FORECASTING THE RESIDUAL DEMAND FUNCTION IN ELECTRICITY AUCTIONS

Matteo Pelagatti

Dipartimento di Statistica
Università degli Studi di Milano-Bicocca
Via Bicocca degli Arcimboldi 8, Milano, Italy
(e-mail: matteo.pelagatti@unimib.it)

ABSTRACT. In the fast growing literature that addresses the problem of the optimal bidding behaviour of power generation companies that sell energy in electricity auctions it is always assumed that every firm knows the aggregate supply function of its competitors. Since this information is generally not available, real data have to be substituted with forecasts. In this paper we propose two alternative approaches to the problem and apply them to the hourly prediction of the residual demand function of the main Italian generation company.

1 INTRODUCTION

The last twenty years have witnessed in most European and many non European countries a radical reorganization of the electricity supply industry. Government owned monopolies have been replaced by regulated (generally pool) competitive markets, where the match between demand and supply takes place in hourly (in some cases semi-hourly) auctions. The auction mechanism is generally based on a uniform price rule, i.e. once the equilibrium price is determined, all the dispatched producers receive the same price per MWh.

The issue of determining the profit-maximising behaviour of a power company bidding in electricity auctions has been addressed by economists both from the normative (profit optimisation) and positive (market equilibrium) point of views (cf. Wolak (2003), Hortaçsu and Puller (2008), Bosco *et al.* (2010)). If we assume that each firm wishes to maximise its profit in each auction independently from the other auctions (as customary in the literature), then we can summarise the optimisation problem as follows.

Suppose that D is the (inelastic) demand for electricity, $S_{-i}(p)$ is the aggregate supply function of firm i's competitors for any given price p, $C_i(q)$ is the production cost function of firm i for any given quantity of energy q, then for those values of the residual demand $D - S_{-i}(p)$ that the production capacity of firm i can fulfill, the profit function of firm i is given by

$$\pi_i(p) = p \cdot \big(D - S_{-i}(p)\big) - C_i\big(D - S_{-i}(p)\big).$$

This profit function can be extended to include financial contracts as in Hortaçsu and Puller (2008) or vertical integration (i.e. the situation in which the producer is also a retailer and plays in both sides of the auction) as in Bosco *et al.* (2010). Assuming the differentiability of S_{-i} and C, and the concavity of π_i, first order conditions indicate that firm i maximises its profit when it offers the quantity $D - S_{-i}(p^*)$ at price p^* satisfying

$$p^* = C_i'\big(D - S_{-i}(p^*)\big) + \frac{D - S_{-i}(p^*)}{S_{-i}'(p^*)}. \tag{1}$$

Now, the quantity D and the supply function S_{-i} are generally unknown, but while D can be predicted using standard time series techniques, the prediction of the function S_{-i} is more involved. In this paper we propose two techniques for forecasting supply functions based on principal component analysis and reduced rank regression. The techniques are applied to the prediction of the hourly supply functions of the competitors of Enel, the main Italian generation company, as observed in two years of Italian electricity auctions .

2 Auction rules and data

According to the rules of the Italian electricity market, each production unit can submit up to four "packages" of price-quantity pairs. Each pair indicates how many MWh the plant owner would sell for how many Euros, were they dispatched. All the submitted pairs are sorted by price and the corresponding quantities are cumulated. When the cumulated quantity matches the demanded quantity, the system marginal price (SMP) is determined and all the plants offering energy up to that price are dispatched.

Each record of the Italian auction result database[1] contains the price-quantity pair, the name of the offering production unit and the name of the owner of that unit. This allows the construction of the supply function of any bidding firm.

From the above reasoning it is clear that real supply schedules are step functions and, thus, the optimal bidding theory discussed in the previous section is not directly applicable. This issue is generally dealt with by approximating the step functions with differentiable functions obtained though kernel smoothing. Smoothing is also necessary for regularising functions before applying canonical correlation techniques such as reduced rank regression (cf. Sec. 11.5 of Ramsay and Silverman (2005)). Since supply functions are nondecreasing in price, we use the kernel $S(p) = \sum_{k=1}^{K} q_k \Phi\left(\frac{p-p_k}{h}\right)$, where Φ is the standard normal cumulative probability function, h is the bandwidth parameter and (q_k, p_k) are the observed quantity-price pairs. Notice that the total number of offers K may change in each auction. The derivative of the smoothed function needed in equation (1) is given by $S'(p) = \sum_{k=1}^{K} q_k \frac{1}{h} \phi\left(\frac{p-p_k}{h}\right)$, with ϕ standard normal density. The left panel of Figure 1 depicts the actual supply function of Enel's competitors on 3 March 2008 at 10am and the kernel approximation thereof ($h = 3$).

3 Supply functions prediction

Both prediction techniques we propose in this paper entail some common steps.

The first step consists in sampling the (smoothed) function on a grid of abscissa points. This is necessary as the price set on which the function can be evaluated changes in every

[1] It can be downloaded (one file per day) from the market operator web site `www.mercatoelettrico.org`

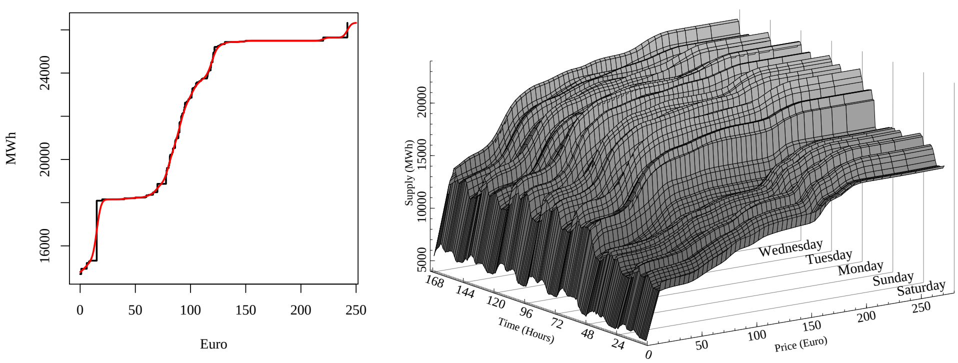

Figure 1. Left: supply function of Enel's competitors on 3.12.2008 at 10am and kernel approximation. Right: the same function sampled at 50-iles over one week.

auction. Since the function can be approximated more accurately where bid pairs are more dense, we sample more frequently in these intervals by using quantiles. In particular, we used 50-iles of unique prices submitted over the entire sample (2007-2008). The forty-nine 50-iles where supplemented with the minimum (0) and the theoretical maximum (500) due to the price capping rule of the Italian market, obtaining 51 time series of ordinate points. The right panel of Figure 1 shows one week of Enel's competitors aggregate supply functions sampled at 50-iles.

The second common step consists in transforming the original ordinate points in a way such that the two features of positivity and non-decreasing monotonicity of the original functions are guaranteed also in their predictions. If we denote with $\{p_0, p_1, \ldots, p_{50}\}$ the points in the price grid and with $S_t(p_i)$ the smoothed supply function at time t for price p_i, then the time series we work on are obtained as

$$q_{i,t} := \begin{cases} \log S_t(p_i), & \text{for } i = 0; \\ \log\big(S_t(p_i) - S_t(p_{i-1}) + c\big), & \text{for } i = 1, \ldots, 50, \end{cases}$$

where c is a small constant. If we assume that the predictor of $q_{i,t}$, say $\hat{q}_{i,t}$, is unbiased, and the prediction error is approximately normal with standard error $s_{i,t}$, unbiased forecasts of the original function can be recovered as

$$\hat{S}_t(p_i) = \begin{cases} \exp\big(\hat{q}_{i,t} + s_{i,t}^2/2\big), & \text{for } i = 0; \\ \exp\big(\hat{q}_{i,t} + s_{i,t}^2/2\big) + \hat{S}_t(p_{i-1}) - c, & \text{for } i = 1 \ldots, 50. \end{cases}$$

Now, since we expect the 51 time series to share information, it is natural to seek some form of dimension reduction. The two alternative algorithms we propose are based on principal component (PC) analysis and reduced rank regression (RRR). We base the choice of dimension reduction on one month of k-step-ahead out-of-sample predictions, where $k = 1$ for one-hour-ahead predictions and $k = 24$ for one-day-ahead predictions. In particular, the model fit is carried out using the hourly observations of the years 2007-2008, while the dimension assessment is based on Jan-2009 prediction mean square errors (MSE).

In describing the two algorithms we collect the 51 transformed time series in vector $\mathbf{q}_t$ and the original supply function ordinate-points in vector $\mathbf{S_t}$. The predictions will be based on lagged responses and deterministic regressors.

Algorithm 1 (Principal components based) *For* $r = \{51, 50, \dots, 1\}$ *iterate through*

1. *Take the first r PCs of* $\mathbf{q}_t$ *(supply function log increments) based on its in-sample covariance matrix, and name the scores* $\mathbf{y}_t$.
2. *Regress each score* $y_{i,t}$ *on its lags* $\mathbf{x}_{i,t}$ *and deterministic regressors* $\mathbf{z}_t$ *and compute predictions* $\hat{\mathbf{y}}_t$.
3. *Regress the vector* $\mathbf{q}_t$ *on the predicted scores* $\hat{\mathbf{y}}_t$ *(and a constant).*
4. *Compute the out-of-sample predictions of the supply function* $\mathbf{S}_t$ *and the relative MSE.*

Pick the rank r that minimize the out-of-sample MSE.

In the PC approach, the time series are first reduced in number by taking the best linear approximation to the original data and then these are be predicted using standard time series models. The main advantage of this approach is the freedom left to the researcher to choose the model to predict the PC scores. The drawback is that the rank reduction is not optimised for prediction.

The second approach is based on reduced rank regression, but since this technique is less popular than principal component analysis, we briefly survey the main points. Consider the model

$$\mathbf{y}_t = \Pi \mathbf{x}_t + \Gamma \mathbf{z}_t + \varepsilon_t$$

where $\mathbf{x}_t$ and $\mathbf{z}_t$ are regressors, Γ is a full-rank coefficient matrix, Π is a reduced-rank coefficient matrix and ε_t is a sequence of zero-mean random errors uncorrelated with the regressors. The fact that Π is reduced-rank means that few linear combinations of the regressors $\mathbf{x}_t$ are sufficient to take account of all the variability of $\mathbf{y}_t$ due to $\mathbf{x}_t$. Now, suppose that Π is $n \times m$ with rank $r < \min(m, n)$, then Π can be factorised as $\Pi = \mathbf{AB}'$, with $\mathbf{A}$ $n \times r$ and $\mathbf{B}$ $m \times r$ matrices. The matrices $\mathbf{A}$ and $\mathbf{B}$ are not uniquely identified, but if one restricts the r (column) vectors forming $\mathbf{B}$ to be orthonormal, then a least squares solution for $\mathbf{B}$ is found by solving the following eigenvalue problem:

$$\mathbf{S}_{xx|z}\mathbf{V}\Lambda = \mathbf{S}_{xy|z}\mathbf{S}_{yy|x}^{-1}\mathbf{S}_{yx|z}\mathbf{V},$$

where $\mathbf{S}_{ab|c}$ indicates the partial product-moment matrix of $\mathbf{a}$ and $\mathbf{b}$ given $\mathbf{c}$, $\mathbf{V}$ is an orthonormal matrix and Λ is a diagonal matrix. The first r columns of $\mathbf{V}$ provide least square estimates of $\mathbf{B}$. Least squares estimates of $\mathbf{A}$ and Γ are found by regressing $\mathbf{y}_t$ simultaneously on $\mathbf{w}_t := \mathbf{B}'\mathbf{x}_t$ and $\mathbf{z}_t$. For details on RRR refer to the excellent monograph by Reinsel and Velu (1998).

Algorithm 2 (Reduce rank regression based) *For* $r = \{51, 50, \dots, 1\}$ *iterate through*

1. *Regress the vector* $\mathbf{y}_t = \mathbf{q}_t$ *on its lags* $\mathbf{x}_t$ *imposing rank r on the matrix* Π *and on the deterministic regressors* $\mathbf{z}_t$ *without restricting the rank.*

2. *Compute the out-of-sample predictions of the supply function* $\mathbf{S}_t$ *and the relative MSE.*

Pick the rank r that minimize the out-of-sample MSE.

The main advantage of the RRR-based algorithm is that the rank reduction is optimised for prediction. The drawback is that only (vector) autoregressive models with exogenous variables are allowed.

4 Application to the Italian electricity auctions

We apply the two algorithms to the hourly Italian electricity auction results for the years 2007-2008 (17544 auctions); Jan-2009 (744 auctions) is used for choosing the rank r as explained in the previous section.

We build models for predicting one-hour-ahead and models for forecasting one-day-ahead. As for the deterministic regressors ($\mathbf{z}_t$) we implement the following three increasing set of variables.

1. Linear trend, $\cos(\omega_j t)$, $\sin(\omega_j t)$, with $\omega_j = 2\pi j/(24 \cdot 365)$ and $j = 1, \ldots, 20$.
2. Regressors at point 1. plus dummies for Saturday, Sunday and Monday.
3. Regressors at point 2. plus $\cos(\lambda_i t)$, $\sin(\lambda_i t)$, with $\lambda_i = 2\pi i/24$ and $i = 1, \ldots, 6$.

Both autoregressive models and error correction mechanisms are explored. In particular, we regress

Level 1-step: $\mathbf{y}_t$ on $\mathbf{y}_{t-1}, \mathbf{y}_{t-24}, \mathbf{y}_{t-168}, \mathbf{z}_t$;
Diff 1-step: $\Delta\mathbf{y}_t$ on $\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-1}, \Delta\mathbf{y}_{t-24}, \Delta\mathbf{y}_{t-168}, \Delta\mathbf{z}_t$;
Level 24-step: $\mathbf{y}_t$ on $\mathbf{y}_{t-24}, \mathbf{y}_{t-168}, \mathbf{z}_t$;
Diff 24-step: $\Delta_{24}\mathbf{y}_t$ on $\mathbf{y}_{t-24}, \Delta_{24}\mathbf{y}_{t-24}, \Delta_{24}\mathbf{y}_{t-168}, \Delta_{24}\mathbf{z}_t$.

The chosen rank and the actual root MSE (RMSE) are summarized in Table 1. Three

Table 1. Out-of-sample root mean square error for the two algorithms and 12 models.

	one-hour-ahead						one-day-ahead					
	Reg. 1.		Reg. 2.		Reg. 3.		Reg. 1.		Reg. 2.		Reg. 3.	
	rank	RMSE	rank	RMSE	rank	RMSE	rank	RMSE	rank	RMSE	rank	RMSE
RRR-Level	50	76.5	50	76.5	50	**67.1**	44	216.1	44	216.1	44	215.5
RRR-Diff	51	*73.7*	51	*73.7*	51	73.7	51	198.7	51	198.7	51	198.7
PC-level	50	80.9	50	80.9	50	72.7	41	*191.1*	41	*191.1*	41	**188.8**
PC-Diff	37	*73.7*	47	*73.7*	37	73.7	51	198.8	51	198.8	51	198.8

remarks appear evident from these figures: i) the optimal rank of both PC and RRR models is very close to the full rank (51), indicating that almost all the information that time series carry is relevant for forecasting; ii) there is no clear indication about the choice of the algorithm: the best algorithm for hour-ahead predictions is RRR, while that for day-ahead predictions is PC; iii) many deterministic regressione are better then few.

Figure 2 depicts the out-of-sample RMSE as a function of time (first panel) and of price (second panel). It appears clear that the precision of the predictions vary over time and price. In particular, the first half of Jan-2009 seems to be harder to predict then the following part of that month. Indeed, those days are characterised by holidays and school vacations that we have not explicitly modelled. As for the precision of the prediction at different points of the supply function, the hardest quantities to predict are those in correspondence of the price intervals $[0,20]$ and $[220,250]$. Most observed SMPs are in the range $[50,100]$, and so this interval is probably the most interesting for a bidding firm. The RMSE in that interval is not particularly large, even if it peaks in a neighbourhood of 90 Euros.

The proposed prediction algorithms, maybe supplied with other relevant regressors such as temperature and whether forecasts, seem to represent a valuable tool for helping generation companies to design their bidding strategies in a more profitable way.

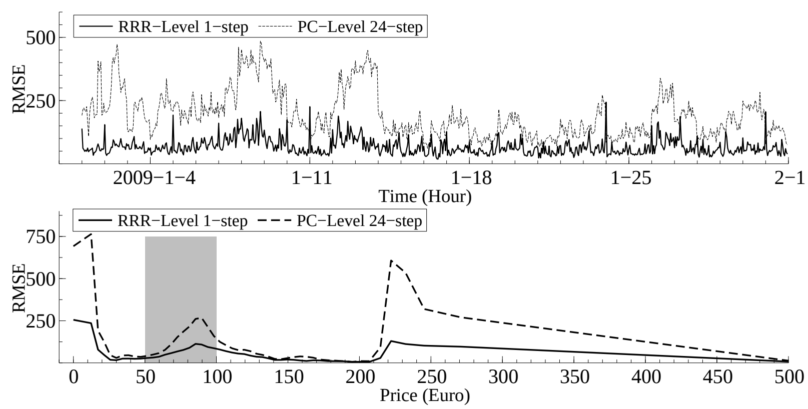

Figure 2. Root mean square error of prediction as functions of time and price.

REFERENCES

BOSCO, B., PARISIO, L., PELAGATTI, M. (2010): Estimating Marginal Costs and Market Power in the Italian Electricity Auctions, Statistics Department Working Paper n.20100201, Università degli Studi di Milano-Bicocca.

HORTAÇSU, A., PULLER, S.L. (2008): Understanding strategic bidding in multi-unit auctions: a case study of the Texas electricity market, *Rand Journal of Economics*, *39(1)*, 86–114.

RAMSAY, J.O., SILVERMAN, B.W. (2005): *Functional data analysis*. Second edition. Springer Series in Statistics. Springer, New York.

REINSEL, G.C., VELU, R.P. (1998): *Multivariate Reduced Rank Regression*. Lecture Notes in Statistics. Springer, New York.

WOLAK, F.A. (2003): Identification and estimation of cost functions using observed bid data: an application to electricity markets. In: M. Dewatripont, L.P. Hansen and S. J. Turnovsky (Eds.), *Advances in economics and econometrics: theory and applications, Eighth World Congress, Volume II*. Cambridge University Press, New York, 133–169.