Working Paper Series
Department of Economics
University of Verona

Forecasting Italian Electricity Zonal Prices with Exogenous Variables

Angelica Gianfreda, Luigi Grossi

WP Number: 1 January 2011

ISSN: 2036-2919 (paper), 2036-4679 (online)

Forecasting Italian Electricity Zonal Prices with Exogenous Variables

Angelica Gianfreda and Luigi Grossi[a]

[a]*Department of Economics, University of Verona, Italy*

Abstract

In the last few years we have observed deregulation in electricity markets and an increasing interest of price dynamics has been developed especially to consider all stylized facts shown by spot prices. Only few papers have considered the Italian Electricity Spot market since it has been deregulated recently. Therefore, this contribution is an investigation with emphasis on price dynamics accounting for technologies, market concentration and congestions. We aim to understand how technologies, concentration and congestions affect the zonal prices since these ones combine to bring about the single national price (*prezzo unico d'acquisto*, PUN). Hence, understanding its features is important for drawing policy indications referred to production planning and selection of generation sources, pricing and risk–hedging problems, monitoring of market power positions and finally to motivate investment strategies in new power plants and grid interconnections. Implementing Reg–ARFIMA–GARCH models, we assess the forecasting performance of selected models showing that they perform better when these factors are considered.

Key words: Electricity prices, Production technologies, Market power (HHI, RSI), Congestions, Fractional Integration, Forecasting

1. Introduction

Several empirical features of electricity prices observed at daily frequency have been widely discussed: mean–reversion, seasonality, time varying and clustered volatility, inverse leverage effect and extreme values called spikes or jumps, see for instance Escribano et.al. (2002), Knittel and Roberts (2005), Koopman et al. (2007) and Gianfreda and Bunn (2010) among others. While seasonality and clustered volatility are well–known features, the remaining stylized facts require to be better explained. Mean-reversion is the tendency that prices show tending to a long-run mean level. The inverse leverage effect, discovered by Knittel and Roberts (2005), is the inverse reaction to shocks: electricity price volatility tends to rise in presence of positive shocks more than in presence of negative ones. Extreme values or spikes are results of abnormally large variations in price caused by weather conditions, outages or transmission failures.

A stylized fact that has been fairly neglected so far is the long memory of the process generating electricity prices. When unit root tests are applied usually the presence of unit root is rejected. On the other hand, stationarity tests provide evidence of non stationarity. Moreover, when the empirical correlation function is estimated and visualized through a correlogram, a long memory pattern can be observed because autocorrelation tends to decrease very slowly as the lag increases. These results combined lead to the conclusion that the analyzed time series could be generated by a fractionally integrated process. This feature is explored for the first time in the present paper using data of the Italian market. Previous papers about the Italian Power Exchange (Gianfreda , 2010; Petrella and Sapio, 2009 and Bosco et.al., 2007) have completely discarded this peculiarity and have focused on prices and returns.

Beside fractional integration, we considered also the conditional heteroscedasticity of residuals which has been captured by GARCH models. The final correlation structure of electricity prices has been then captured by Reg–ARFIMA–GARCH models. An interesting results is that the Italian market does not show the presence of leverage effect, nor direct or indirect, as explored by Gianfreda (2010).

The procedure suggested in this paper is aimed to correctly identify the appropriate stochastic generating process for electricity prices which is important for several reasons. First of all, the price dynamics can be used to understand the deregulation process, verify the competition in this electricity market and give indications on spot and forward price definitions (Giulietti et al., 2010). Secondly, a good model identification leads to proper managing of network congestions for needs of continuous real time balancing. Thirdly, modeling is important for forecasting, for trading, for generation planning and plants availability, for risk management and hedging purposes in such market given the recent launch of the Forward Electricity Market (MTE)[1] on 3 November 2008.

Another original contribution of this work is the use of exogenous variables to explain the electricity price dynamics. Exploiting the massive information provided by the managing authority of the Italian electricity market (GME: Gestore del Mercato Elettrico), we have analyzed the effect of technologies, market power and network congestions on prices. It is well-known that electricity prices depend on prices of generation sources employed, however there is no evidence on the degree and sign of these influences. Moreover we control for exercise of market power from the generation side. Therefore we find answers on how generation sources, market power and congestions interact with the zonal price determination. Having a clear picture of these relationships, then it would be easy to obtain policy indications for future investments on an optimal technology mix, investments in additional capacity and in network interconnections. A procedure has been followed to select the variables which most significantly influence prices. Finally combined models have been estimated for each zone in which the Italian market is organized.

The last contribution of the paper is the use of models with explanatory variables for short term forecasting. A rolling window procedure has been applied to assess the forecasting performance of the best model for each zone, analyzing the superiority of selected models with respect to the simplest ones.

To summarize our contribution, we propose a forecasting model for the Italian electricity price dynamics accounting for seasonality, volatility clustering, long memory, technologies, market concentration and congestions. The forecasting performance of the suggested models seems quite promising, providing evidence that it is very important to consider the special market structure when these prices are modeled.

The paper is structured as follows: Section 2 links our research to the existing literature. The Italian zonal structure is explained in Section 3, where technologies, market concentration and congestions are also introduced and defined. Model specifications and results are studied in Section 4, whereas the forecasting performance is presented in Section 5. Section 6 concludes.

[1] All the abbreviations refer to the Italian definitions.

2. Background and literature review

Earlier contributions proposed several specifications for the electricity price process, taking into account traded volume, as in Goto and Karolyi (2002), or price volatility, demand and margin as in Karakatsani and Bunn (2008) and again power consumption and water supply as in Koopman et al. (2007). Hence we have found precedents, but none of these has been employed in the first empirical investigations on the Italian market, to the authors' knowledge. In addition, considering recent data from 2007 to 2008, we detect important features of Italian spot prices implementing models with daily median prices accounting for spiky behavior, technologies determining zonal prices, indicators of market concentration and also congestions among contiguous zones. Following Haldrup and Nielsen (2006), we propose to consider possible congestions among zones, where a *congestion* is identified every time we observed different zonal prices. The technical factors underlying transmission network congestions may have a crucial influence over the behaviour of generators resulting in the allocation of production and this may affect the final prices paid for electricity. Hence generation, congestions and market power are strongly interdependent factors as in Furió and Lucia (2009). Therefore as Zarnikau and Lam (1998) and Lisea et al. (2008) point out, the transmission capacity plays an important role in controlling congestions, reducing the impact of market power and improving market competitiveness.

In simple words, a generator has market power if it is able to raise the electricity price above marginal cost without experiencing a significant decline in demand. Previous studies focussed on this topic in the electricity generation sector relying on oligopoly theory, implementing simulation techniques to model the electricity generators' behaviour, see Green and Newbery (1992), Newbery (1998), and Wolfram (1998, 1999). Some others proposed empirical research as Wolak and Patrick (1997), Wolak (2000), and Borenstein et al. (2000), Helman (2006), Bask and Widerberg (2009). For a survey on models to detect market power see Fridolfsson and Tangeras (2009).

Traditionally, analysts and anti–trust regulators investigate market power issues using various measures of market concentration such as the popular *Hirschmann-Herfindahl* index (HHI), computed as the sum of the shares of the volumes sold in the market by market participants (see Murry and Zhu, 2008 and Blumsack et al., 2002 among others); and the *Residual Supply* index (RSI), which gives indications on the presence of residual market participants necessary to cover demand. Since there is not a consensus on which measure is the best indicator of market power for the electricity markets, because there is a number of factors to account for (transmission constraints are an example), we have decided to consider both of these structural indexes.

Moreover, we address the issue of forecasting electricity prices since market participants need specific information on a short–term period to set their optimal bidding strategies, or on a longer term to base bilateral contracts. Therefore price forecasting is essential to both agents and practitioners.

As Karakatsani and Bunn (2008) and Trapero and Pedregal (2009) suggest, it is possible to move from classical methods for the analysis of time series to models for unobserved components, considering dynamic regressions (as in Nogales et.al., 2002 and Karakatsani and Bunn, 2008), structural time series and ARIMA models (as Cuaresma et.al., 2004; Conejo et.al., 2005; Bowden and Payne, 2008; Gianfreda and Grossi, 2009), jump diffusion (see Johnson and Barz, 1999; Skantze and Illic, 2000; Knittel and Roberts, 2005) and regime-switching (Huisman and Mahieu, 2003; Haldrup et.al., 2010), among others techniques. However only few papers consider fundamental drivers or explanatory variables in assessing the forecasting performance. In details, demand, margin and scarcity were implemented in Karakatsani and Bunn (2008), whereas load and air temperature were used in Weron and Misiorek (2008). Here we consider production technologies, concentration and congestions in assessing the forecasting performance of selected models for zonal prices.

3. The Italian Zonal Market Structure

The Italian wholesale electricity market started its operations in April 2004 but became an Exchange only in 2005 registering an increasing in traded volumes from 73 TWh in 2004 to 232 TWh in 2008. It is important to emphasize that since this market is comparatively young there are continuous structural changes as for instance the abolition of the Calabria zone and its inclusion in the Southern zone starting from the beginning of 2009[2]. As other electricity markets, the Italian spot market consists of the day–ahead market (*Mercato del Giorno Prima*, MGP), the adjustment market (*Mercato di Aggiustamento*, MA) and the ancillary services market (*Mercato dei Servizi di Dispacciamento*, MSD).

[2]Hence investigations refer only to a time period going from January 2005 to the end of 2008.

The Italian independent system operator, Gestore del Mercato Elettrico (GME), operates on the day–ahead market (MGP) which is an auction market where participants start to submit their offers for sales and purchases nine days before and up to at 09:00 of the delivery day, when the MGP closes. Then according to the economic merit order criterion and to the capacity limits of the transmission lines between zones, offers and bids can be accepted. The accepted supply offers are evaluated at the clearing price of the zone. This price is the equilibrium price determined on hourly basis by the intersection of the demand and supply curves. Hence the *zonal market clearing prices* are those prices observed on several zones or areas, and they can differ across zones if a proportion of the grid becomes congested and so separated from the entire network (Weron, 2006). On the other hand, the accepted demand bids are evaluated at the single national price (*Prezzo Unico d'Acquisto*, PUN) which is the purchase price for end customers and it is computed as the average of the zonal prices weighted by zonal consumptions. On the adjustment market (MA) opening at 10.30 and closing at 14.00, participants can modify their positions resulting from the MGP market submitting additional supply offers and demand bids but now the zonal prices are used to evaluate the accepted purchase bids.

At 14.30 the transmission system operator, Terna S.p.A., starts its operations on the ancillary services market (MSD) and until 16.00 manages and controls the power system, cross zonal congestions and real-time balancing.

3.1. Technologies

Italian electricity is produced by the following plants: thermal power plants only with coal, or with fuel oil or with natural gas; as well as multifuel thermal power plants with oil and coal or with oil and natural gas; combined cycle gas turbines (CCGT); hydro power plants with pumped storage, with run of the river (fluent) or with reservoirs (modulation); gas turbine plants (GT); wind power plants and finally other generation plants not included in the previous ones. These twelve technologies have been used in a previous investigation of Italian zonal price dynamics (Gianfreda and Grossi, 2009) to detect influences of generation sources on price and volatility[3]. Contrary to

[3]They firstly used the *marginal technology index* (MTI) which gives indications on the technology fixing the price over one zone.

what done by the GME[4], we have decided to cluster all previous technologies into the following six types of the MTI index for a better representation of zonal generations and distinguishing between oil, gas and coal producing plants[5]: *Coal* (all multifuel and thermal power plants with coal), *Thermal* (plants without coal), *Hydro*, *Wind* (renewables), *CC* that is combined cycles (CCGT and GT) and finally *Other* plants not included in the previous ones. As proposed by Gianfreda and Grossi (2009), we compute for every group of technology the number of hours (frequency) in which it has fixed the price over the corresponding zone and we built a set of 6 dummies, one for each group, and we attributed one to the group with the maximum frequency over the day and zero to the others. Formally, let f_{rjt} the number of hour for the r-th technology group used in zone j on day t, the dummy variable for the r-th group in zone j is then defined as

$$\begin{aligned} &d_{rjt} = 1 \text{ if } f_{rjt} = max_r(f_{rjt}) \\ &d_{rjt} = 0 \text{ otherwise.} \end{aligned}$$

	Coal	CC	Thermal	Wind	Hydro	Other
North	73	632	366	0	449	40
CNorth	122	462	702	0	218	26
CSouth	143	362	817	0	183	26
South	151	356	815	0	185	25
Calabria	188	351	810	0	156	28
Sicily	18	325	1106	0	59	1
Sardinia	296	274	700	0	251	20

Table 1: Frequencies (number of days) of Technologies fixing the price over individual zones

From the summary reported in Table 1, it is possible to exclude two technologies, *Wind* and *Other* in all zones, from our analysis since they have a low influence compared to the other sources.

[4]In the annual report GME, 2008b the following groups of technologies have been considered yearly and so at the national level: other, pumped storage, modulation, fluent, CCGT, thermal conventional, see page 99.

[5]It is well documented that oil and gas have similar and correlated dynamics whereas coal has a dissimilar behavior.

3.2. Market Concentration

The number of operators has increased progressively through years growing from 66 and 76 of the sale side bidding and of the demand side bidding respectively on May 2005 to 98 and 95 sales and purchases operators on December 2008. These numbers refer to participants of both exchange and bilateral markets (GME, 2008a). Participants registered on the IPEX market increased from 51 in 2005 to 151 in 2008 (GME, 2008b). It could be possible to consider the number of market participants because as this number increases the market becomes more competitive and the price should decrease. Hence we expect to observe a reduction of national and zonal prices with the progressive increasing of competition. However this information can be used only at a national level. Instead we have decided to use two popular indexes, the *Hirschmann–Herfindahl* index (HHI) and the *Residual Supply* index (RSI), since they are available for all market zones[6].

3.2.1. The Hirschmann–Herfindahl Index (HHI)

The HHI measures the degree of concentration and dispersion of volumes sold (and/or offered[7]) by market participants for each hour and each zone. It is the sum of the shares of the volumes sold in the market by market participants as indicated in the following equation

$$HHI(j,h) = \sum_{i=1}^{N} [Q_i(j,h) * 100]^2 \tag{1}$$

with

$$Q_i(j,h) = \frac{V_i(j,h)}{\sum_{i=1}^{N} V_i(j,h)} \tag{2}$$

where $i = 1, \ldots, N$ are market participants, j represents the individual zones, h is the considered hour and finally V_i are volumes sold by the i–th participant.

[6] It is well–known that the HHI is a traditional structural index which measures static concentration and it represents just one of major sign of market power (see Hellmer and Warell, 2009), however these indexes have been used as common and initial screening.

[7] The shares are defined by considering the volumes sold and/or offered (including those covered by Bilateral Contracts) by individual market participants aggregated on the basis of the group to which they belong.

The range values of this index are 0 when there is *perfect competition* and 10,000 points when there is *monopoly*. It is common practice to distinguish among the following intervals: if $HHI \leq 1000$ then the market is said to be *unconcentrated* equivalent to N firms with equal market shares, if $1000 < HHI < 1800$ then the market is said to be *moderately concentrated* and finally if $HHI \geq 1800$ then the market is *highly concentrated* or *poorly competitive* which is equivalent to have between 50% or 60% of N firms with equal market shares.

From a preliminary analysis of the Italian zonal HHI provided by GME (Table 2), it is possible to state that in all Italian zones (apart from North) there is a poor competition on the generation side producing expectations on a direct relation between price and HHI, since when the latter increases then the price should increase as an effect of market concentration (or market power). Looking at time series of certain hours belonging to delivery periods

	Unconcentration $HHI \leq 1000$	Moderate Concentration $1000 < HHI < 1800$	Concentration $HHI \geq 1800$
North	1,06	90,49	8,45
CNorth	1,08	0,97	97,95
CSouth	1,93	0,11	97,96
South	2,01	2,57	95,42
Calabria	2,05	0,00	97,95
Sicily	2,03	1,02	96,95
Sardinia	2,05	0,00	97,95

Table 2: Percentages of HHI levels with respect to the employed sample of 35064 hours

off–peak 1, off–peak 2 and peak[8], it is possible to see that there is a sensible shift in level in the HHI hourly series during the entire month of November 2008. In that period we observed a shift in the HHI dynamics but similar behaviors can be seen neither in the quantities sold nor in the Residual Supply Index (RSI)[9].

[8]The delivery periods for the Italian market refer to the following groups of hours: *off peak 1* from 00.00 to 06.00 until the end of 2005 then from 2006 to 07.00; *peak* is from 07.00 (08.00 from 2006) to 22.00 (to 20.00 from 2006); *off peak 2* from 23.00 (or 21.00 from 2006) to 24.00.

[9]In the last quarter of 2008 across all zones and groups of hours, there has been observed a drop on sold volumes determined by the principal operator in favor of all other

3.2.2. The Residual Supply Index (RSI)

The Residual Supply Index measures the presence of residual market participants necessary to cover the total demand, thus the index measures the ex-post residuality. The hourly zonal RSI published by GME has the following formulation

$$RSI_i(j,h) = \sum_{l=1,l\neq i}^{N} S_l(j,h) - V_i(j,h) \qquad (3)$$

where $l, i = 1, \ldots, N$ are market participants, j represents the individual zones, h is the considered hour and finally V_i are volumes sold by the i–th participant. This difference between the total supply and the sum of ith sellers' supply (or in other words the quantity offered by other market participants) represents the *non-contestable volumes.* Hence dividing this quantity by the total quantities sold in one zone at one particular hour, we determined the hourly and daily aggregated *true residual supply index*, $TRSI_i$. If the index is less than 1, then the ith firm is necessary to cover the demand and so it is a *pivotal* supplier in the market; if the index is greater or equal to 1, then the ith firm is not necessary and the market can be considered competitive, see Manuhutu and Owen (2009) and Rahimi and Sheffrin (2003).

3.3. Congestions

The Italian market has been then segmented into several zones as a consequence of congestions. In this paper we do not include into the analysis either the foreign virtual zones[10], the limited production poles[11] or islands, but we only consider five physical national zones which are (from 2004 to the end of 2008) the following regional zones: North (*North*), Central North (*CNorth*)[12], Central South (*CSouth*), South (*South*), Calabria (*Calb*), di-

competitors, for details see GME, 2008b page 96.

[10]The *foreign virtual zones* ones are neighboring markets as Austria, Corsica, France, Greece, Slovenia and Switzerland.

[11]The *limited production poles* only inject electricity into the systems. We find Brindisi and Rossano among others.

[12]We assume that CNorth directly connected is connected with Sardinia even if it happens through Corsica.

Figure 1: Italian market structure

rectly connected with Sicily. Electricity flows in both directions[13], and so a congestion occurs every time the transmission capacity is exceeded. Figure 1 represents the Italian zonal market structure with circles indicating the limited production poles, gray arrows represent direct electricity flows whereas black ones are flows *assumed* as direct. Therefore transmission limits or, in addition, dissimilar suppliers' behavior can cause differences between zonal marginal prices.

Preliminary investigations performed on couples of zonal daily median prices provided evidence on the importance of congestion state. Gianfreda and Grossi (2009) indeed defined the difference between zonal price and the single national price (PUN) as a marginal congestion cost and showed that the Italian market is inefficient since not all zonal prices are equal to the PUN prices. Instead of using congestions costs as defined in Hadsell and Shawky (2006) and implemented in Gianfreda and Grossi (2009), we identify and define daily time series of frequencies of congestions every time we observe different zonal prices among contiguous zonal couples which are are North–

[13]In addition to the previous assumption, we also consider a direct connection through South and Calabria even if it happens through Rossano, a limited production pole.

CNorth, CNorth–CSouth, South–Calb, Calb–Sici and finally CNorth–Sard. In addition, since one zone as CNorth is connected with North, CSouth and Sard, we have added frequencies of congestions at all borders adjusting for total hourly congestions[14]. Details on occurrences of these daily frequencies referred to studied years are reported in Table 3.

	2005	**2006**	**2007**	**2008**
North	1511	3035	2927	1040
CNorth	5296	5639	5552	4687
CSouth	1048	353	580	1438
South	704	2144	361	587
Calabria	5017	6005	4926	6175
Sicily	4341	3926	4593	5744
Sardinia	2765	2316	2073	2365

Table 3: Sum of daily frequencies of zonal congestions

4. Model Specifications and Empirical Results

4.1. Preliminary Analysis and Model Selection

A preliminary empirical analysis of the Italian zonal market carried out using daily medians of prices has provided evidence of the presence of seasonality at daily level and a long memory autocorrelation structure (see Gianfreda and Grossi (2009)).

Figure 2 represents correlograms (AutoCorrelation Function, ACF, and Partial AutoCorrelation Function, PACF) for seasonally adjusted prices collected in the Northern zone. Seasonal adjustment has been carried out by using a linear model with dummy regressors for days of the week and calendar effects (*CalEf*). The pattern of ACF is typical of a long memory process, while the first few lags of the PACF are outside the probability bands. This means that the seasonal adjustment did not capture all the weekly seasonal dependence of the series. For this reason we decided to use the original series

[14]For example, we have counted 46 congestions in one day in CNorth adding up observed frequencies of congestions at all three borders. Then we have divided the daily amounts by the daily total possible congestions for that zone, that is by 72 (accounting for 24 hours in a day and for 3 zones). Similarly for the other zones.

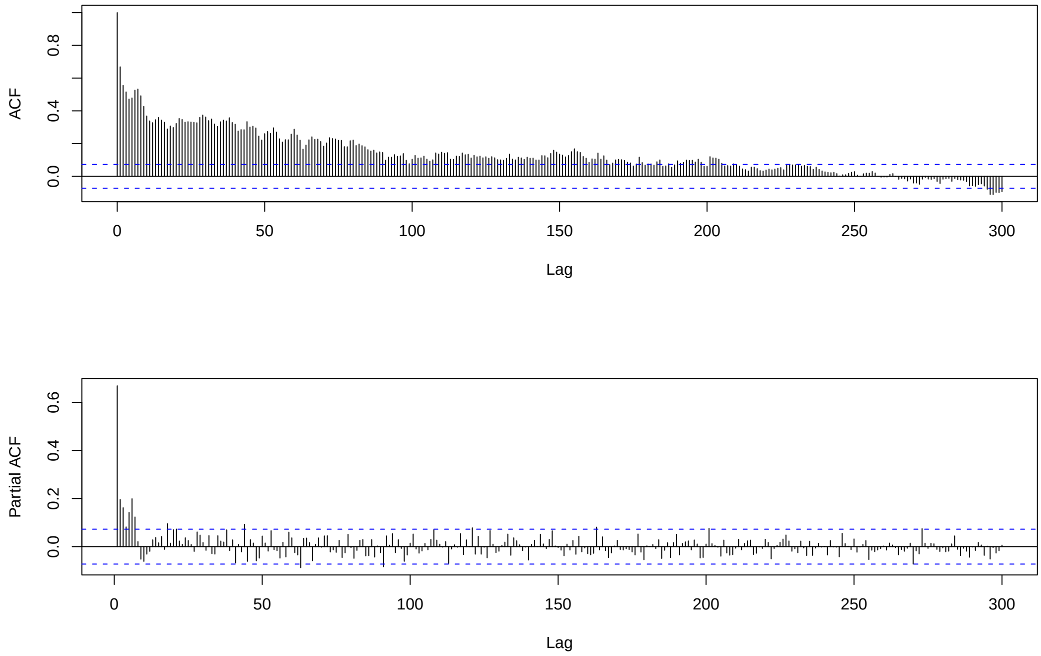

Figure 2: Autocorrelation functions of seasonally adjusted prices for North

as dependent variable and explicitly model the the seasonal pattern of the series.

	North	**CNorth**	**CSouth**	**South**	**Calb**
PP test	-19.859	-20.565	-19.086	-19.154	-20.006
KPSS test	0.435	0.295	0.406	0.395	0.389

Table 4: Stationarity (KPSS) and unit roots (PP) tests for daily prices. Thresholds at 1% and 5% level are -3.970 and -3.415 for PP test and 0.216 and 0.146 for KPSS test

In order to test the hypothesis of long memory process, the KPSS test for stationarity and the Philips-Perron test for unit roots have been applied (see Table 4 for results). Both tests reject the null hypothesis confirming the long memory process as generating the prices series.

A significant debate has considered the properties of tests for unit root, or long memory in the presence of structural breaks. It has been shown that persistence tests are severely compromised, in terms of their size and power properties, in series which display breaks (Banerjee and Urga, 2005). However, there are no reasons to consider the presence of structural breaks in the Italian market in the considered time period. Long memory is usually captured by fractionally integrated processes. Then, taking into account the autocorrelation structure we could estimate ARFIMA models with seven Autoregressive terms, that is a ARFIMA(7,0) or seven Moving Average terms, that is a ARFIMA (0,7). Another stylized fact that should not be neglected is the mean reversion of electricity prices which is usually estimated by a one–lag autoregressive term. For this reason we added one AR term also in the case of the ARFIMA model with seven Moving Average terms. Finally, the two estimated ARFIMA models are: ARFIMA(7,0) and ARFIMA(1,7). Residual diagnostics[15] from these models show that ACF and PACF functions are inside the confidence regions, but the null hypothesis of homoscedasticity, according to the Engle LM test, cannot be accepted. For this reason, we estimated ARFIMA–GARCH models and residuals diagnostics, reported in Table 5, lead to the acceptance of the ARFIMA(1,7) combined with a GARCH(1,1) against an EGARCH(1,1) model, hence supporting the evidence against a *leverage effect* characterizing Italian prices, provided in Gianfreda (2010).

The best model, that is the ARFIMA(1,7)–GARCH(1,1), also according to the information criteria (AIC and BIC), has been used as the basic one for testing the influence of exogenous explanatory variables on wholesale zonal prices. Moreover, we estimated the models under the assumption of different distributions for residuals to take into account the presence of many extreme values and consequent fat tails of the distribution of electricity prices. The best performance has been obtained using a Student–t distribution. Hence, the effect of exogenous factors on wholesale prices has been measured implementing Reg–ARFIMA–GARCH models, as in Koopman et al. (2007), with dummies for group of technologies, frequencies of congestions and the market concentration indexes. In the next subsection, the model specification will be formalized, while in the next three subsections we are going to explore the relation between prices and single groups of explanatory variables (technologies, congestions and market power) within the framework of

[15]Tables are not reported for lack of space but are available on request.

	NORTH	CNORTH	CSOUTH	SOUTH	CALB
Q–Statistics on Standardized Residuals					
Q(15)	0.302	0.406	0.340	0.369	0.251
Q(20)	0.234	0.560	0.656	0.724	0.454
Q(30)	0.112	0.373	0.242	0.394	0.165
Q–Statistics on Squared Standardized Residuals					
Q(15)	0.581	0.915	0.704	0.627	0.749
Q(20)	0.640	0.972	0.718	0.609	0.775
Q(30)	0.635	0.552	0.042	0.715	0.929
Diagnostic test based on the news impact curve (EGARCH vs. GARCH)					
Sign Bias Test	0.542	0.594	0.459	0.274	0.535
Negative Size Bias Test	0.851	0.620	0.820	0.674	0.835
Positive Size Bias Test	0.712	0.780	0.719	0.682	0.512
Joint Test	0.749	0.891	0.720	0.479	0.618
LM Engle test					
ARCH 1–2 test	0.742	0.984	0.851	0.750	0.744
ARCH 1–5 test	0.390	0.990	0.362	0.294	0.342
ARCH 1–10 test	0.575	0.835	0.475	0.385	0.511

Table 5: P–values of residuals diagnostic tests for ARFIMA(1,7)–GARCH(1,1) models estimated in five zones

the ARFIMA–GARCH processes. The combined effect of most significant variables will be estimated in the final subsection.

4.2. Model Specifications

The proposed models can be formalized as follows:

$$(1-L)^d(y_t-\mu_t) = \varepsilon_t + \theta_1\varepsilon_{t-1} + \ldots + \theta_q\varepsilon_{t-q} \qquad \varepsilon_t|I_{t-1} \sim NID(0,\sigma_t^2) \quad (4)$$

where

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1} + \beta\sigma_{t-1}^2 \quad (5)$$

for $t = 1, ..., T$, y_t is the zonal median electricity price at time t, L is the lag operator defined by $Ly_t = y_{t-1}$ and $\mu_t = E(y_t|I_{t-1})$ is the mean equation conditioned to the set of information available at time $t-1$. The $\theta_j\varepsilon_{t-j}$ terms represent the moving average component of the price dynamics with coefficients θ_j for $j = 1, \ldots, q$. The following specification has been considered for the conditional mean function:

$$\mu_t = \nu_1 D_1 + \ldots + \nu_7 D_7 + \gamma CalEf + \phi_1 y_{t-1} + \lambda_v x_t \quad (6)$$

where D_j with $j = 1, \ldots, 7$ are dummies for days of the week and ν_j are the corresponding coefficients; $CalEf$ is a dummy accounting for calender effects and γ is the corresponding coefficient; the $\phi_1 y_{t-1}$ term represents the autoregressive component of the price dynamics for $t = p+1, ..., T$; x_t represent explanatory dummy variables with $v = 1, 2, 3$ indicating respectively technologies determining the price when $v = 1$ and $x_t = Tech_t$, the index of market power when $v = 2$ and $x_t = MarPow_t$, and finally congestions for daily events when $v = 3$ and $x_t = Cong_t$; λs are regression coefficients. Therefore we have initially tested effects of individual explanatory variables to understand their implications on zonal prices, then we have selected the significant ones to verify their combined effects on prices.

4.2.1. Effects of Technologies

To control for technologies, we have estimated an ARFIMA(1,7)–GARCH(1,1) with $x_t = Tech_t$ in the conditional mean equation. Looking at Table 6, it is possible to draw the following comments on the employed *groups of technologies* determining zonal prices. First of all, *Hydro* and *Other* are never found to be significant in all zones, so even if determining the zonal prices these factors are uninfluential to the price dynamics. *Coal* is found to be

	Technology	Coef	Std. Err.	t–stat	p–value	AIC	
Calb	Coal	-1.073	0.590	-1.819	0.069	6.952	(*)
	CC	-2.507	0.655	-3.827	0.000	6.926	(***)
	TNC	3.031	0.723	4.194	0.000	6.919	(***)
	Hydro	0.514	0.652	0.788	0.431	6.955	
	Other	1.466	1.849	0.793	0.428	6.954	
CNorth	Coal	-1.110	0.635	-1.747	0.081	6.845	(*)
	CC	-1.177	0.522	-2.256	0.024	6.840	(**)
	TNC	2.113	0.698	3.026	0.003	6.831	(***)
	Hydro	0.048	0.606	0.080	0.937	6.849	
	Other	1.597	2.736	0.584	0.560	6.847	
CSouth	Coal	-0.737	0.598	-1.233	0.218	6.927	
	CC	-2.272	0.661	-3.436	0.001	6.903	(***)
	TNC	2.852	0.725	3.935	0.000	6.894	(***)
	Hydro	0.040	0.628	0.064	0.949	6.928	
	Other	0.977	1.680	0.582	0.561	6.928	
North	Coal	-1.334	1.001	-1.332	0.183	6.902	
	CC	-2.491	0.606	-4.108	0.000	6.875	(***)
	TNC	3.849	0.999	3.853	0.000	6.873	(***)
	Hydro	0.791	0.649	1.219	0.223	6.902	
	Other	2.347	1.609	1.458	0.145	6.900	
South	Coal	-0.503	0.581	-0.866	0.387	6.931	
	CC	-2.321	0.659	-3.523	0.001	6.906	(***)
	TNC	2.877	0.733	3.925	0.000	6.899	(***)
	Hydro	0.035	0.649	0.054	0.957	6.932	
	Other	1.567	1.903	0.823	0.411	6.930	

Table 6: Estimates of Reg–ARFIMA–GARCH models for the only effects of technologies

marginally significant (at 10% confidence level) only in CNorth and Calabria, on the contrary *Combined Cycles*, CC, and *Thermal* power, TNC, are always significant. In details, the former technology reduces electricity zonal prices, whereas the latter increases them. According to these results, we have selected only the last two variables to be included in the final model formulation.

4.2.2. Effects of Market Concentration

Concentration or exercise of market power has been investigated using two indexes in the conditional mean, then meaning respectively that $x_t = HHI_t$ in the first case and $x_t = TRSI_t$ in the second one. Looking at results (Table 7), and in view of our expectations[16] and previous considerations on the problematic behavior of the HHI time series, we have decided to use the TRSI to account for market power in the following analysis.

	Index	Coef	Std. Err.	t–stat	p–value	AIC	
Calb	HHI	0.000	0.001	-0.056	0.955	6.956	
	TRSI	0.204	0.105	1.947	0.052	6.952	(*)
CNorth	HHI	-0.001	0.001	-0.947	0.344	6.847	
	TRSI	0.123	0.040	3.081	0.002	6.832	(***)
CSouth	HHI	0.001	0.001	1.660	0.097	6.924	(*)
	TRSI	0.046	0.024	1.901	0.058	6.921	(*)
North	HHI	No convergence					
	TRSI	0.837	0.087	9.638	0.000	6.726	(***)
South	HHI	0.001	0.001	1.133	0.258	6.929	
	TRSI	0.254	0.074	3.411	0.001	6.910	(***)

Table 7: Estimates of Reg–ARFIMA–GARCH models for the only effects of market power

4.2.3. Effects of Congestions

Considering Table 8, we find that *congestions* affect zonal prices but sometimes with surprising signs. First of all we observe that these events do not affect South and CNorth. Specifically, in the former case the motivation could lie in the presence of a limited production pole, Brindisi, which injects electricity into the system even if the zone is congested so separated from Calabria and CSouth; in the latter case, congestions with North and CSouth do not affect CNorthern prices, whereas they can be increased by electricity flows with Sardinia. As expected, congestions increase CSouthern prices and prices in Calabria when the lines with South are congested. Moreover

[16]Market concentration seems to affect all zones apart one, hence both indexes are expected to be significant, and so inducing increasing prices as result of exercise of market power.

Calabrian prices decrease when there is a congestion with Sicily since here there is another limited production pole, Rossano. Interestingly, Northern prices are affected by congestions, but surprisingly these decrease zonal prices (at 1% confidence level) and it can be argued that when congestions occur the Northern demand could be satisfied by imports from foreign markets. Subsequently, only significant congestion events have been considered.

Zonal Prices	Couples	Coef	Std. Err.	t–stat	p–value	AIC	
Calb	South↔Calb	3.310	1.787	1.852	0.065	6.949	(*)
	Calb↔Sici	-0.911	0.469	-1.941	0.053	6.951	(*)
CNorth	North↔CNorth	0.827	0.636	1.300	0.194	6.846	
	CNorth↔CSouth	-1.601	1.117	-1.433	0.152	6.845	
	CNorth↔Sard	1.053	0.560	1.881	0.060	6.844	(*)
CSouth	CNorth↔CSouth	3.333	1.706	1.954	0.051	6.917	(*)
	CSouth↔South	2.261	1.234	1.832	0.067	6.927	(*)
North	North↔CNorth	-3.904	0.736	-5.303	0.000	6.858	(***)
South	South↔Calb	0.662	1.368	0.484	0.628	6.932	
	CSouth↔South	1.184	2.646	0.448	0.655	6.932	

Table 8: Estimates of Reg–ARFIMA–GARCH models for the only effects of congestions on zonal prices

4.2.4. Combined Effects

After the appropriate variable selection, we have tested the model for all significant explanatory variables in the conditional mean, that is considering

$$\mu_t = \phi_1 y_{t-1} + \lambda_1 Tech_t + \lambda_2 MarPow_t + \lambda_3 Cong_t. \qquad (7)$$

Table 9 shows the maximum likelihood estimates of Reg–ARFIMA–GARCH parameters applied to time series of daily median prices.

	NORTH		CNORTH		CSOUTH		SOUTH		CALB	
Const (mean)	54.09	(***)	50.64	(***)	54.35	(***)	51.89	(***)	36.30	(***)
CC	-1.37	(**)	-0.68		-1.44	(**)	-1.55	(**)	-1.69	(***)
TNC	1.30		1.38	(*)	1.84	(***)	1.74	(**)	2.16	(***)
TRSI	0.77	(***)	0.11	(**)	0.04		0.17	(**)	0.18	
Cong	-2.89	(***)	1.18	(**)	2.15		0.80		1.96	
CalEf	-12.55	(***)	-21.83	(***)	-22.32	(***)	-22.89	(***)	-21.81	(***)
Mon	8.22	(***)	20.14	(***)	20.17	(***)	19.80	(***)	18.99	(***)
Tue	11.11	(***)	21.27	(***)	21.14	(***)	20.89	(***)	20.41	(***)
Wed	10.84	(***)	21.07	(***)	21.17	(***)	20.96	(***)	20.58	(***)
Thu	10.11	(***)	20.51	(***)	20.41	(***)	20.18	(***)	19.89	(***)
Fri	8.76	(***)	19.86	(***)	19.29	(***)	18.89	(***)	18.86	(***)
Sat	2.16	(**)	5.59	(***)	5.90	(***)	6.44	(***)	6.09	(***)
d–Arfima	0.50	(***)	0.49	(***)	0.48	(***)	0.47	(***)	0.49	(***)
ϕ_1	0.19	(**)	0.22	(*)	0.19	(***)	0.13	(**)	0.18	(**)
θ_1	-0.26	(**)	-0.23	(**)	-0.16		-0.14		-0.16	
θ_2	0.00		-0.10	(**)	-0.14	(***)	-0.14	(***)	-0.12	(***)
θ_3	-0.02		-0.03		-0.05		-0.07		-0.08	(*)
θ_4	-0.01		0.00		0.00	(***)	0.01		0.02	
θ_5	-0.03		-0.01		0.01		-0.01		-0.01	
θ_6	0.04		0.07	(*)	0.03		0.03		0.02	
θ_7	0.22	(***)	0.22	(***)	0.24	(***)	0.24	(***)	0.22	(***)
ω	1.08		2.39	(*)	1.74	(*)	2.24	(*)	2.03	
ARCH α	0.06	(***)	0.12	(***)	0.10	(***)	0.12	(***)	0.10	(***)
GARCH β	0.92	(***)	0.86	(***)	0.89	(***)	0.86	(***)	0.88	(***)
Student(DF)	6.37	(***)	4.42	(***)	4.85	(***)	5.21	(***)	5.10	(***)

Table 9: Reg–ARFIMA estimates (with p–values in brackets) for Italian Electricity Zonal Prices

4.3. Preliminary Comments

Looking at previous estimates, we can draw the following preliminary conclusions:

1. *Calendar effects*, *seasonality*, *fractional integration* as well as *volatility clustering* are important and salient features to take into account since the estimates of $CalEf$, days of the week, d, α and β are always significant. Moreover d is less than 0.5 for all zones, as found previously in Gianfreda and Grossi (2009), hence confirming that these price processes have long memory.
2. The *autoregressive structure*, that is the ϕ_1 term, is found to be important to capture the stylized fact of mean–reversion of electricity prices. Whereas the inclusion of moving average terms has been used to obtain white noise residuals.
3. The employed *groups of technologies* determining the zonal prices are generally significant across zones. And the final formulation confirms that *Combined Cycles* (CC) always reduce electricity zonal prices, whereas *Thermal* power (TNC) generally increases them.
4. *Concentration* has been analyzed making use of the TRSI. More precisely, this index indicates competitive markets when it approaches one, hence it should reduce zonal prices. However, we observe a positive sign in all considered zones implying that the residual supply was not sufficient to cover zonal demand hence inducing prices to increase. Interestingly, in the final formulation it turned out to become non–significant in CSouth and Calabria[17].
5. Finally, *congestions* are important only in North and CNorth. In the first case, it can be argued that when congestions affect the Northern zone, demand could be satisfied by imports. On the other hand, when congestions occur in CNorth electricity prices raise because of an excess of demand. Now prices in CSouth and Calabria turn not to be influenced by congestions.

[17] Plotting time series of TRSI, it can be seen (graphs not reported but available on request) that firstly it is almost always equal to 1 in Calabria, which is a special case give its interconnection through Rossano, and secondly it is fairly close to one in CNorth and CSouth.

5. Assessing forecasting performance

In this section the forecasting performance of Reg–ARFIMA–GARCH model with conditional mean equation specified in (6) is investigated. We assume to have knowledge of history up to the end of June 2008 and try to assess the performance ability of the model. In other words, we use daily data from 01/01/2007 until 30/06/2008 as a sort of "training data set" and measure the forecasting performance of the model until the end of 2008.

We evaluate the out–of–sample forecasting performance of the models we use a "rolling windows" procedure. In order to make clear how this technique works, the whole time period is divided in two sub-periods, the first going from $t = 1$ to $t = T-m$ and the second covering the period from $t = T-m+1$ to T. The procedure is iterative as we use a different set of information for estimating purposes rolling a windows of $T-m$ observations over the original data-set. Every time the estimated parameters are used to get a one-step-ahead forecast. Going into details, the rolling windows procedure works as follows:

- at time $T-m$ the vector of estimates θ_{T-m} is obtained through different models (RW1, RW7, Basic and Final) using data for $t = 1, \dots, T-m$; the forecast in $T-m+1$, is then given by

$$y_{T-m+h|T-m} = f(\theta_{T-m}, y_{T-m}).$$

- at time $T-m+1$ the forecast for time in $T-m+h+1$ is obtained on data for $t = 2, \dots, T-m+1$, that is

$$y_{T-m+h+1|T-m+1} = f(\theta_{T-m+1}, y_{T-m+1})$$

- ...

- the last forecast is estimated at time $T-h$, using data for $t = d, \dots, T-h$

$$y_{T|T-h} = f(\theta_{T-h}, y_{T-h}).$$

At the end of the iterative procedure, $m-h+1$ h–step ahead forecasts are obtained. Analyzing daily data of electricity prices, if $m = 180$, we can check the h–day ahead forecasting performance for the last six months of data.

To evaluate the gain obtained by using exogenous variables we considered three benchmark models: a simple random walk (RW1), a weekly random walk (RW7) and the ARFIMA(1,7)–GARCH(1,1) model without regressors (form now on, called "Basic Model". The RW1 is a classical benchmark model whose forecasts are commonly called "naive" predictions. The forecast function of the random walk is $y_{t|t-1} = y_{t-1}$, that is the observed average price of yesterday is the forecast for today. We take the value of two days ago if the there was a holiday yesterday. The number of days in the past is increased accordingly when there are two or more contiguous holiday days. The RW7 is a forecast method which has been used as benchmark model in previous papers on electricity loads forecasting (Taylor and McSharry, 2007; Dordonnat *et al.*, 2008). The forecast function for the RW7 is $y_{t|t-1} = y_{t-7}$, that is the average price observed one week ago is the forecast for today. We take the value of two weeks ago if there was a holiday one week ago. Anyway special days, holidays included, arise many problems, thus we have deleted these forecasts for the RW1 and RW7 since we consider them only as benchmark models. The Basic Model is a restricted version of the general model with regressors (from now on, called "Final Model"). The comparison between the Basic and the Final models is carried out to evaluate how the exogenous variables can improve the forecasting performance.

Figure 3 presents the one-day-ahead relative forecast errors (forecast errors divided by observed prices) for the North macro-region as well as their autocorrelation function using the Final model. On the whole the forecasts seem unbiased. The largest relative errors correspond to the final days of 2008. The correlations are slightly outside the confidence region only for few lags, but in general the correlogram does not show any particular pattern that could be considered as a clue of a missed structural dynamic feature in the time series.

We use a set measures to assess the predictive goodness of each model: the root mean squared forecast error (RMSE), the mean absolute percentage forecast error (MAPE), the Theil's U index. For a set of m forecasts they are given by:

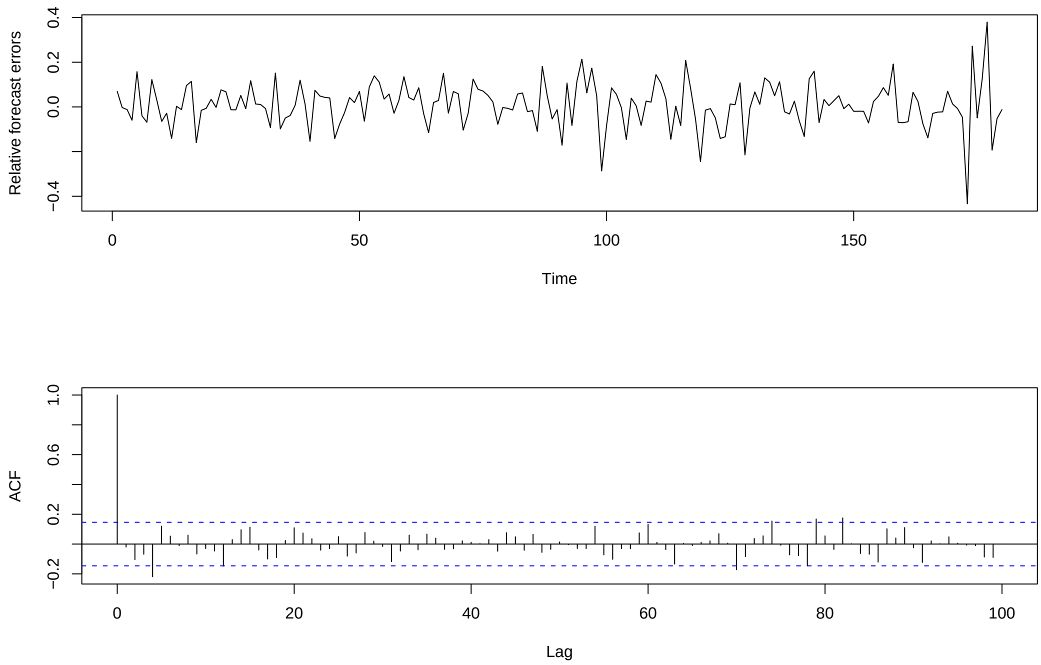

Figure 3: Out-of-sample one-step-ahead relative forecast error (upper panel) and corresponding empirical autocorrelation function (lower panel) for North region

$$RMSE = \left[\frac{1}{m} \sum_{t=1}^{m} (e_t^h)^2 \right]^{1/2} \tag{8}$$

$$MAPE = \frac{1}{m} \sum_{t=1}^{m} 100 |e_t^h / y_t| \tag{9}$$

$$U = \frac{\left[\frac{1}{m} \sum_{t=1}^{m} (e_t^h / y_{t-1})^2 \right]^{1/2}}{\left[\frac{1}{m} \sum_{i=t}^{m} ((y_t - y_{t-1})/y_{t-1})^2 \right]^{1/2}} \tag{10}$$

where $y_{t|t-h}$ and $e_t^h = y_t - y_{t|t-h}$ are the actual forecast and the forecast error, respectively, at day t for $t = T - d + 1, \ldots, T$, with d being the number of available forecasts and h being the forecasting horizon, in our case $h = 1, \ldots, 7$. We also apply the Diebold–Mariano test (Diebold and Mariano,

1995) to compare different estimated models. Given two alternative models A and B, the relevant statistic is given by

$$DM = \frac{\bar{d}}{\sqrt{\omega}} \sim N(0,1) \tag{11}$$

where $\bar{d} = \frac{1}{m}\sum_{t=1}^{m} d_t$, $d_t = (e_{t,A}^h)^2 - (e_{t,B}^h)^2$, with $e_{t,A}^h$ and $e_{t,B}^h$ are the forecast errors at time t made by the forecasts form model A and B respectively. The symbols ω in (11) is the asymptotic variance of the average difference $\bar{d}$. Diebold and Mariano (1995) suggest estimating ω by an unweighted sum of the autocovariances of d_t, denoted $\gamma_i(d)$, that is

$$\hat{\omega} = \sum_{i=-(h-1)}^{h-1} \hat{\gamma_i(d)}, \tag{12}$$

where h is again the forecast horizon for which the prediction errors are compared. As we use the one–sided test to evaluate the superiority of one model the null hypothesis of equal performance is rejected at 1% level when $|DM| > 2.33$.

Table 10 shows the forecast assessment indexes for one-day-ahead forecasts obtained from the four models cited above and for each macro–region. The MAPE and RMSE of the Basic and Final models are always far lower than those of the Random Walk models. Moreover the Theil's index always shows that considering the autocorrelation structure of the time series always lead to an improvement of the forecasts with respect to the "naive" model. Finally, the last two rows for each macro-region report the DM tests and the corresponding p-values. Models are compared by couples moving from the simplest to the more complex model. For instance in the second column the RW7 model is compared with RW1 model, in the third column the Basic model is compared with the RW7 and in the last column the Final model is compared with the Basic. The evidence from using the DM test can be summarized as follows: the RW7 model is never better than the RW1, at 1% level of significance the Basic models is always better than the RW7 and the Final model is better than the Basic in all the macro-regions but the Calabria where the p-value is slightly higher than 0.01.

Table 11 presents the MAPE index and the DM test for multi–day ahead forecasts from the Basic and Final model and considering a forecasting horizon ranging from two to seven days. The MAPE index is always lower for the Final model and the DM test is almost always significant at the 5% level with

		1 step–ahead			
		RW1	RW7	Basic	Final
NORTH	RMSE	14.072	13.001	10.710	9.577
	MAPE	10.696	10.103	8.384	7.252
	Theil's U	-	0.892	0.759	0.675
	DM	-	0.693	2.393	2.325
	p–value	-	0.244	0.008	0.010
CNORTH	RMSE	15.852	14.789	11.563	10.941
	MAPE	11.603	11.089	9.192	8.724
	Theil's U	-	0.965	0.751	0.711
	DM	-	0.571	2.344	3.109
	p–value	-	0.284	0.010	0.001
CSOUTH	RMSE	17.567	17.506	13.310	12.267
	MAPE	12.634	13.402	9.937	9.164
	Theil's U	-	0.942	0.747	0.690
	DM	-	1.020	3.090	3.489
	p–value	-	0.154	0.001	0.000
CALB	RMSE	16.947	17.233	12.886	11.968
	MAPE	12.382	13.319	9.689	9.025
	Theil's U	-	0.969	0.758	0.700
	DM	-	0.853	3.339	2.262
	p–value	-	0.197	0.000	0.012
SOUTH	RMSE	17.067	17.585	13.172	12.193
	MAPE	12.541	13.459	9.829	9.096
	Theil's U	-	0.975	0.765	0.703
	DM	-	0.530	3.327	2.555
	p–value	-	0.298	0.000	0.005

Table 10: Assessment for one-day ahead forecasts. The following indexes are reported: Root Mean Squared Error (RMSE), Mean Absolute Percentage forecast Error (MAPE), Theil's inequality coefficient (Theil 's U), Diebold and Mariano test (DM) and corresponding p-value.

MAPE

		Basic	Final	DB.stat	DB.p.value
NORTH	2 days	9.293	8.215	1.669	0.048
	3 days	9.433	8.450	1.825	0.034
	4 days	9.256	8.363	1.472	0.070
	5 days	9.129	8.443	1.419	0.078
	6 days	9.128	8.484	1.640	0.050
	7 days	9.442	8.751	1.781	0.037
CNORTH	2 days	9.695	9.171	1.795	0.036
	3 days	9.625	9.138	2.066	0.019
	4 days	9.535	9.041	1.938	0.026
	5 days	9.170	8.783	1.999	0.023
	6 days	9.276	8.913	2.168	0.015
	7 days	9.470	8.987	2.307	0.011
CSOUTH	2 days	10.935	9.977	2.024	0.022
	3 days	10.901	9.847	1.742	0.041
	4 days	10.714	9.741	1.602	0.055
	5 days	10.501	9.581	1.633	0.051
	6 days	10.613	9.654	1.775	0.038
	7 days	10.709	9.820	1.791	0.037
CALB	2 days	10.796	10.073	1.642	0.050
	3 days	10.741	9.893	1.620	0.053
	4 days	10.649	9.838	1.559	0.060
	5 days	10.409	9.743	1.650	0.050
	6 days	10.383	9.741	1.749	0.040
	7 days	10.459	9.793	1.747	0.040
SOUTH	2 days	10.852	10.051	1.811	0.035
	3 days	10.754	9.891	1.719	0.043
	4 days	10.669	9.893	1.652	0.049
	5 days	10.465	9.771	1.734	0.041
	6 days	10.383	9.706	1.810	0.035
	7 days	10.530	9.810	1.869	0.031

Table 11: Assessment for two–seven–days ahead forecasts. The following indexes are reported: Root Mean Squared Error (RMSE), Mean Absolute Percentage forecast Error (MAPE), Theil's inequality coefficient (Theil 's U), Diebold and Mariano test (DM) and corresponding p-value.

the most evident exception in the North region where the p-value are around 7% for the four and five-days ahead predictions. To conclude this section we must remember that, since the Final model includes explanatory variables, the forecasting accuracy could be biased either on realized daily values of the explanatory variables or on their one-day-ahead forecasts. There advantages and disadvantages in both approaches. The former (using realized values) maybe preferred to avoid having to discuss external inaccuracies due to explanatory variables forecast errors, while the latter (using forecasted values of explanatory variables) may be preferable since the corresponding model could be used as in real situations. As we do not have access to any type of forecasts for our explanatory variables, we can only use the first approach. Anyway a scenario analysis could be applied in case of binary variables such as congestion events and technologies. For instance, we could compare the models predicting the price for tomorrow assuming a congestion event and a given production technology.

6. Conclusions

This paper is an analysis of effects of technologies, concentration and congestions on Italian Electricity zonal prices. According to the most recent contributions in time series analysis applied to electricity prices, we took into account the long memory feature of the generating stochastic process estimating a parameter of fractional integration, which turned out to lie very close to 0.5. A causal analysis in the framework of Reg–ARFIMA–GARCH models confirmed the significant impact of production technologies, market concentration and congestions on these price dynamics. Moreover we have tested the forecasting ability of several models from a naive one to the selected one, hence showing that the performance is improved including the selected explanatory variables.

Concluding, we have provided firstly insights on relationships between zonal electricity spot prices, technologies, concentration and congestions and secondly proved that the Reg–ARFIMA–GARCH models with exogenous variables outperform better than other models when forecasting zonal prices. These results can have important implications when programming the medium–long term energy policy in Italy or future investment strategies with respect to the technology mix and, especially, the network grid since generators can serve only if there exists adequate transmission capacity whereas the installation of new power plants is expected to produce even more and sudden bottleneck problems.

Acknowledgements

The authors would like to thank the Market Monitoring Area at the GME and in particular the director Cosimo Campidoglio, Pia Saraceno, Sandro Sapio and participants at the *IEFE–FEEM* seminar at Bocconi University and at EUI in Fiesole. Special thanks go to Derek W. Bunn and Mike Clements for useful comments which lead to improve substantially the paper.

References

Banerjee A. and Urga G., (2005), *Modelling structural breaks, long memory and stock market volatility: an overview*, Journal of Econometrics, 129, pp. 1–34

Bask M. and Widerberg A., (2009), *Market structure and the stability and volatility of electricity prices*, Energy Economics, 31, pp. 278–288

Battaglia F. and Orfei L., (2005), *Outlier Detection and Estimation in Non-Linear Time Series*, Journal of Time Series Analysis, 26, 1, pp. 107–121

Blumsack S., Perekhodtsev D. and Lave L.B., (2002), *Market Power in Deregulated Wholesale Electricity Markets: Issues in Measurement and the Cost of Mitigation*, The Electricity Journal, 15, pp.11-24

Bosco B., Parisio L. and Pelagatti M., (2007), *Deregulated Wholesale Electricity Prices in Italy: An Empirical Analysis*, International Advances Economic Research, 13, pp. 415–432

Bowden N. and Payne J.E., (2008), *Short term forecasting of electricity prices for MISO hubs: Evidence from ARIMA–EGARCH models*, Energy Economics, 30, pp. 3186–3197

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

Cuaresma J.C., Hlouskova J., Kossmeier S. and Obersteiner M., (2004), *Forecasting electricity spot–prices using linear univariate time–series models*, Applied Energy, 77, pp. 87–106

Diebold F.X., Mariano R.S. (1995), Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13, pp. 253-263.

Diongue A.K., Gugan D. and Vignal B., (2009), *Forecasting electricity spot market prices with a k–factor GIGARCH process*, Applied Energy, 86, pp. 505–510

V. Dordonnat V., Koopman S.J., Ooms M., Dessertaine A. and Collet J., (2008), *An hourly periodic state space model for modelling French national electricity load*, International Journal of Forecasting, 24, pp. 566–587

Escribano A., Peña J., Villaplana P., (2002), *Modelling electricity prices: international evidence*, Economic Series 08, Working Paper 02-27, Universidad Carlos III de Madrid

Fridolfsson S. and Tangeras T., (2009), *Market power in the Nordic electricity wholesale market:a survey of the empirical evidence*, Energy Policy, 37, pp. 3681–3692

Furió D. and Lucia J.J., (2009), *Congestion management rules and trading strategies in the Spanish electricity market*, Energy Economics, 31, pp. 48-60

Garcia A. and Arbelaez L.E., (2002), *Market power analysis for the Colombian electricity market*, Energy Economics, 24, pp. 217–229

Gianfreda A., (2010), *Volume and Volatility Effects in European Electricity Markets*, Economic Notes, 39, 1, forthcoming

Gianfreda A. and Bunn D., (2010), *Integration and Shock Transmissions across European Electricity Forward Markets*, Energy Economics, 32, 2, pp. 278–291

Gianfreda A. and Grossi L., (2009), *Zonal Price Analysis of the Italian Wholesale Electricity Market*, IEEE CNF Conference Proceedings of the *European Energy Markets* Conference (EEM09), available on IEEE Xplore

Giulietti M., Grossi L. and Waterson M. (2010), *Price transmission in the UK electricity market: was NETA beneficial?*, Energy Economics, 32, pp. 1165–1174

GME, (2002), *Istruzioni alla disciplina del mercato italiano*, available at http://www.mercatoelettrico.org

GME, (2003), Technical Report April 2003 - March 2004, available at http://www.mercatoelettrico.org

GME, (2008), Monitoring Monthly Reports 2008, available at http://www.mercatoelettrico.org

GME, (2008), Technical Report 2008, available at http://www.mercatoelettrico.org

Goto M. and Karolyi G.A., (2002), *Understanding Electricity Price Volatility Within and Across Markets*, working paper available at http://papers.ssrn.com sol3/papers.cfm?abstract_id=576982

Hadsell L. and Shawky H.A., (2006), *Electricity Price Volatility and the Marginal Cost of Congestion: An Empirical Study of Peak Hours on the NYISO Market, 2001-2004*, The Energy Journal, 27, 2, pp. 157–179

Haldrup N. and Nielsen M.O., (2006), *A regime switching long memory model for electricity prices*, Journal of Econometrics, 135, 1-2, pp. 349–376

Haldrup N., Nielsen F.S. and Nielsen M.O., (2010), *A vector autoregressive model for electricity prices subject to long memory and regime switching*, Energy Economics, available on–line, doi:10.1016/j.eneco.2010.02.012

Helman U., (2006), *Market power monitoring and mitigation in the US wholesale power markets*, Energy, 31, pp. 877-904

Hellmer S. and Warell L., (2009), *On the evaluation of market power and market dominance – The Nordic electricity market*, Energy Policy, 37, pp. 3235–3241

Huisman R. and Mahieu R., (2003), *Regime jumps in electricity prices*, Energy Economics, 25, 5, pp. 425–434

Johnson B. and Barz G., (1999), *Selecting Stochastic Processes For Modelling Electricity Prices*, Energy Modelling and the Management of Uncertainty, Risk Publications

Kanamura T., Ohashi K., (2007), *A structural model for electricity prices with spikes: measurement of spike risk and optimal policies for hydropower plant operation*, Energy Economics, 29 (5), pp. 1010–1032

Karakatsani N. and Bunn D.W., (2008), *Forecasting electricity prices: The impact of fundamentals and time-varying coefficients*, International Journal of Forecasting, 24, 4, pp. 764–785

Knittel C.R. and Roberts M., (2005), *An empirical examination of restructured electricity prices*, Energy Economics, 27, 5, pp. 791–817

Koopman S. J., Ooms M., and Carnero M.A., (2007), *Periodic Seasonal Reg-ARFIMA-GARCH Models for Daily Electricity Spot Prices*, Journal of the American Statistical Association, 102, 477, pp. 16–27

Kupper G., Delarue E., Delvaux B., Meeus L., Bekaert D., Willems B., Proost S., Dhaeseleer W., Deketelaere K. and Belmans R., (2009), *Does More International Transmission Capacity Increase Competition in the Belgian Electricity Market?*, The Electricity Journal, 22, 1, pp. 21–36

Lisea W., Hobbsb B.F. and Hers S., (2008), *Market power in the European electricity market – The impacts of dry weather and additional transmission capacity*, Energy Policy, 36, pp. 1331-1343

Manuhutu C. and Owen A.D., (2009), *Gas–on–gas competition in Shangai*, Energy Policy, doi:10.1016/j.enpol.2009.06.027

Misiorek A., Trueck S. and Weron R., (2006) , *Point and Interval Forecasting of Spot Electricity Prices: Linear vs. Non–Linear Time Series Models*, Studies in Nonlinear Dynamics & Econometrics, 10, 3, article 2

Nogales F.J., Contreras J., Conejo A.J. and Espinola R., (2002) *Forecasting Next–Day Electricity Prices by Time Series Models*, IEEE Transactions on Power Systems, 17, 2, May, pp. 342–348

Mount T.D., Ning Y. and Cai X., (2006), *Predicting price spikes in electricity markets using a regime–switching model with time–varying parameters*, Energy Economics, 28, pp. 62–80

Murry D. and Zhu Z., (2008), *Asymmetric price responses, market integration and market power: A study of the U.S. natural gas market*, Energy Economics, 30, pp. 748-765

Olbermann, B.P., Lopes, S.R.C. and Reisen, V.A. (2006),Invariance of the first difference in ARFIMA models, *Computational Statistics*, 21, pp.445-461

Petrella A. and Sapio S., (2009), *A time series analysis of day–ahead prices on the Italian Power Exchange*, IEEE CNF Proceedings of the *European Energy Markets* Conference (EEM09), available on IEEE Xplore

Rahimi A.F. and Sheffrin A.Y., (2003), *Effective market monitoring in deregulated electricity markets.*, IEEE Transactions on Power Systems, 18, (2), pp. 486–493

Skantze P. and Illic M., (2000), *The Joint Dynamics of Electricity Spot and Forward Markets: Implications of Formulating Dynamic Hedging Strategies*, Energy Laboratory Report No. MIT-EL 00-005

Tamaschkea R., Docwrab G. and Stillman R., (2005), *Measuring market power in electricity generation: A long–term perspective using a programming model*, Energy Economics, 27, pp. 317-335

Taylor J.W. and McSharry P.E (2007), Short-term load forecasting methods: An evaluation based on European data *IEEE Transactions on Power Systems*, 22, 2213–2219.

Juan R. Trapero J.R. and Pedregal D.J., (2009), *Frequency domain methods applied to forecasting electricity markets*, Energy Economics, 31, pp. 727–735

Weron R., (2006), *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*, Wiley, New York

Weron R. and Misiorek A., (2008), *Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models*, International Journal of Forecasting, 24, pp. 744–763

Zarnikau J. and Lam A., (1998), *Market Concentration and Market Power Among Electricity Generators in Texas*, The Electricity Journal, November, pp.60–64