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

3.3. Benchmark Models

The first benchmark, denoted by **Naive1**, is a very popular reference model in day-ahead EPF (Contreras et al., 2003; Weron, 2014). It is defined by $\widehat{P}_t = P_{t-168}$ for Monday, Saturday and Sunday, and $\widehat{P}_t = P_{t-24}$ otherwise. The second, denoted by **Naive2**, also belongs to the class of self-similar techniques (Weron, 2006). It is defined by $\widehat{P}_t = P_{t-4}$, i.e., we assume that the last

observed intraday price (i.e., P_{t-4} ; note, that we have a 3 hour lag due to using the ID3 index) is the best estimate of the intraday price at time t . The third, denoted by **Naive3**, is based on the assumption that 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 day-ahead price for the same day and hour (recall, that it is set at noon on day $d - 1$).

Finally, the fourth benchmark is a parsimonious autoregressive structure inspired by the well-performing *expert_{DoW,nl}* model of Ziel and Weron (2018). Within this model, denoted by **ARX**, the VST-transformed price at time t is given by:

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

where X_{t-24} , X_{t-48} and X_{t-168} account for the autoregressive effects of the previous days (the same hour yesterday, two days ago, and one week ago), X_{t-4} is the last observed intraday price and Y_t is the VST-transformed day-ahead price for same day and hour. The seven dummy variables $D_1, \dots, D_7$ account for the weekly seasonality, and are defined as $D_i = 1$ for $d = i$ and zero otherwise. Finally, the ε_t 's are assumed to be independent and identically distributed normal variables.

3.4. LASSO-estimated Models

An advantage of using automatic variable selection is an almost unlimited number of initially considered explanatory variables. In this study we utilize a baseline model with 349 potential regressors: 165 last known prices from the intraday market (i.e., nearly the whole week), 177 prices from the day-ahead market (i.e., one week of past prices plus nine, already known ‘future’ hourly prices) and seven dummy variables (to account for the weekly seasonality, like in the **ARX** benchmark):

$$X_t = \underbrace{\sum_{i=4}^{168} \beta_{i-3} X_{t-i}}_{\text{past intraday prices}} + \underbrace{\sum_{j=-8}^{168} \beta_{174+j} Y_{t-j}}_{\text{day-ahead prices}} + \underbrace{\sum_{k=1}^7 \beta_{343+k} D_k}_{\text{weekday dummies}} + \varepsilon_t. \quad (4)$$

Note, that the price for each hour is predicted with a 3-hour time lag (hence the first sum in the above formula starts with $i = 4$) and using the most recent information. Note also, that we do not consider any fundamental regressors in Eqn. (4). 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 day-ahead price for the same day and hour already includes this information.

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

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

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

The *least absolute shrinkage and selection operator* (LASSO) of Tibshirani (1996) can be treated as a generalization of linear regression, where instead of minimizing only the *residual sum of squares* (RSS), the sum of RSS and a linear penalty function of the β 's is minimized:

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

where $\lambda \geq 0$ is a *tuning* (or *regularization*) parameter. Note that for $\lambda = 0$, we get the standard least squares estimator, for $\lambda \rightarrow \infty$, all $\beta_{h,i}$ s tend to zero, while for intermediate values of λ , there is a balance between minimizing the RSS and 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 $i = 1, \dots, 10$. The obtained models are denoted by **LASSO**(λ_i) or simply by λ_i . Because of a relatively short dataset compared to the one used in Uniejewski and Weron (2018), we have not opted for reselecting λ based on model performance in a validation period, instead we have decided to show the results for all 852 days in the out-of-sample test period and all ten λ 's.

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 Figure 1) as the main evaluation criteria:

$$\text{MAE} = \frac{1}{24D} \sum_{t=1}^{24D} |\mathcal{E}_t| \quad \text{and} \quad \text{RMSE} = \sqrt{\frac{1}{24D} \sum_{t=1}^{24D} \mathcal{E}_t^2}, \quad (7)$$

where $\mathcal{E}_t = P_t - \widehat{P}_t$ 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 obtained MAE and RMSE values can be used to provide a ranking of models, but do not allow to draw statistically significant conclusions on the outperformance of the forecasts of one model by those of another. Therefore, we use the Diebold and Mariano (1995) test, which is simply an asymptotic z -test of the hypothesis that the mean of the loss differential series:

$$\Delta_{A,B,t} = |\mathcal{E}_{A,t}|^i - |\mathcal{E}_{B,t}|^i \quad (8)$$

is zero, where $\mathcal{E}_{Z,t}$ is the prediction error of model Z at time t and $i = 1, 2$ correspond to the absolute and squared loss, respectively; by ‘model Z ’ we mean here one of the four benchmarks or the ten LASSO-estimated models. For each model pair and each dataset we compute the p -value 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 obtained loss differential series are covariance stationary.

Table 1: *Mean Absolute Errors (MAE) and Root Mean Squared Errors (RMSE) for the four benchmarks (Naive1, ..., ARX) and the ten LASSO(λ_i) models with $\lambda_i = 10^{-\frac{19-i}{6}}$, $i = 1, \dots, 10$, in the 852-day long out-of-sample test period, see Figure 1.*

	Naive1	Naive2	Naive3	ARX	LASSO									
					λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}
MAE	10.4423	9.8295	5.0391	4.7152	5.3636	5.0500	4.8157	4.6396	4.5118	4.4398	4.4170	4.4511	4.5416	4.6691
RMSE	16.9180	14.2516	8.0862	7.6111	8.2947	7.8127	7.4896	7.2815	7.1266	7.0633	7.0747	7.1705	7.3603	7.6098

Figure 3: Results of the Diebold-Mariano (DM) test for the absolute (*left*) and squared (*right*) prediction errors, and the same 14 models as in Table 1. A heat map is used to indicate the range of the 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 the forecasts of a model on the Y-axis (worse).

In Table 1, we report the MAE and RMSE errors for the four benchmarks and the ten LASSO models, while in Figure 3 we depict the results of the DM test for these 14 models. We use a heat map to indicate the range of the p -values – the closer they are to zero ($\rightarrow$ dark green) the more significant is 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 Figure 3 is green, indicating that – irrespective of whether we are considering absolute or squared losses – the forecasts of the **Naive1** benchmark are significantly outperformed by those of all other models. On the other hand, the column which corresponds to the **LASSO(λ_7)** model is green in the left panel, meaning that – when considering absolute losses – this model leads to significantly better forecasts than all other.

Looking at Table 1 we can clearly see that the **Naive1** and **Naive2** benchmarks are the worst predictors. While this is not so surprising for the former model, the poor performance of **Naive2**, which is based on the most recent intraday price P_{t-4} , may be unexpected. Apparently, the 3-hour

lag leads to large errors because of the intraday seasonality. On the other hand, **Naive3** works reasonably well and for absolute losses it even significantly outperforms the worst LASSO model, i.e., λ_1 , see the green squares in the left panel of Figure 3 in the **Naive3** column. Obviously, the day-ahead 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 benchmarks and the three (for absolute losses) or two (for squared losses) worst LASSO models, i.e., λ_1 , λ_2 and λ_3 . Now, comparing the LASSO-estimated models among themselves, we can see that the best predictions are obtained for λ_7 (according to MAE) or λ_6 (according to RMSE). However, while for absolute losses the λ_7 is a clear winner, for squared losses there are no significant differences in predictive accuracy between λ_6 and λ_7 . The overall picture is consistent with other recent studies that use the LASSO for day-ahead EPF (Uniejewski et al., 2016; Uniejewski and Weron, 2018; Ziel, 2016; Ziel and Weron, 2018).

4.2. Variable Selection

Let us now comment on variable selection by looking at the results of the best performing model, i.e., **LASSO**(λ_7). In Table 2 and 3 we report the mean occurrence (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 top row in Table 2 with ‘100’ in all columns), as well as the day-ahead price corresponding to the same (i.e., Y_t ; see the row labeled ‘0’ in Table 3 with ‘100’ in nearly all columns) or nearby hours (lags -2 , -1 and 1). Note, however, that the most recent intraday price by itself does not yield an accurate prediction – the **Naive2** benchmark, i.e., $\widehat{P}_t = P_{t-4}$, performs rather poorly in Table 1 and Figure 3.
- 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 and Keynia, 2009; Karakatsani and Bunn, 2008; Keles et al., 2016; Uniejewski et al., 2016; Ziel and Weron, 2018). The ‘greener’ part of the row for $h = 20, \dots, 23$ can be interpreted as the impact of the late evening hours (see the next point).
- As observed by Maciejowska and Nowotarski (2016) and Ziel (2016), the price for hour 24, but also the nearby evening hours are important predictors. This can be seen as yellow-green diagonals in Table 2. Note, that the first full diagonal (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.
- What also comes as a surprise, the dummies are usually removed by the LASSO (except for three morning hours on Sunday), see the last 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 and Weron, 2018).

Table 2: Mean occurrence (in %) of model parameters across the 852-day long out-of-sample test period. Columns represent the hours ($h = 1, \dots, 24$) for which the price predictions are made and rows represent the parameters (β_{i-3}) corresponding to the past intraday prices (X_{t-i} , $i = 4, \dots, 168$) of the **LASSO**(λ_7) model, see Eqn. (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
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
5	62	2	40	6	6	0	32	42	39	0	33	78	0	0	0	1	50	0	0	0	0	11	23	35
6	15	67	40	0	1	3	58	9	54	2	4	40	60	3	5	4	19	67	0	0	3	1	3	20
7	19	23	86	26	0	2	3	1	8	8	0	17	50	27	18	73	6	65	16	0	12	3	13	13
8	6	19	34	70	19	0	21	0	2	0	0	0	26	15	36	0	80	44	38	2	3	1	0	31
9	1	21	4	1	25	35	31	42	0	0	0	6	18	26	25	40	0	96	85	2	23	3	32	0
10	0	0	0	12	0	11	25	26	53	1	4	3	15	4	15	37	42	8	63	65	10	12	11	5
11	7	1	0	0	7	0	3	50	21	27	0	15	9	14	6	0	4	22	56	15	42	4	33	4
12	11	3	2	0	21	0	3	0	37	11	38	4	6	7	2	10	0	7	37	79	11	2	2	19
13	4	0	31	2	0	24	0	0	15	55	37	68	5	3	21	1	0	0	5	43	71	0	5	0
14	8	0	42	15	0	27	27	0	0	3	71	13	53	20	1	0	6	0	0	6	53	70	1	2
15	13	0	4	74	8	4	26	25	0	2	13	44	20	52	10	1	9	0	3	0	8	50	12	18
16	4	10	10	2	22	1	0	10	17	0	20	0	11	4	29	20	13	16	0	6	0	2	16	3
17	0	15	20	0	0	14	2	9	15	0	7	16	2	11	0	6	24	37	4	0	2	0	30	4
18	2	0	3	3	0	0	1	0	9	8	11	20	23	3	2	1	8	1	45	2	0	0	4	7
19	31	12	15	0	0	2	0	7	0	10	8	16	3	19	33	0	9	53	7	44	2	0	4	12
20	15	44	0	39	0	4	0	10	10	0	37	6	31	3	22	36	0	3	18	14	52	11	8	9
21	16	41	15	11	14	0	2	1	3	0	22	56	16	17	6	9	76	0	0	11	17	55	0	23
22	39	8	30	42	14	9	0	0	1	13	0	29	44	28	6	6	5	88	0	40	18	23	17	5
23	0	40	27	38	47	29	46	11	0	7	22	0	19	30	19	11	48	8	100	18	81	38	44	37
24	0	1	35	20	31	36	19	41	15	0	10	18	0	15	0	0	18	41	3	100	57	73	81	47
25	16	11	29	0	0	9	2	17	30	0	0	1	0	0	3	1	0	0	0	48	0	26	68	
26	46	11	13	3	0	3	0	0	9	26	1	3	4	0	6	34	0	26	3	0	4	10	14	
27	3	7	17	2	0	0	0	10	0	1	20	3	7	24	2	15	25	42	10	100	23	0	15	14
28	10	0	1	16	15	9	0	1	2	5	0	19	12	0	29	0	22	17	46	30	96	46	0	28
29	19	17	0	33	26	6	4	0	9	2	4	3	10	3	0	2	0	12	9	6	5	98	2	1
30	2	24	26	17	48	0	6	4	1	1	2	0	10	43	2	29	2	0	0	4	0	20	52	28
31	13	1	19	10	7	5	21	10	0	12	2	0	0	28	38	13	33	4	2	6	11	0	7	19
32	0	7	1	1	22	0	14	10	0	2	6	0	0	0	24	1	20	9	1	6	0	4	0	3
33	0	0	0	2	4	3	0	21	10	0	17	0	0	0	0	0	1	22	1	1	0	0	0	0
34	7	0	0	5	0	15	3	1	12	4	0	8	0	5	4	18	17	12	17	2	2	2	0	0
35	0	0	0	11	0	16	31	18	1	10	0	0	0	6	10	11	8	1	14	10	28	4	9	0
36	0	0	1	0	5	0	3	5	14	0	13	0	1	0	10	2	4	14	9	12	6	41	8	0
37	0	0	0	10	0	7	2	8	22	9	0	17	29	3	0	8	2	3	17	17	19	7	3	6
38	16	0	15	2	0	0	0	5	5	1	2	0	16	45	0	4	0	6	2	9	3	20	0	0
39	0	20	0	9	0	7	1	0	1	5	7	31	0	7	29	0	39	5	3	1	1	28	8	10
40	54	21	13	0	0	12	4	2	0	10	9	29	13	0	13	19	0	20	0	22	0	0	53	21
41	8	41	31	10	0	22	10	25	4	14	12	0	62	18	0	3	4	2	14	3	13	1	0	9
42	44	27	19	25	20	0	25	3	12	6	3	0	3	62	11	17	0	0	2	6	3	15	4	9
43	0	14	10	18	26	12	3	8	8	5	0	0	7	11	3	14	17	10	5	1	0	3	38	8
44	0	6	17	9	5	0	1	10	0	17	4	4	18	0	0	0	3	0	17	7	0	0	3	14
45	0	27	15	2	14	0	0	0	0	9	6	12	37	23	0	0	10	0	38	2	4	0	15	
46	4	0	16	27	26	1	30	0	1	5	1	1	1	24	48	76	4	0	3	0	25	2	0	0
47	0	11	2	30	28	15	24	37	0	5	5	8	0	5	18	29	92	24	3	10	2	21	0	11
48	36	0	15	1	7	9	0	23	3	0	7	10	8	3	0	0	20	77	9	15	50	52	33	12
49	0	32	0	7	31	0	12	0	25	0	1	15	4	3	0	9	0	0	16	9	8	13	19	7
50	6	0	5	0	4	13	0	33	15	17	4	1	3	2	0	0	2	7	0	26	0	7	1	
51	1	13	0	1	29	0	10	1	41	21	12	0	0	4	1	8	10	0	1	6	7	28	1	0
52	50	34	27	1	0	38	0	5	0	8	13	0	2	0	10	13	19	16	19	41	24	51	24	3
53	18	13	24	30	0	0	5	0	0	3	4	8	15	3	18	27	46	57	7	23	16	4	2	
54	0	0	14	13	35	2	0	1	0	0	0	6	6	11	14	2	20	34	26	4	0	7	5	
55	0	3	5	1	0	14	1	4	3	0	0	1	8	18	5	19	0	14	4	28	4	0	13	9
56	0	0	0	1	0	0	44	1	5	3	0	0	8	0	36	0	7	0	9	0	40	0	0	4
57	0	19	0	9	0	0	6	8	33	6	9	0	0	5	1	13	9	4	2	0	5	42	1	6
58	0	0	20	0	5	2	16	0	48	49	4	1	0	0	1	36	29	10	0	27	2	24	42	4
59	0	1	3	12	3	4	10	12	4	33	61	3	6	3	0	7	39	45	25	2	27	11	21	3
60	0	0	1	7	38	0	14	4	18	1	54	54	4	5	0	0	10	23	20	1	30	27	1	
61	0	0	0	0	10	38	12	9	8	30	4	43	18	0	0	0	15	0	30	0	5	11	4	
62	5	0	0	0	0	0	33	33	10	19	35	5	30	5	1	6	0	0	6	0	23	4	3	1
63	2	0	0	27	0	0	6	2	34	1	39	2	7	23	15	16	0	0	0	1	8	29	35	1
64	4	3	0	0	0	0	8	4	0	2	12	32	0	13	4	9	0	1	0	1	2	17	36	4
65	8	21	1	2	0	0	9	0	10	0	1	18	5	2	12	2	12	17	0	0	25	0	21	22
66	3	2	28	5	5	2	2	2	1	9	5	1	40	10	19	0	17	19	21	0	8	13	24	13
67	12	0	5	44	2	0	6	3	4	15	6	36	3	21	0	23	0	11	2	33	0	0	4	0
68	2	18	0	1	17	8	0	8	2	9	15	0	30	19	21	0	21	10	0	0	1	0	0	0
69	0	11	17	0	0	49	46	5	7	2	28	11	0	18	33	17	1	1	5	2	6	4	10	0
70	2	5	30	39	4	5	2	33	1	0	9	11	3	3	20	40	12	1	6	0	6	24	0	0
71	6	5	21	14	14	25	5	2	3	0	7	0	11	0	2	10	40	28	1	7	23	25	20	0
72	0	7	30	24	50	12	18	29	7	8	4	0	0	0	0	5	8	31	11	1	18	26	37	23
73	61	1	23	11	25	9	0	13	35	9	2	13	5	0	0	3	2	0	5	0	0	0	26	9
74	0	59	21	45	27	30	6	4	2	32	12	0	0	17	2	3	0	0	2	11	0	1	7	4
75	0	0	35	1	22	25	13	5	15	12	6	0	1	18	1	7	0	0	1	7	17	12	86	20
76	34	5	7	0	0	14																		

Table 2: *Cont.*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
91	0	9.2	13	5.5	2.8	4.3	0	0.1	25	17	0	17	0	1.8	0	26	11	5.4	0	13	10	0	26	3.6
92	0	3	6	17	7	1	0	0	5	5	12	1	17	0	5	4	0	0	5	31	3	0	0	0
93	1	0	4	2	27	12	18	4	1	3	11	13	2	16	23	19	9	8	0	2	32	0	17	1
94	3	9	6	2	14	22	10	16	12	0	6	4	6	6	4	15	13	16	3	0	0	0	0	19
95	14	0	5	4	5	5	4	4	25	1	0	15	3	7	0	0	37	3	52	3	2	11	0	0
96	13	0	4	5	15	21	26	11	3	32	3	0	16	2	2	0	0	46	20	74	66	38	17	0
97	0	12	0	0	14	32	0	1	1	4	23	11	1	3	0	0	0	0	34	21	1	17	22	58
98	25	0	9	0	2	28	4	18	15	1	4	1	16	13	0	0	0	0	35	1	0	5	0	0
99	8	7	0	2	0	2	49	0	56	29	0	0	6	5	20	0	15	5	1	0	0	0	1	7
100	0	0	0	0	7	11	8	18	5	44	44	1	0	0	19	0	0	0	12	6	43	0	0	1
101	0	0	0	10	6	0	0	0	0	18	24	39	45	0	2	18	0	0	0	6	3	8	6	1
102	0	0	7	0	0	1	0	0	0	6	19	31	11	40	0	31	24	1	0	0	18	5	26	0
103	12	0	3	4	1	3	1	0	0	0	42	7	16	12	35	0	3	0	0	0	0	24	3	0
104	17	2	0	25	6	3	2	0	6	0	9	6	0	12	3	13	1	0	0	0	0	0	2	8
105	19	10	0	1	0	32	2	0	0	0	13	3	2	0	22	0	2	9	2	0	7	10	9	9
106	15	27	0	2	0	1	35	3	0	2	0	5	0	7	0	13	6	9	28	28	1	8	12	17
107	4	31	0	19	1	3	45	40	8	0	13	6	3	0	11	0	17	1	68	9	24	4	20	4
108	2	10	12	0	12	3	4	47	58	4	0	11	0	0	1	49	27	2	0	45	10	7	10	9
109	5	0	3	2	0	18	5	1	41	48	5	2	15	0	0	3	35	0	4	16	38	3	33	12
110	2	1	17	4	1	3	0	2	0	31	27	1	0	10	0	0	5	0	0	0	14	3	24	23
111	0	0	0	0	3	0	0	14	0	5	14	0	10	10	0	0	7	0	1	1	0	4	57	0
112	0	0	12	0	3	1	0	0	6	7	3	1	4	12	6	4	2	0	10	0	0	1	2	1
113	21	0	0	8	4	6	21	0	0	6	6	27	3	2	0	7	6	0	0	1	0	0	10	1
114	33	38	0	1	0	18	6	0	0	22	15	14	8	1	17	17	4	8	0	0	0	0	1	2
115	6	45	48	20	26	0	11	1	0	11	16	9	11	7	8	22	8	0	29	1	0	0	1	4
116	0	0	34	45	40	2	5	14	0	6	20	18	6	12	12	0	24	4	0	2	0	0	0	7
117	0	0	0	0	0	6	26	9	8	19	8	2	22	14	3	13	36	12	3	5	0	2	6	0
118	0	0	0	9	0	20	54	0	26	2	3	16	24	1	1	20	2	1	4	0	2	6	3	1
119	0	0	2	0	0	6	26	9	8	19	8	2	22	14	3	13	36	12	3	5	0	2	6	0
120	1	0	0	0	0	9	0	33	36	14	2	2	0	2	11	8	6	27	9	69	5	9	2	5
121	2	0	0	0	0	2	9	1	1	51	4	0	0	1	0	0	16	10	8	48	13	20	7	0
122	0	0	0	1	0	0	0	12	5	6	48	0	0	0	1	6	0	2	43	1	0	1	52	5
123	0	33	1	0	32	0	10	0	11	1	20	8	0	0	0	0	2	0	0	39	14	0	0	42
124	41	1	45	5	0	52	0	6	14	9	0	41	4	12	2	0	0	0	0	19	25	1	13	0
125	0	37	7	34	0	0	0	5	11	16	3	1	13	12	0	3	0	2	11	0	0	27	4	0
126	19	0	22	0	16	25	0	0	0	12	10	4	0	20	32	10	3	5	0	1	0	0	10	1
127	6	38	0	34	0	23	39	9	6	1	9	9	3	5	30	29	17	10	0	0	20	0	0	2
128	5	3	31	1	2	0	20	16	9	0	5	3	2	8	21	12	10	4	6	7	0	1	0	0
129	14	0	0	25	7	0	2	22	12	2	0	5	11	2	12	16	5	5	6	2	0	1	1	3
130	2	29	1	0	1	2	0	27	30	10	0	0	6	9	3	3	24	12	18	0	0	3	13	20
131	7	0	8	0	0	0	8	5	14	24	14	0	0	0	16	2	0	17	24	5	0	3	2	1
132	0	4	2	4	32	0	0	2	11	24	29	10	0	1	0	10	1	0	1	26	5	1	7	12
133	5	6	11	0	2	36	0	2	13	9	23	8	8	0	0	9	0	29	0	0	11	4	11	0
134	21	7	0	9	10	7	39	2	0	0	4	44	2	16	0	8	22	0	31	0	10	5	4	0
135	0	9	4	33	6	1	1	2	1	3	0	6	22	2	1	0	1	3	6	26	6	16	9	2
136	16	0	0	0	32	1	0	0	0	9	0	7	17	14	3	10	0	0	0	4	2	26	27	0
137	65	44	0	0	0	34	2	0	4	0	8	7	10	38	5	10	3	3	0	0	3	1	14	15
138	19	66	64	0	0	0	9	1	8	7	0	21	24	18	34	5	43	5	0	0	1	0	2	6
139	0	12	47	76	0	5	13	0	20	0	12	1	36	21	16	11	0	39	9	0	0	0	32	3
140	9	1	0	47	63	0	8	20	0	16	4	16	2	26	3	2	0	11	33	16	29	0	0	24
141	0	14	0	0	38	45	17	24	20	12	0	0	20	2	18	0	0	1	21	47	12	38	11	0
142	3	0	4	12	0	0	8	12	0	31	2	0	6	32	12	12	0	0	0	5	49	0	33	9
143	1	6	12	0	3	0	26	0	45	0	46	34	2	6	32	11	27	5	6	3	0	19	11	9
144	16	3	0	13	6	17	11	22	0	28	4	27	25	3	0	34	33	50	39	49	21	0	10	14
145	4	14	3	10	0	0	28	31	16	8	1	1	0	11	9	5	35	26	10	0	39	21	0	5
146	0	0	16	0	0	7	0	0	6	6	5	2	0	12	9	15	4	13	1	0	0	14	36	0
147	0	4	1	0	0	0	0	0	16	14	0	4	3	7	25	1	11	0	0	3	5	0	8	61
148	0	11	25	1	2	0	0	10	0	0	0	1	7	9	2	22	2	0	0	15	41	9	3	14
149	38	0	17	33	0	0	0	1	0	2	9	11	0	6	2	0	0	2	0	1	21	32	13	0
150	0	13	0	10	25	0	7	2	1	0	3	4	32	0	0	0	1	20	9	0	14	0	39	6
151	16	4	2	0	12	47	0	10	1	1	2	0	19	37	5	0	0	3	2	4	2	11	0	17
152	1	3	0	0	0	30	16	3	0	0	0	0	0	31	36	22	0	8	0	0	4	8	19	0
153	6	0	5	0	0	24	33	7	0	0	0	0	0	57	12	20	9	14	1	1	1	17	44	
154	29	0	14	22	0	3	38	38	19	2	0	14	2	0	8	32	0	35	13	18	2	11	8	6
155	5	40	0	49	3	3	0	30	42	24	11	5	5	3	0	9	1	2	11	8	60	16	10	2
156	1	0	27	6	50	0	29	11	10	49	1	15	0	2	1	0	4	1	0	24	17	37	25	0
157	0	0	0	12	26	8	0	6	2	0	37	0	24	0	7	1	0	0	0	0	1	3	21	0
158	2	1	0	0	0	28	0	6	5	4	3	29	0	11	0	0	0	22	0	1	9	4	7	0
159	0	3	1	0	0	2	1	23	2	0	14	29	0	0	8	1	0	17	19	12	0	0	17	5
160	0	2	1	6	0	5	2	7	14	2	0	11	0	3	17	22	0	2	2	15	24	0	5	51
161	10	0	0	0	0	10	0	13	8	6	0	0	2	8	11	43	79	0	31	0	32	38	0	14
162	1	14	0	0	0	4	0	8	6	15	8	0	0	0	14	15	25	42	84	3	28	9	6	8
163	0	0	13	0	1	1	0	0	0	10	29	6	1	7	39	1	27	56	66	0	23	5	0	0
164	0	0	0	0	3	0	0	23	2	3	31	30	0	4	3	35	0							

Table 3: Mean occurrence (in %) of model parameters across the 852-day long out-of-sample test period. Columns represent the hours ($h = 1, \dots, 24$) for which the price predictions are made and rows represent the parameters (β_{174+j} and β_{343+k}) corresponding to the ‘future’ and past intraday prices (Y_{t-j} , $j = -8, \dots, 168$), and the seven weekday dummies (D_k , $k = 1, \dots, 7$) of the **LASSO**(λ_T) model, see Eqn. (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
-8	0	11	70	11	0	8	50	57	29	73	61	28	71	75	63	63	100	33	45	13	0	39	72	33
-7	0.4	0	0	53	0	0	0	0.7	76	25	26	41	23	0	31	17	0	82	20	51	5.8	0	1.8	21
-6	53	0	0	4	17	0	1	18	28	46	43	20	35	18	0	9	0	1	80	23	27	0	0	32
-5	46	53	0	0	48	0	0	14	7	53	14	52	26	27	17	5	6	0	0	99	35	6	0	0
-4	13	51	40	0	7	11	13	41	0	4	36	32	57	35	4	3	1	3	0	0	100	38	0	1
-3	46	20	59	29	48	48	2	0	38	0	21	72	2	61	50	9	0	0	0	0	10	100	21	36
-2	52	72	36	75	17	95	60	1	22	67	21	39	76	36	41	64	0	3	12	0	0	0	100	37
-1	97	53	70	84	28	47	97	100	50	60	94	69	76	93	60	75	63	26	37	42	64	40	0	80
0	100	100	63	100	100	100	96	100	100	99	100	100	100	100	100	100	100	100	100	100	100	100	96	0
1	37	0	0	25	9	6	0	0	100	72	0	4	50	36	8	31	34	44	55	7	63	5	12	100
2	36	9	0	0	0	0	1	0	0	88	2	0	0	0	0	0	0	2	13	0	6	0	0	0
3	0	0	0	0	0	0	3	15	0	0	4	44	0	1	0	0	0	0	0	2	8	6	0	28
4	12	0	4	0	5	0	97	61	56	21	75	100	100	46	33	63	76	29	11	12	35	63	89	89
5	44	28	0	65	0	0	8	84	86	45	24	23	55	93	71	36	65	71	47	80	65	6	41	47
6	17	50	60	13	96	0	50	15	51	65	2	0	26	33	35	76	32	83	95	32	84	54	10	40
7	0	10	44	21	30	100	61	39	25	46	25	5	0	24	10	26	43	68	23	67	10	58	43	7
8	2	0	1	15	2	0	92	77	34	18	12	3	7	3	22	0	9	22	63	19	31	2	17	0
9	0	4	0	2	0	0	0	39	71	40	0	0	3	8	1	22	19	15	39	38	0	56	20	24
10	7	0	6	0	0	0	1	0	41	55	11	0	0	0	8	18	28	0	13	29	2	9	13	6
11	0	0	0	2	0	2	41	0	6	36	86	4	0	0	0	14	8	26	1	16	14	3	6	0
12	11	6	11	0	0	0	4	18	35	3	6	48	78	1	0	0	33	5	34	0	9	1	0	0
13	0	1	10	0	0	7	13	66	34	0	25	18	63	0	0	3	0	52	9	37	6	0	20	0
14	3	5	26	8	0	0	14	13	72	16	0	16	38	50	0	0	12	0	43	7	37	19	29	7
15	14	4	8	9	3	0	25	5	11	45	0	1	30	43	50	1	0	19	3	17	8	38	0	41
16	24	16	0	7	26	0	0	25	31	38	33	0	0	33	44	51	18	5	37	8	3	3	38	12
17	0	24	0	0	20	17	0	9	33	35	35	0	29	12	34	75	30	10	17	4	0	3	50	0
18	40	0	0	0	0	0	50	9	1	0	26	12	14	40	1	48	17	23	38	32	24	0	0	7
19	0	29	0	0	0	0	0	0	0	0	3	3	19	17	8	26	12	39	41	4	9	0	0	10
20	0	0	0	45	0	0	12	6	7	0	1	16	0	10	20	1	0	3	4	23	1	0	0	2
21	0	0	3	0	77	0	0	23	7	4	0	0	15	0	1	12	5	0	0	7	10	0	0	0
22	0	0	9	7	1	55	14	0	28	0	4	0	0	4	0	0	0	16	0	0	0	4	3	0
23	0	0	3	38	12	11	42	0	15	6	0	11	0	0	0	0	0	0	30	0	4	2	0	6
24	3	0	0	19	15	4	2	26	16	49	0	2	5	0	0	0	0	10	2	17	30	37	34	0
25	1	0	8	0	16	0	1	10	6	10	35	4	1	0	1	0	0	0	10	0	26	29	11	33
26	23	1	1	0	0	0	2	2	7	2	0	18	4	6	0	21	0	0	0	0	0	0	10	0
27	0	24	19	0	0	0	0	2	7	3	3	0	9	1	5	0	13	0	2	15	0	0	3	0
28	4	3	2	6	0	0	0	0	0	0	0	0	1	0	0	7	1	0	28	25	17	12	23	1
29	3	2	0	7	0	0	0	0	0	0	0	0	0	0	3	1	3	5	24	23	1	0	0	0
30	0	2	0	2	11	10	13	0	0	0	0	0	0	0	0	0	0	2	0	18	14	1	2	0
31	33	0	14	0	0	25	0	0	0	0	0	0	2	0	2	24	0	0	0	0	31	6	0	0
32	0	37	0	13	2	7	0	0	0	0	0	0	1	4	0	0	30	20	0	0	32	0	4	0
33	50	22	23	1	11	5	1	2	1	3	0	36	12	0	0	3	42	2	0	0	0	2	0	0
34	8	33	13	18	0	15	6	0	64	0	8	0	41	0	2	0	10	3	7	0	0	0	0	14
35	17	2	45	17	0	0	0	10	0	62	8	16	1	39	0	0	0	43	1	6	0	0	22	1
36	0	4	6	24	1	1	0	8	6	6	19	10	11	0	17	0	0	0	17	13	4	0	1	41
37	0	0	1	0	13	0	14	0	1	5	0	11	9	2	0	0	0	0	6	4	1	0	0	0
38	0	0	1	0	1	0	2	2	2	1	2	0	0	5	0	0	0	17	0	0	17	0	0	1
39	10	0	0	0	0	5	2	7	1	3	4	21	0	0	1	0	3	0	6	0	5	18	7	10
40	0	4	0	0	0	0	5	2	0	0	11	3	25	0	0	1	0	4	24	0	0	1	45	12
41	3	0	6	1	0	0	3	17	0	0	9	6	0	27	0	0	0	0	2	8	9	1	7	0
42	13	0	4	3	15	1	3	3	1	0	8	16	0	1	9	0	0	0	3	0	21	4	0	0
43	10	14	4	6	4	6	5	4	2	1	0	23	3	3	18	3	0	0	4	1	0	4	5	12
44	0	34	3	0	1	0	10	0	4	0	0	0	8	0	1	10	0	0	0	2	0	10	5	0
45	0	24	44	1	0	0	13	0	2	2	0	0	4	2	0	0	6	0	4	2	2	0	0	22
46	0	5	0	44	4	18	0	0	8	16	1	0	4	2	0	0	0	0	0	8	1	1	4	0
47	0	7	15	3	35	10	19	2	0	0	7	7	7	4	0	0	0	21	39	10	1	24	19	8
48	6	2	0	4	11	20	25	22	8	0	5	0	12	4	0	0	1	33	40	88	47	18	57	19
49	25	20	0	8	24	0	0	37	5	0	0	3	0	8	0	0	0	34	27	37	9	15	21	0
50	11	0	10	0	13	20	0	0	37	0	0	0	4	19	1	0	0	0	9	0	15	17	0	0
51	1	0	0	9	1	10	0	0	0	33	12	0	1	23	36	0	0	0	1	0	0	0	0	0
52	4	0	0	0	19	1	1	0	4	0	32	0	0	13	47	46	5	0	14	0	2	5	2	0
53	1	0	6	0	0	0	0	0	1	0	0	0	27	0	17	17	27	38	6	1	0	0	5	0
54	0	0	3	9	9	0	0	0	3	0	0	0	26	13	52	1	18	17	4	0	0	0	8	0
55	29	0	1	1	4	0	3	0	0	1	0	0	0	26	21	52	4	31	7	1	0	2	3	43
56	28	2	0	3	0	0	0	10	0	6	3	3	0	8	21	49	0	1	1	0	0	0	0	0
57	1	38	0	0	1	9	5	0	3	0	4	0	5	41	1	19	14	0	0	1	11	2	0	5
58	5	0	2	0	0	1	23	0	0	0	0	2	2	17	39	0	27	0	0	0	7	13	3	0
59	1	0	0	0	0	5	0	7	4	2	0	2	1	14	15	16	0	4	19	2	10	0	7	5
60	9	0	0	0	0	9	2	0	7	0	0	3	18	0	0	3	0	0	4	21	2	4	0	6
61	0	0	0	0	0	0	0	6	0	0	0	3	1	11	0	1	0	0	0	12	3	0	1	0
62	0	0	0	0	0	4	0	4	0	0	0	11	1	2	6	0	2	0	5	0	15	9	0	0
63	0	1	1	0	0	2	17	0	8	0	0	8	12	0	6	0	0	22	2	0	0	5	1	9
64	0	0	11	9	0	0	0	16	0	3	0	1												

Table 3: *Cont.*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
81	1	0	0	31	30	64	9	2	0	0	0	0	0	4	0	14	12	2	17	0	4	0	0	2
82	0	0	0	5	21	12	0	8	2	0	2	0	0	0	0	0	0	2	2	2	0	0	0	25
83	19	0	0	0	0	20	3	25	2	0	1	0	0	3	0	0	0	0	4	1	9	0	0	0
84	0	15	0	0	0	3	0	11	22	0	1	18	0	1	6	0	0	0	9	1	6	10	0	0
85	2	0	9	0	0	0	0	0	0	10	0	3	1	0	0	3	0	1	2	10	6	0	0	3
86	4	5	0	9	6	14	21	0	0	19	6	0	17	0	0	5	5	15	2	7	0	1	5	5
87	0	0	0	0	3	10	12	17	0	0	20	9	7	16	1	0	0	5	6	13	2	0	0	4
88	17	5	0	0	0	2	47	4	11	0	0	9	6	10	13	1	0	4	8	3	3	0	1	0
89	0	52	1	0	0	0	9	2	0	0	3	0	7	4	2	1	6	21	0	1	4	0	0	0
90	0	12	31	0	0	0	0	0	0	0	1	16	3	0	1	0	4	13	3	16	0	1	15	0
91	0	0	43	28	0	0	0	0	0	0	5	3	15	9	0	0	0	0	41	2	3	0	5	0
92	0	0	1	33	1	0	10	0	0	2	2	48	16	6	11	1	0	0	2	35	0	0	0	0
93	0.7	0	0	20	25	0	0	0	2.1	9.5	15	5.2	36	2.8	0	4.3	0	0	0.7	0	14	0	0	0
94	14	0	1	0	5	0	0	2	0	4	20	29	0	1	4	3	0	0	9	1	0	0	0	0
95	0	12	0	1	0	7	3	6	19	3	36	4	10	0	1	29	33	6	0	6	15	25	0	28
96	19	0	8	0	7	0	7	2	1	20	0	23	4	0	0	11	41	74	21	42	25	23	40	9
97	2	1	0	3	6	4	0	20	13	1	3	0	0	6	0	0	15	18	55	10	2	0	7	26
98	0	0	0	0	2	0	0	4	5	35	0	0	7	26	2	0	0	0	2	0	3	1	1	0
99	0	17	0	0	0	2	0	0	0	21	9	1	0	12	12	44	0	0	5	3	0	3	13	3
100	12	2	1	0	0	0	0	11	29	13	0	4	19	31	18	0	0	6	0	14	29	6	14	0
101	15	14	6	31	0	4	1	0	0	3	15	5	4	8	13	5	0	0	0	7	36	6	3	0
102	1	1	9	1	12	2	0	0	0	0	2	21	6	16	1	40	0	0	2	0	0	0	4	0
103	0	0	7	0	1	1	0	0	0	0	0	7	12	4	26	0	34	0	11	0	0	0	0	0
104	0	0	0	15	0	1	0	2	2	0	1	3	1	1	8	66	0	0	4	23	0	0	0	0
105	1	0	16	25	36	0	0	4	0	4	2	6	0	0	0	2	40	2	11	4	21	0	2	0
106	13	0	0	4	29	19	18	0	10	0	0	20	5	0	0	5	10	47	1	0	0	3	5	6
107	0	2	1	0	6	2	10	15	0	26	0	14	2	1	0	1	0	15	10	1	3	0	1	1
108	3	3	2	1	0	0	2	10	18	3	11	0	19	0	0	0	0	0	0	17	0	10	0	4
109	0	0	3	2	0	0	1	0	16	10	12	7	14	5	0	2	0	0	0	24	0	9	0	0
110	1	0	0	8	2	1	0	4	1	15	6	13	8	18	0	2	0	0	1	0	41	8	7	4
111	0	0	8	4	0	0	0	6	0	0	3	0	0	0	17	8	0	0	2	3	0	4	8	0
112	0	5	0	0	0	7	0	8	0	0	0	0	0	0	0	29	9	0	2	2	5	0	0	0
113	1	17	5	0	6	0	0	0	0	0	3	0	0	0	0	4	3	0	0	2	0	1	0	2
114	4	1	15	4	0	0	8	0	2	0	0	2	11	6	0	0	1	0	0	0	6	0	0	0
115	0	11	19	31	5	0	17	0	24	1	0	1	11	12	5	3	4	0	2	0	8	2	0	1
116	0	0	0	3	25	2	0	6	19	14	1	0	8	12	36	2	5	0	1	0	0	0	0	1
117	0	0	0	14	6	0	4	16	0	4	18	0	19	0	5	9	30	0	8	0	14	0	0	0
118	0	0	0	0	0	23	12	15	8	0	0	0	13	0	14	12	4	10	9	5	1	2	0	7
119	0	0	9	0	0	0	24	2	10	0	1	0	16	10	0	7	0	2	1	0	3	1	19	0
120	0	0	0	0	11	7	0	4	26	4	9	2	0	5	4	1	5	18	1	15	18	7	13	42
121	0	0	0	0	17	13	0	0	6	7	0	7	2	4	0	0	0	24	4	2	13	2	38	39
122	8	0	0	0	0	14	0	25	0	13	9	8	5	0	3	4	0	6	2	0	11	14	0	14
123	0	0	4	0	0	0	0	0	34	0	33	1	1	0	0	1	4	0	4	0	3	12	31	0
124	3	6	2	17	0	0	0	0	11	7	9	31	1	1	0	2	2	6	7	0	8	3	45	30
125	31	0	5	0	6	0	0	0	0	41	0	1	12	0	1	0	9	6	7	0	1	53	3	36
126	3	4	0	0	0	3	0	4	8	2	40	0	1	1	0	0	5	6	8	0	5	0	0	0
127	10	5	0	3	0	0	1	0	1	12	4	38	0	0	0	3	0	20	1	2	2	0	0	0
128	0	4	16	0	4	0	0	1	0	0	0	0	5	0	0	21	0	0	22	0	0	3	0	0
129	0	0	9	0	0	0	45	3	0	0	1	0	3	0	0	0	8	0	2	11	0	0	0	0
130	0	3	0	2	8	0	0	23	20	0	3	0	0	0	0	5	2	0	26	2	6	12	0	31
131	0	2	3	0	0	2	0	16	8	0	2	0	0	4	0	0	0	3	10	12	0	4	16	0
132	0	1	6	8	0	0	12	0	1	9	0	4	0	0	3	0	0	0	3	14	4	0	2	46
133	0	0	17	5	0	4	6	0	1	1	0	0	14	0	0	0	0	0	0	30	3	1	4	0
134	0	0	0	23	0	5	5	0	0	10	1	1	9	14	3	0	12	0	3	0	4	0	0	0
135	12	2	0	1	14	0	21	0	0	0	8	0	10	0	0	27	0	1	6	39	0	0	7	0
136	3	12	1	12	0	15	0	21	0	0	0	0	0	0	0	33	10	6	2	13	0	18	37	0
137	19	27	2	15	0	11	6	1	23	7	0	0	0	22	0	6	15	58	2	1	3	36	0	8
138	12	13	16	0	25	4	2	2	0	20	0	10	0	0	5	0	9	12	14	0	12	0	29	0
139	0	15	26	31	1	15	0	1	0	0	16	3	5	6	2	5	0	7	8	0	0	0	3	1
140	0	0	0	6	30	0	4	0	0	0	0	15	1	0	22	4	1	0	3	20	6	0	2	2
141	3	0	3	0	14	25	26	6	11	0	3	0	3	0	1	27	1	0	6	0	11	0	0	1
142	0	0	0	1	1	0	23	20	3	0	0	2	4	0	0	9	4	0	0	0	3	0	0	0
143	0	0	4	3	0	1	0	0	0	1	0	0	0	2	1	19	6	1	2	16	4	0	0	0
144	4	1	4	2	11	8	0	0	0	2	2	1	2	3	13	0	16	0	48	85	73	35	25	10
145	0	0	6	3	2	4	13	6	0	0	0	0	1	2	4	0	0	0	16	7	18	41	30	5
146	1	0	2	0	0	17	0	0	0	2	0	4	0	11	10	0	0	0	0	1	0	4	0	6
147	0	4	0	0	0	0	0	0	0	2	22	4	1	12	16	32	0	0	2	2	12	0	0	2
148	0	0	4	0	0	0	0	0	0	0	10	24	17	0	20	17	13	2	1	0	28	3	2	28
149	50	6	1	6	0	0	33	13	0	0	6	27	12	0	14	0	0	0	15	0	0	0	8	0
150	1	23	8	11	37	0	1	0	41	27	0	0	1	6	0	0	0	2	0	0	1	0	0	1
151	0	0	9	4	6	2	10	0	0	32	30	0	0	4	4	0	3	0	21	0	1	3	0	0
152	0	0	0	0	8	0	0	0	0	0	1	7	0	0	17	24	0	0	24	13	0	2	8	4
153	0	0	4	3	2	1	0	0	3	0	0	1	2	0	0	0	27	0	1	31	10	5	0	1
154	10	0	3	7	0	14	5	0	20	0	0	0	4	0	0	0	0	18	8	11	27	23	6	8
155	0	32	0	0	1	14	21	11	9	17	0	1	6	0	0	0	0	4	23	10	10	37	5	0

5. Conclusions

Using a unique set of electricity prices from the German EPEX market, we have addressed 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 forward 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.

To this end, we have considered 14 models, including three naive benchmarks, a parsimonious autoregressive structure inspired by the well-performing *expert_{Dow,nl}* model of Ziel and Weron (2018) and a LASSO-estimated model with nearly 350 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 all 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 price corresponding to the same or nearby hours. However, the most recent intraday price by itself was a rather poor predictor. Also the intraday and – to a lesser extent – the day-ahead prices for late evening hours could 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.

Acknowledgements

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) 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).
- Barnes, A. K., Balda, J. C., 2013. Sizing and economic assessment of energy storage with real-time pricing and ancillary services. *PEDG 2013 Conference Proceedings*, DOI: 10.1109/PEDG.2013.6785651.
- Contreras, J., Espínola, R., Nogales, F., Conejo, A., 2003. ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems* 18 (3), 1014–1020.
- Diebold, F. X., Mariano, R. S., 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253–263.
- EPEX, 2018. Annual Report 2017. [Http://www.epexspot.com/en/extras/download-center](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 and 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.
- 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., 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).
- Nowotarski, J., Weron, R., 2018. Recent advances in electricity price forecasting: A review of probabilistic forecasting. *Renewable and 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 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 intra-day market. In: Zamasz, K. (Ed.), *Capacity Market in Contemporary Economic Policy*. Difin, pp. 124–138.
- 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. *IEEE Conference Proceedings - EEM17*, DOI: 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.

HSC Research Report Series 2018

For a complete list please visit <http://ideas.repec.org/s/wuu/wpaper.html>

- 01 *An empirical analysis of green energy adoption among residential consumers in Poland* by Anna Kowalska-Pyzalska
- 02 *Efficient forecasting of electricity spot prices with expert and LASSO models* by Bartosz Uniejewski and Rafał Weron
- 03 *A note on averaging day-ahead electricity price forecasts across calibration windows* by Katarzyna Hubicka, Grzegorz Marcjasz and Rafał Weron
- 04 *Household willingness to pay for green electricity in Poland* by Anna Kowalska-Pyzalska and David Ramsey
- 05 *Probabilistic electricity price forecasting with NARX networks: Combine point or probabilistic forecasts?* by Grzegorz Marcjasz, Bartosz Uniejewski and Rafał Weron
- 06 *Selection of calibration windows for day-ahead electricity price forecasting* by Grzegorz Marcjasz, Tomasz Serafin and Rafał Weron
- 07 *Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO* by Bartosz Uniejewski, Grzegorz Marcjasz and Rafał Weron