Carlo Pozzi

The Relationship between Spot and Forward Prices in Electricity Markets

Abstract

The functional relationship linking spot and forward power prices has been long debated. In this chapter, we rely on a modified interpretation of the storage theory and draw on an approximation of residual generation capacity in the German power system, to model the difference between future and spot prices (price *basis*) registered at the European Energy Exchange (EEX). We accommodate the various econometric specifications to three years of daily data time series. Statistical significance is achieved in all cases. Best results are obtained with an exponential GARCH estimation. Restated residual capacity is able to accurately drive the observed basis. This provides some evidence of the increasing rationality of power markets and their dependence on production and distribution constraints.

Keywords: adjusted basis, ARMA, auto-regressive error, backwardation, base load, blackout, commodity, conditional heteroskedasticity, contango, convenience yield, cost-of-carry, demand rigidity, disturbance generating process, EGARCH, error normality, forward price, futures, generalized estimation, GARCH, GMM, homoschedasticity, hydroelectric reserves, instrumental variable, inventories, Jarque-Bera test, opportunity cost, peak load, power exchange, power trading, price basis, price kurtosis, skewness, price spikes, price volatility, production capacity, reserve capacity, residual load, robust estimation, spot price, storability, storage theory, two stage least square, vector covariance, vector orthogonality, White heteroskedasticity tests, white noise.

The liberalization of the electricity sector has brought about new marketplaces where power can be traded in standardized form, in a manner similar to the way in which other traditional commodities like oil, ores or crops are traded. To cite a few, Nordpool, PJM, EEX or Powernext, are today familiar names for commodity traders in Europe and North America. They identify financial exchanges, matched to one or more power grids, where producers of electricity (or traders having access to production) can offer, for a fixed price, the supply of a predetermined amount of energy (usually measured in megawatts, MW) during one or more hours of the next day, while buyers (such as industrial consumers or local distribution companies) can bid for the purchase of an equal amount of energy, during the same time-slot.

According to the settlement model followed in each marketplace, bids and offers can be matched in diverse ways. In marketplaces where trades are organized on a continuous basis, bids and offers are paired on the spot. A bid can thus be placed for a time slot in the immediate future (like the next hour) and is settled—and a sale contract established—as soon as a seller makes available an offer (1) for an equal or lower price and (2) a corresponding amount of energy to be delivered during the same period. If, instead, trades are settled by auction, buyers and sellers must communicate their undisclosed bids and offers to the market authority generally one day ahead of their delivery time. Bids and offers are subsequently stacked according to their proposed prices, and different demand and supply schedules are built for every future time slot in which they are to be delivered (in descending and ascending order respectively). The intersection between each pair of schedules then yields the settlement price at which power will be exchanged in every next-day time period of reference. Accordingly, this price is taken as the performance basis for agents who are assigned contracts in the auction process.

This brief illustration provides some insight into *spot* power trading, specifically on the settlement mechanism of bids and offers placed for quasi-immediate delivery (from a few minutes to an entire day ahead). But power markets need not only accept the placement of bids and offers for very short maturities. Generators may in fact commit to provide power to their customers well ahead of when it is needed. Likewise, buyers can forecast their seasonal necessities by following their patterns of consumption and place bids accordingly. In several existing power exchanges, contracts for *forward* delivery have thus thrived, giving rise to *futures* markets where agents can trade electricity for short-to-medium maturities.

The coexistence of spot and forward power markets—markets where contracts for the supply of electricity in a given transmission region at a number of future times are simultaneously traded—requires the availability of *spot* and *forward* prices for a single megawatt-hour (MWh) to be consumed at a particular location in a power system. In this regard, power markets have thus developed similarly to other commodity markets, whose prices for immediate or future delivery have long been available to traders to guide their speculative activities. Yet, power prices seem to escape the application of the traditional asset-pricing relationships which are commonly employed to link spot and term prices in other commodity markets. It is indeed still largely unexplained why spot electricity prices may trade for some time below future prices, then suddenly soar well above the latter and reach levels several times in excess of their previous values. This limitation in many ways thwarts the liquid functioning of electricity exchanges. In the absence of well-understood pricing processes, market agents cannot easily perform inter-temporal trades and simultaneously contain the assumption of excessive price exposures. For many financial practitioners, power markets are thus somewhat problematic: whether to be safe and prudent rather than an eager speculator. On the other hand, for researchers, power exchanges offer an interesting area of investigation, where some advances are still to be made.

The non- or limited storability of electricity is often invoked to justify the lack of a well-defined relationship between spot and forward power prices. Electricity cannot be directly amassed in reserves (since existing batteries have very limited capacity) and is thus stored as potential energy through its means of production (water, coal, oil, natural gas and uranium). The *storage theory* (Kaldor, 1939; Working, 1948) illustrates why this may have a significant effect on power prices. According to the theory, firms trading storable commodities (hence not power) hold inventories in order to respond to unanticipated demand oscillations. This surely exposes them to storage and opportunity costs (greater working capital), but makes possible the selling of retained stocks when goods are most desired—a valuable advantage commonly called *convenience yield*. Therefore, in tight market conditions (when demand is high and commodity reserves scarce), traders dislike the postponed delivery associated with forward contracts and prefer to gain immediate possession of contracted goods. As a result, storage and opportunity costs become secondary, the convenience yield acquires a crucial importance, spot prices rise above forward prices, and the market is said to *backward*. Conversely, in loose market conditions, the opposite trend prevails: low demand and abundant inventories increase the importance of storage and opportunity costs, the convenience of having reserves is quasi-irrelevant, and spot prices quote below forward prices (*contango*).[1] With electricity, this is not easily observed. Power inventories have a blurred nature and very indistinct magnitude, so no reliable metric is available to functionally link the significant oscillations of the difference between future and spot power prices to power reserves.

However, if power reserves could be measured in an alternative fashion, it is, in principle, admissible that the storage theory could also have some explanatory power with reference

---

[1] An alternative explanation (Bresnahan and Spiller, 1986) relates the difference between spot and forward commodity prices to inventories via the implicit performance guarantee that reserves provide to firms that short their products in future markets. Market agents tend to anticipate purchases when a commodity is expected to be in scarce supply; therefore having reserves yields value which pushes spot prices above forward prices. This is consistent with the idea that a futures price may include two components: a forecast of the spot price and a risk premium that possibly depends on the risk preferences of market agents (Keynes, 1930). Fama and French (1987) provide an empirical comparison of the storage and the risk-premium theories using the prices of several traded commodities which affords greater significance to the first of these theories.

to electricity trading. In this chapter, using empirical data from the German power market (EEX), we test the hypothesis that an implicit measure of power reserves may provide a good foundation to explain the oscillations followed by the *basis*—the algebraic difference between forward and *adjusted* spot power prices. In order to do this, (1) we rely on timely measures of power load (expressed in MWh) as communicated by the four Transmission System Administrators (TSOs) that manage the entire German grid and (2) we extract from them a measure of available power reserves by proxy (hereinafter, *implicit reserves*) to which we econometrically link the simultaneous basis observed on the most liquid futures contract traded at the EEX. The remainder of the chapter is organized as follows. Section 9.2 illustrates the existing state of the research on the subject. Section 9.3 explains the explicit hypotheses which are subjected to econometric testing. Section 9.4 discusses the employed econometric methods. Section 9.5 describes the dataset under investigation. Section 9.6 presents the estimation results. Finally, Section 9.7 provides a discussion of the conclusions which can be drawn from this study.

§ 9.2 - Background: The Storage Theory and the Forward Price of Commodities

For the storage theory, the relationships linking forward and spot prices of a commodity can be derived from the *cash-and-carry* rationale. Agents agreeing to sell an asset at a future date may replicate (thus cover) their commitment by immediately buying and carrying until maturity what they will (then) need to deliver. In this way, they incur the opportunity cost of using liquidity to readily purchase the asset, but profit from the utility of possessing and being able to trade it until maturity.[2] Therefore, for these agents the return of buying a commodity today (that is, at time t) and delivering it at maturity (T) should at least be equal to,[3]

$$F\left(t,T\right)-S\left(t\right)=S\left(t\right)R\left(t,T\right)+W\left(t,T\right)-Y\left(t,T\right). \qquad [9.1]$$

Here, $F(t,T)$ represents the commodity forward price, $S(t)$ is the spot price, $S(t)R(t,T)$ is the opportunity cost of investing cash in a unit of commodity, $W(t,T)$ is the marginal cost of storing the commodity through the delivery period, and $Y(t,T)$ tracks the convenience yield of holding the asset. By moving the second term on the left-hand side of [9.1] to the right-hand side, a statement for the forward price of a commodity is obtained. This statement, in complete markets, yields the theoretical value at which commodity futures written on storable commodities for maturities equal to T should trade at t.

[2] According to financial theory, this functional dependence is a *no-arbitrage* relationship. By considering all costs and revenues involved in physically replicating the contractual obligation of delivering a commodity at a certain future maturity, the theoretical forward price of a commodity may be correctly assessed. In this manner, observed forward prices may differ from theoretical forward prices only randomly and, thus, differences do not provide market agents with arbitrage opportunities (profit occasions at no risk)..

[3] Here, forward prices are modeled through the formulation proposed by Fama and French, 1987.

As explained in the introduction, when commodity inventories become scarce, prices tend to back up. Therefore, with low inventories, the left-hand side of [9.1] becomes significantly negative and matches the growth of $Y(t,T)$ on the right-hand side; while the reverse is true with abundant inventories. The literature on the empirical estimation of this prediction is relatively extensive. With reference to surveys conducted on various types of storable commodities (either seasonal, like agricultural products or semi-processed foods, or non-seasonal, like metal ores or hydrocarbons), the difficulty of conducting econometric estimations generally lies in the problem of modelling storage costs and convenience yields. Researchers may not, in fact, find direct data for them and, thus, often need to resort to their implicit modelling. Nonetheless, studies have been able to significantly show that, as expected, convenience yields and the timely differences between forward and spot prices $[F(t,T) - S(t)]$ decrease when the inventory level of a commodity declines relative to its trading volumes. In fact, they have not only confirmed the basic insight of the storage theory, but have also shown that inventory levels drive the difference between forward and spot prices in a strongly non-linear fashion, since in all tested markets the absolute value of convenience yields sharply increases with the growing scarcity of reserves. For several examples, readers may refer to the studies of Working, 1948 and 1949; Telser, 1958; Fama and French, 1987 and 1988; Brennan, 1991; Deaton and Laroque, 1992; Ng and Pirrong 1994; and Pyndick, 1994.

Electricity, on the other hand, is patently non-storable. Hence, many deem that trying to use the storage theory to estimate power prices is nonsense. But this is perhaps an excessively exaggerated standpoint. In fact, since power can be stored in potential form, power inventories may possibly be tracked in some analogous form. For instance, in power systems where electricity, as a secondary form of energy, is mainly produced with hydroelectric reserves (as in the Nordic countries), researchers have discussed the relationship between water levels in hydraulic reservoirs and the forward-spot difference, finding similar (but less significant) evidence to that presented by the literature cited above (see Gjoilberg, 2001, and Botterud *et al.*, 2003). This seems to posit that estimating no-arbitrage statements akin to [9.2] on power prices may not be altogether futile. The challenge is to measure indirect power reserves in a way that validly approximates inventories as they are tracked in other commodity markets; particularly when water reserves are not available or just unimportant. The following section illustrates how we propose to tackle this task using German data.

§ 9.3 – Hypothesis: Power Implicit Reserves

In order to accumulate power reserves and be ready to respond to additional demand, power generators need to indirectly store electricity through its means of production: water, coal, oil, natural gas and uranium. In addition, they need to have production plants capable of processing more raw materials and generate more MW when they are needed. In developed economies, this ability to cope with greater demand must also be firm. Consumers in Europe and North America are, in fact, not used to power failures. They assume the continuous supply of electricity to their premises to be a basic right.[4] Hence, their unanticipated demand swings must be satisfied. This entails that -

(1) The maximum overall supply capacity in a western power system is greater than what is normally needed; this additional capacity is defined as *reserve capacity*;

(2) Power producers as a whole need to resort to primary energy reserves to supply more power when needed. Therefore, they respond with varying delays to additional demand, depending on several explanatory factors such as the production fuel, the generation technology, the location of the plant, the availability of transmission capacity, the atmospheric conditions, etc.

When more demand requires the injection of more supply into the system, reserve capacity gets eaten up. The use of additional capacity translates to additional supply with some delay. When present, the mechanism of balancing markets takes care of the very-short term re-equilibration of the system towards greater supply and this has a signalling effect. Bidders start to increase prices ahead of time in order to secure readily available output. Settlement prices jump above their normal levels and, the greater the reserve capacity to be used, the higher the pressure on prices to avoid blackouts, hence the higher the spikes in spot price processes.

Reserve capacity makes up for direct inventories and measures the ability of a power system to resort to primary sources of energy, in a timely manner, in order to cope with demand swings. Assuming that raw materials are available for production, the greater the level of reserve capacity with respect to the normal level of output, the lower the convenience it provides. Conversely, the lower the capacity to be set aside for use in normal circumstances, the higher the utility of possessing an extra MW to satisfy demand. We refer to *power implicit reserves* as the floating level of reserve capacity in a power system with respect to its normal level of supply and we formally define it as described below.

Since the maximum production capacity in a power system is a relatively stable measure (it basically represents the summation of the capacity of all existing and operating plants) and is probably never reached in actual terms, this datum can be approximated as the highest supply level attained over a sufficiently long period of time. Let us thus first define the timely level of residual production capacity in a power system as the difference between its maximum and timely levels over the time window (t-n, t-n+1, … , t-1, t) as,

$$RL(t) = \max_t \left[L(t) \right] - L(t), \qquad [9.2]$$

[4] In less developed economic systems—even if heavily industrialized as in China—blackouts are more common and accepted, to a degree, by industrial consumers and households.

where $L(t)$ represents the total load of electricity supplied in a power system at time t expressed in MW and $RL(t)$ stands for residual load.

It follows that *power implicit reserves*, $IR(t)$, can be defined as,

$$IR(t) = f\left(RL(t) - \frac{\sum_{t=1}^{t-n}\left[RL(t)\right]}{n}, RL(t) - \frac{\sum_{t=1}^{t-p}\left[RL(t)\right]}{p} \right). \qquad [9.3]$$

If time is measured in days, equation [9.3] hypothesizes that implicit reserves are a function of, (1) the conditional expectation of the residual load level (over the entire set of n days considered) and (2) some short-run mean specification (over a subset of p observations, with $p < n$). The idea is that market agents track inventory levels by looking at two pieces of information: the current residual capacity with respect to its normal level and the latest trend in its evolution (possibly on a weekly basis, ($p \leq 5$)). This provides insight both on long-run consumption intensity and on the immediate possibility to cope with demand oscillations and, hence, on the overall utility of possessing available residual capacity.

Now, ignoring marginal storage costs[5], equation [9.1] can be rewritten as,

$$F(t,T) - \left[S(t) + S(t)R(t,T)\right] = -Y(t,T) \quad . \qquad [9.4]$$

In the expression above, the square bracket on the left side represents the future value of the spot price on maturity. Using continuously compounded rates, equation [9.4] thus becomes,

$$F(t,T) - S(t)e^{r(T-t)} = -Y(t,T) \qquad [9.5]$$

where r is an approximation of the risk-free continuous rate. We define the left-hand side of [9.5] as *adjusted basis* (adjusted by the opportunity cost of capital) and we posit that this term represents the profit of having reserve capacity of power production. This profit is, therefore, a sort of convenience yield in electricity markets and may be significantly driven by implicit power reserves as tracked by equation [9.3].

---

[5] Power requires generators to build large facilities in order to store water or fuels. Within certain ranges, additional storage may actually have marginal costs close to zero, until the long-term investment of building a new facility needs to be undertaken.

§ 9.4 – Econometric Methodology

In order to test this hypothesis, an explanatory relation linking $[S(t) - S(t)e^{r(T-t)}]$ to $IR(t)$ should be estimated based on trading data. With electricity prices, this poses some methodological complications. First, the hypothesis that residual production capacity drives forward-spot price differentials may not be thoroughly accommodated by the way $IR(t)$ are modelled by [9.3]. Equation [9.3] indeed tries to provide an intuitive specification of implicit reserves. But price time series may have complex lag structures in their functional dependence on exogenous drivers. Buyers and sellers of electricity in a competitive market may, in fact, use information they learn at different points in time to orient their exchange activities. This implies that, most likely, significant serial correlation will affect estimation residuals after simple regression models are initially fit to price and load data.[6] Second, the rigidity of power demand, paired with the impossibility of directly storing it, causes power prices to oscillate greatly when consumption surges unexpectedly. Spot power price time series are, in fact, characterized by the periodic observation of high positive jumps followed by immediate negative jumps (that is, *spikes*), which tend to cluster in times of market crisis. Hence, on the one hand, numerous price spikes confer significant non-normality to power price data.[7] On the other hand, they also cause large estimation errors (which also

---

[6] Serial correlation means that estimated errors are correlated with each other. In OLS regression, besides having zero expected value, disturbances should be assumed, (1) to have constant variance and (2) to be serially uncorrelated. Let us assume that we have fitted to a dataset a regression model of the general form $\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$, where $\mathbf{y}$ is a vector of n dependent variable observations, $\mathbf{X}$ is a $n\times k$ matrix of k regressors, $\boldsymbol{\beta}$ are k estimated regressor coefficients, and $\boldsymbol{\varepsilon}$ is a vector of n estimation errors. The first assumption—constant variance, or $\mathrm{Var}[\boldsymbol{\varepsilon}|\mathbf{X}]=\sigma^2$—is generally referred as homoschedasticity, a word whose literal meaning can be expressed as a situation in which errors present stable scale. For instance, in the case of univariate OLS models, homoschedasticity obtains if actual pairs of dependent and independent observations ($y_i\in\mathbf{y}$ and $x_i\in\mathbf{x}$) will plot at similarly scaled distances from the regression line, so that the variance of the error will not show a consistently growing or diminishing value. The opposite of homoschedasticity is heteroskedasticity, a situation in which errors show some kind of growing or diminishing schedule as one moves along the observations (that is, as one moves along the regression line). The second assumption—non-autocorrelation, or $\mathrm{Cov}[\varepsilon_i,\varepsilon_j|\mathbf{X}]=0$, . $i\neq j$—implies that disturbances do not reveal any sort of functional dependence between them. Hence, if errors are uncorrelated, it is not possible to extract any information about one error term from previous observations. In a multivariate regression setting (that is, when $\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$) both conditions (homoskedasticity and non-correlation) can be summarized by defining a matrix of error variances and covariances. By recalling that the covariance of a variable with itself is the variance of a variable, this matrix is generally indicated with $\boldsymbol{\Sigma}$ and assumes the form $\boldsymbol{\Sigma}\equiv E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\mathbf{X}]=\sigma^2\mathbf{I}$, where $\mathbf{I}$ is an identity matrix. With this notation, it is indicated that residuals from the estimation are not correlated between them ($\mathrm{E}[\varepsilon_1,\varepsilon_2|\mathbf{X}]=0$, $\mathrm{E}[\varepsilon_1,\varepsilon_3|\mathbf{X}]=0$, … , $\mathrm{E}[\varepsilon_1,\varepsilon_n|\mathbf{X}]=0$), while their variance ($\mathrm{E}[\varepsilon_1,\varepsilon_1|\mathbf{X}]= \mathrm{E}[\varepsilon_2,\varepsilon_2|\mathbf{X}]=\ldots=\mathrm{E}[\varepsilon_2,\varepsilon_2|\mathbf{X}]=\sigma^2$) is constant. In this case errors are also said to be *spherical*. With *autocorrelation* the second assumption is violated, hence, $\boldsymbol{\Sigma}\equiv \mathrm{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\mathbf{X}]\neq\sigma^2\mathbf{I}$ and $\boldsymbol{\Sigma}$ is thus a matrix of error variances and covariances where terms off the central diagonal have non-zero values.

[7] Price spikes can be seen as observations significantly off the conditional price mean over the entire sample of n power prices. When a distribution accommodates numerous extreme values, its bell-shaped curve has relatively fat tails. It is then said to be leptokurtic. Comparing, for instance, the spot price time series of different trading assets over the same period provides a vivid image of this fact. With reference to three North American price indexes—the Standard & Poor 500 (an equity index), the West Texas Int. oil FOB

concentrate in time), when estimation models are fit to actual datasets. This fact generates, in turn, significant heteroskedasticity in cross-sectional disturbances.[8]

Now, normality in regression estimation errors is the fundamental assumption to derive the properties and the statistical significance of ordinary least square (OLS) estimators. The same holds true for the assumption of their non-autocorrelation and homoskedasticity.[9] With power prices, the validity of OLS estimations may, therefore, be seriously limited. This requires us to tackle the problem of estimating a functional relationship between [$S(t) - S(t)e^{r(T-t)}$] and $IR(t)$ by using different techniques. Let us review the viable alternatives.

The presence of serial correlation in disturbances after fitting a simple OLS regression on data between $-Y(t)$ and $IR(t)$, may require us to find alternative ways to model the independent variable, so as to mimic the possible trading behaviour of market agents, given their information on power load data. This can be done, (1) by relaxing the way equation [9.3] models residual power loads $RL(t)$ and (2) with the express insertion of auto-regressive (AR) and/or moving average (MA) terms in the model specification, in order to more accurately capture the relationship between past observations of $RL(t)$ in the generating process of $IR(t)$ and current observations of $-Y(t)$. Therefore, an ARMA model, whose specification will be guided by the measurement of partial serial correlation statistics between error terms (see Section 9.6 below) provides a first methodological improvement.[10]

---

price (a storable commodity index) and the New England Pool Daily Mean Power price (a non-storable commodity index)—spot prices standardized by their minimum value ($S(t)$ /min[$S(t)$]) and tracked over a three-year period (January 2000/December 2002) present the following descriptive statistics, the informational content of which is self-evident (the kurtosis of a perfectly normal distribution is equal to 3).

Index	Mean	Stand. Dev.	Skewness	Kurtosis
Equities	1.55	0.25	-0.21	2.01
Oil	1.57	0.22	-0.39	3.16
Power	7.94	8.88	21.52	534.30

[8] Heteroskedasticity occurs when observations on the central diagonal of the variance-covariance matrix of estimated errors ($\mathbf{\Sigma} \equiv E[\boldsymbol{\varepsilon\varepsilon}'|\mathbf{X}]$) are different from σ^2, so the scale of estimation errors is not constant (see note 6 above)

[9] As mentioned above, disturbances are spherical when their matrix of variance-covariance is $\mathbf{\Sigma} \equiv E[\boldsymbol{\varepsilon\varepsilon}'|\mathbf{X}] = \sigma^2\mathbf{I}$. Therefore, disturbances are non-spherical when their matrix of variance-covariance is $\mathbf{\Sigma} \equiv E[\boldsymbol{\varepsilon\varepsilon}'|\mathbf{X}] = \sigma^2\mathbf{\Omega}$, where $\mathbf{\Omega}$ is another matrix of some known or unknown form which differs from $\mathbf{I}$.

[10] ARMA models take care of the task of capturing serial correlation in estimation errors and use this piece of information as an explanation of the dependant variable. Intuitively, once they are fit on a dataset, the original serial correlation in OLS residual errors should get significantly reduced. In the absence of a theory to guide the definition of the explanatory structure of residuals to be inserted in the estimation (*i.e.*, whether one, two, three, etc., days of lagged errors should be considered), it is advisable to use a trial-and-error approach. In this case, one can use partial autocorrelation statistics (see Section 6) to learn which OLS lagged estimation errors have more predicting power for the last error in the time series. Accordingly, the higher the partial correlation, the greater the explanatory power on the dependent variable; hence, it is better to include AR and/or MA terms in the estimated model specification, until the serial correlation observed across residuals is eliminated to the maximum possible extent.

The simultaneous (and interrelated) presence of heteroskedasticity and non-normality in estimation errors may then suggest a second direction in which further estimation improvements may be secured. EViews—the econometric software application used in this work—supports, in fact, *robust estimation* methods to account for heteroskedasticity in the fitting of the ARMA model to a dataset. In this way, regression coefficients linking reserve load observations to the adjusted basis can be corrected to consider the varying scale of estimation errors.

However, if after an ARMA model is specified and *robustly* estimated, disturbances still remain highly non-normal, the most credible hypothesis is that the econometric estimation is not thoroughly able to cope with the occurrence of high power price spikes. Estimation errors may, in fact, be large when spot prices jump well above or below forward prices, sending $|-Y(t)|$ to very extreme levels. If this is the case, given the chosen ARMA specification, it means that some correlation between the chosen regressors (the explanatory variables) and the estimated disturbances (after the ARMA structure has been considered) still exists. In this instance, it may be possible to try to further improve estimation with the support of *instrumental variables*. Instrumental variables are a set of alternative regressors that enter the estimation model instead of the original ones. A correct identification of instrumental variables requires them to be significantly correlated with the original regressors, but not with estimation disturbances (in other words, they should therefore respect the *orthogonality* condition with respect to the disturbance vector)[11]. Therefore, if a set of instrumental variables is available, it can be profitably employed in estimation techniques like the *two stage least squares* (TSLS) and the *generalized method of moments* (GMM) that may afford some better results.

An alternative and, possibly, more powerful approach is to simultaneously take care of heteroskedasticity and non-normality in estimation errors (due to price spikes), by using a *generalized auto-regressive conditional heteroskedastic* (GARCH) model. This type of approach relies, in fact, on the separate estimation of two regression equations—a *mean* and a *variance* equation—which take into account both the conditional mean and conditional variance of estimation errors. Specifically, the *mean equation* regresses the dependent variable on a set of exogenous variables not differently from what is done in most regression techniques. Hence, in our case, it fits the adjusted basis, $-Y(t)$, on present and past reserve load data, using an ARMA as specification seen above. Whereas the *variance equation* just models the estimation error in the first equation by treating its variance as a dependent variable of two separate terms: (1) the square of one or more estimation errors at different lags from time t, and (2) the variance of the same lagged errors, up to a different delay order. To reiterate, the variance of the error in the estimation of the first equation at time t is treated as a dependent variable of (1) squared errors observed at previous times, between $t-1$ and $t-p$, and (2) their observed variance over a

[11]Two vectors are said to be orthogonal when their product equals zero. Orthogonal vectors are therefore completely uncorrelated.

different time window, between $t-1$ and $t-q$. In this manner, GARCH models allow for explaining the concentration of price spikes in times of high market volatility (which necessarily translates into a high variance of estimation errors) as a function of the quadratic increase of past errors and of their variance. Therefore, this type of model is able to anticipate times of large price swings—that is, times of large estimation errors in the mean equation—by exploiting the tendency of power price spikes to cluster over time. The implication is that GARCH models may possibly provide the best way to normalize, to the highest possible extent, the distribution of estimation errors after all measurable causes of price spikes have been accounted for.[12]

§ 9.5 – Dataset: Power Prices at the EEX Market and Power Load in the German System

In this study we focus on a continental European market, the German power exchange (EEX). In this competitive arena, almost three years of daily price observations and intra-daily power load and consumption data are available. This, combined with the acceptable (though still limited) liquidity of the EEX power exchange and the availability in it of a consistent array of financial forward contracts, provide good grounds for empirical testing.

Located in Leipzig, the EEX market—European Energy Exchange is the result of the merger in 2002 of two separate marketplaces: the Leipzig Power Exchange, originally also located in Leipzig, and the European Energy Exchange, located in Frankfurt. This market looks forward to becoming the trading place of election for power in Central Europe and, for the moment, can be viably matched to the overall power system administered by the four German TSOs: EnBW, EON, RWE and Vattenfal.

Spot trading is available at the EEX both on a continuous and auction basis, with the latter market making up the bulk of trading volume.[13] Every day, two single weighted average price indexes—the *Phelix Base* and *Phelix Peak*—representing that days spot prices during two different time windows, are determined on the basis of 24 hourly prices. These two time windows—the base-load window and peak-load window, from hour 1 through 24 and

---

[12] Formally, GARCH(p,q) models (where the Roman letters in the round brackets define the lag orders of the errors they consider) assume that errors in the mean equation respond to a generation process that is a function of their past values (up to p) and variances (up to q). This implies that $\sigma_t^2 \equiv \mathrm{Var}[\varepsilon_t \in \boldsymbol{\varepsilon}]$ in $\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ (the mean equation), with $\boldsymbol{\varepsilon}=(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-n})$, is written (the variance equation) as follows,

$$\sigma_t^2 = \omega + \sum_{j=1}^{p} \gamma_j \varepsilon_{t-j}^2 + \sum_{i=1}^{q} \gamma_i \sigma_{t-i}^2 \cdot$$

[13] According to EEX data, throughout the first five months of 2005, continuous trading has reported actual trading volumes only in 37 out 138 business days. Mean volume exchanged has been for continuous and auction trading of 1250.3 MWh and 219,032.9 MWh, respectively. The auction market has been, on average, 175 times more liquid than the continuous one.

hour 9 through 20, respectively—are defined according to normal patterns of consumption. These price indexes are taken as a settlement reference for their respective futures contracts. Futures contracts (base-load and peak-load) at the EEX are then available for numerous increasing monthly, quarterly and yearly maturities. For instance, traded base-load monthly futures for which an open interest existed in MWh on 2 May 2005, were for deliveries through May, June, July, August, September, and October 2005; quarterly futures were for the July-2005 through the January-2007 quarters; and yearly futures went up to 2011. All of these mentioned contracts are to be settled in cash against Phelix base-load prices reported through their respective delivery periods. In fact, for standardized forward power contracts like these futures, delivery is over an entire period of time, not at a single date. Hence, the performance of futures begins upon maturity, which is the beginning of the delivery period, and ends with the end of the delivery period. For instance, according to EEX trading rules, a monthly future for delivery in June 2003, traded on 9 May 2003, has 20 days of residual trading and will be performed, and thus cash settled, through that entire month of June.

Various futures have diverse liquidity. Base-load contracts are more liquid than peak-load futures. Among the former, monthly contracts are more traded than quarterly contracts, which in turn are more numerous than yearly ones. Among monthly contracts, the most traded is the one which is to be delivered during the month that follows the month to which a current trading day belongs. Given its higher liquidity, it may be conjectured that this contract presents better pricing data; hence it provides a more adequate testing dataset. Accordingly, we test the hypothesis spelled out in Section 9.3 on its price time series.

In order to perform econometric investigations, future prices must be juxtaposed on spot prices. By using the Greek letter τ to designate the beginning of the maturity period for the one-month base-load futures mentioned above, we indicate with $F_t(\tau,T)$ the future price traded at t for the one-month ahead delivery period (τ, τ+1, …, T).[14] This price can be compared to the Phelix base-load daily mean in t, $S(t)$. Likewise, $F_{t-1}(\tau,T)$ can be compared to $S(t-1)$; $F_{t-2}(\tau,T)$ to $S(t-2)$, and so forth, so that two time series of prices are built backwards to $t-n$. Note that, going from t to $t-n$ over a time set in excess of one month, entails periodically rolling back the beginning and the end of the maturity periods (τ,T) for the tracked futures and choosing forward prices accordingly.[15] (τ,T) are thus also variable dates which are a scaled function of t. To avoid clumsiness, we do not represent this in the (τ,T) notation. However, we employ an algorithm to select, among all available future prices, the one for the contract which is for delivery in the month subsequent to which each trading day in the ($t-n$, $t-n+1$, …, $t-1$, t) set belongs. Since future prices are available at EEX on each working day from Monday through Friday (not for weekends), this procedures yields a dataset (selected after the merger of the power exchange in Leipzig

[14] So, for instance, if t is 9 May 2005, τ is 1 June 2005 and T is 30 June 30 2005.

[15] In other words, starting from 9 May 2005 and going backwards, requires tracking the future price of the [τ=1 June 2005/ T=30 June 2005] futures contract when t belongs to May 2005, the price of the [τ= 1 May 2005/ T=31 May 2005] futures when t belongs to April 2005, etc.

with the one in Frankfurt) of 560 pairs of forward-spot prices, between 3 January 2003, and 25 April 25 2005.

Equation [9.5] requires then that the basis be determined after spot prices are adjusted for their opportunity cost of capital until delivery. This entails determining $S(t)e^{r(T-t)}$ in steps as follows. (1) $S(t)$ is directly observed and does not pose any problem. (2) Each $S(t)$ is then multiplied by an exponential function of r for the $(T–t)$ period that includes a variable number of days to be split in two time slots: $(T–\tau)$, which is always one month and is approximated with the median value of a fortnight; $(\tau–t)$, that, given EEX trading rules, can go from a minimum of three days to a maximum of a month, and is directly determined on t. (3) Finally, continuous-time risk-free rates, r, are approximated with the most appropriate (given $(T–t)$) discrete-time Euribor rate in the weekly-to-sixty-day maturity term-structure, subsequently converted into its continuously-compounded equivalent.

Once price data have been opportunely treated as illustrated in these paragraphs, a time series of adjusted bases $-Y_{t-i}(t,T)$, with $i \in (n, \ldots, 0)$, is obtained. This time series is the dependent variable which, in our tests, must be regressed on a measure of power implicit reserves—the independent variable—as defined in Section 9.4.

To model implicit reserves, we track the evolution of power loads in the German grid so as to determine maximum and retained production capacities. All of the four German TSOs administering the national power grid release historical data on the total amount of power in MW they injected in the system every quarter hour, since June 2003.[16] This information is available online directly from their websites. The summation of each TSO's load provides the German national load. The arithmetic mean of national load across the 96 slots of 15 minutes that make up a base-load day (as set to determine the Phelix price basis), averaged across the four TSOs, provides the daily mean load in the whole German grid (previously indicated as $L(t)$). The maximum load over the time series of daily loads between 1 June 2003 and 25 April 2005 provides—according to [9.2]—the foundation to determine the corresponding time series of daily residual loads, $RL(t)$. This series is thus obtained under the assumption that the maximum observed load over the sampled period represents a quasi-complete utilization of production capacity. As a result, power loads treated in this manner define a (normally weekly) pattern of capacity utilization. This time series,

[16] In these time series of data, some arrays of observations are missing for reasons unspecified by TSOs. Missing data occur during time slots that go from a few quarters of an hour up to a few days. In these cases, given the weekly and hourly pattern of German power consumption, missing data for one TSO have been determined by tracking that TSO's loads in the same weekdays and hours of the previous week. The previous week's loads have then been opportunely adjusted by the mean ratio between the loads in the same missing hours as communicated by the other three TSOs and their loads in the corresponding hours of the previous week.

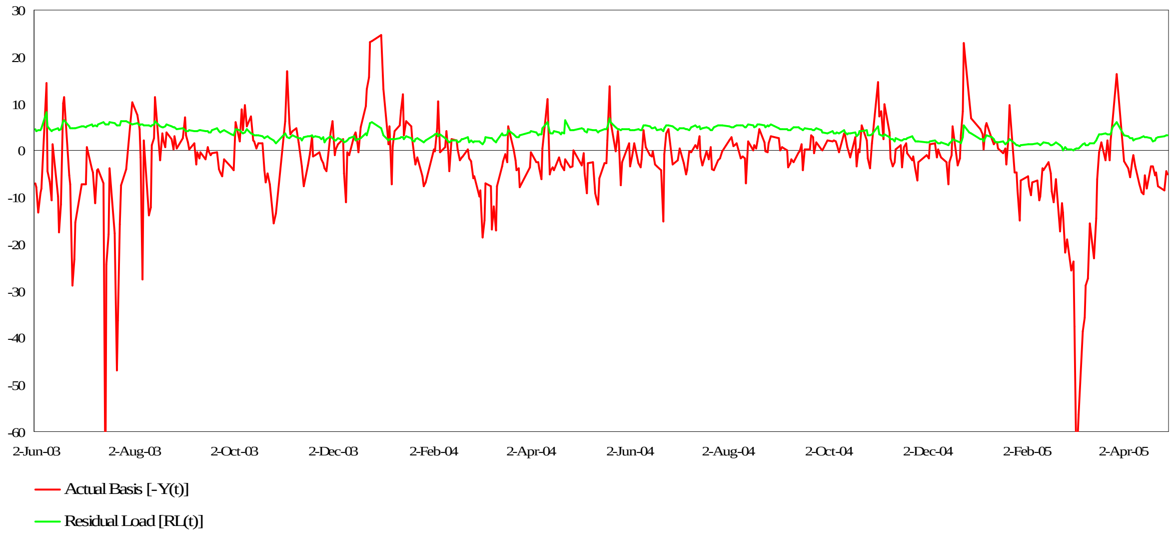
correctly modelled according to [9.3], finally yields the explanatory variable that hypothetically guides the adjusted basis of power prices over the entire sampled period.[17]

§ 9.6 – Estimation and Results

§9. 6.1 – Basic OLS Estimation

The research objective of this paper is to verify the hypothesis that, given the reduced volatility of future prices relative to spot prices, the latter tend to move away from the former (hence, the basis assumes oscillating values) according to some function of the residual capacity that, in a power system, is available to satisfy demand. Figure 9.1 below presents, therefore, a preliminary graphic comparison between the independent variable $RL(t)$ (residual capacity measured in GW of residual load) and the dependent variable $-Y(t)$ (the adjusted basis measured in € per MW).

Figure 9.1 – Adjusted Basis vs. Residual Load



This basic association does not really suggest a functional dependence linking the two variables, although some slight similarities between their trajectories may be at times observed. However, a simple OLS regression of the adjusted basis on a log-restatement of equation [9.3] already yields some interesting results, provided that the whole dataset is divided into single working days, and five estimations per each working day (from Monday through Friday) are separately conducted.

[17] Residual load values are determined with the exclusion of Saturdays and Sundays for which forward prices, hence basis values, are not available.

Here, implicit reserves are simply modelled as,

$$IR(t)=\left\{rl(t)-\frac{\sum_{t=1}^{t-n}\left[rl(t)\right]}{n}\right\}+\left\{rl(t)-\frac{\sum_{t=1}^{t-p}\left[rl(t)\right]}{p}\right\} \quad [9.6]$$

(where $rl(t)=\lambda\ln[RL(t)]$). The estimated model at this preliminary stage of investigation is possibly the most streamlined,

$$-Y(t)=\alpha+\beta IR(t)+\varepsilon(t). \quad [9.\ 7]$$

In it, $IR(t)$ are fed to equation [9.7] and determined as in [9.6], with $p=7$ and $\lambda=10$ for estimation optimization. Estimation statistics are apparently relatively good for all days, with all regression coefficients significantly different from zero at least at the 95 per cent level. Table 9.1 provides some highlights (there are 92 observations per day).

Table 9.1 – OLS Statistics for Single Business Day Estimations

Statistic	Monday	Tuesday	Wednesday	Thursday	Friday
α	-2.147108	-3.565482	-3.484542	-3.425387	-1.396168
t(α)	-3.298438	-3.082761	-4.383150	-4.091367	-2.107090
Prob. t(α)	0.0014	0.0027	0.0000	0.0001	0.0379
β	0.672074	0.117438	0.416443	0.975429	0.821897
t(β)	7.678897	2.081058	4.959429	8.538555	9.532839
Prob. t(β)	0.0000	0.0403	0.0000	0.0000	0.0000
Adjusted R^2	0.389120	0.035310	0.205906	0.441399	0.496890
Durbin-Watson	1.349961	1.281113	1.190776	1.034494	1.559369

However, the Durbin-Watson statistic presented in Table 9.1 is quite bad in all cases, since, with perfectly uncorrelated residuals, this should have a value of two.[18] The Ljung-Box *Q*-statistics and the Breusch-Godfrey LM test further confirm this fact. In both tests, and for all trading days, the probabilities associated with the *Q*-statistics and the χ^2 distribution of the Breusch-Godfrey's $N\times R^2$ statistic (not reported here) reveal a significant autocorrelation in the residuals, at multiple lags. As anticipated in Section 9.4, serial correlation in estimated residuals strongly biases OLS regression coefficients (α,β in equation 7 above) and suggests that we employ more involved methodologies that account for its presence in the estimation. Moreover, OLS residuals from these regressions are also

[18] See footnote 6 for a discussion on serial correlation in estimation residuals.

plagued by the presence of significant heteroskedasticity.[19] This problem is also discussed and tackled in the following subsections.

§ 9.6.2 – ARMA Specification

Serial correlation suggests directly inserting lagged error terms as regressors in the econometric specification to be tested.[20] For this reason, using Ljung-Box *Q*-statistics to target significant lagged disturbance terms, we fit an ARMA (*auto-regressive*, *moving-average*) model directly to residual load values, *RL*(*t*), as determined in [9.2]. The estimation of the ARMA specification below is now conducted on a single sample comprising all business days.

$$-Y(t) = \alpha + \beta^1 RL(t) + \beta^2 u(t-1) + \beta^3 u(t-4) + \beta^4 \eta(t-5) + \varepsilon(t). \qquad [9.8]$$

In equation [9.8], $u(t)$ are AR terms, while $\eta(t)$ is an MA term. Their lag order is specified in parenthesis (in days) and, according to theory, coefficients on AR terms are estimated after the intercept and the explanatory variable have been accounted for, while the MA coefficient is estimated on errors after intercept, explanatory variable and AR terms have been accounted for.[21]

Table 9.2 – ARMA Estimation Statistics

Statistic	α	β^1	β^2	β^3	β^4
Value	-20.23092824	5.036562971	0.6760196261	0.1627462608	0.1220492011
t	-7.658950	10.00908	18.56121	4.247654	2.434448
Prob. t	0.0000	0.0000	0.0000	0.0000	0.0153
F	179.0564				
Prob. F	0.000000				
Adjusted R^2	0.610710				
Durbin-Watson	2.161843				

This ARMA specification captures market pricing patterns within one week of trading and has the highest overall significance among all specifications satisfying the hypothesis

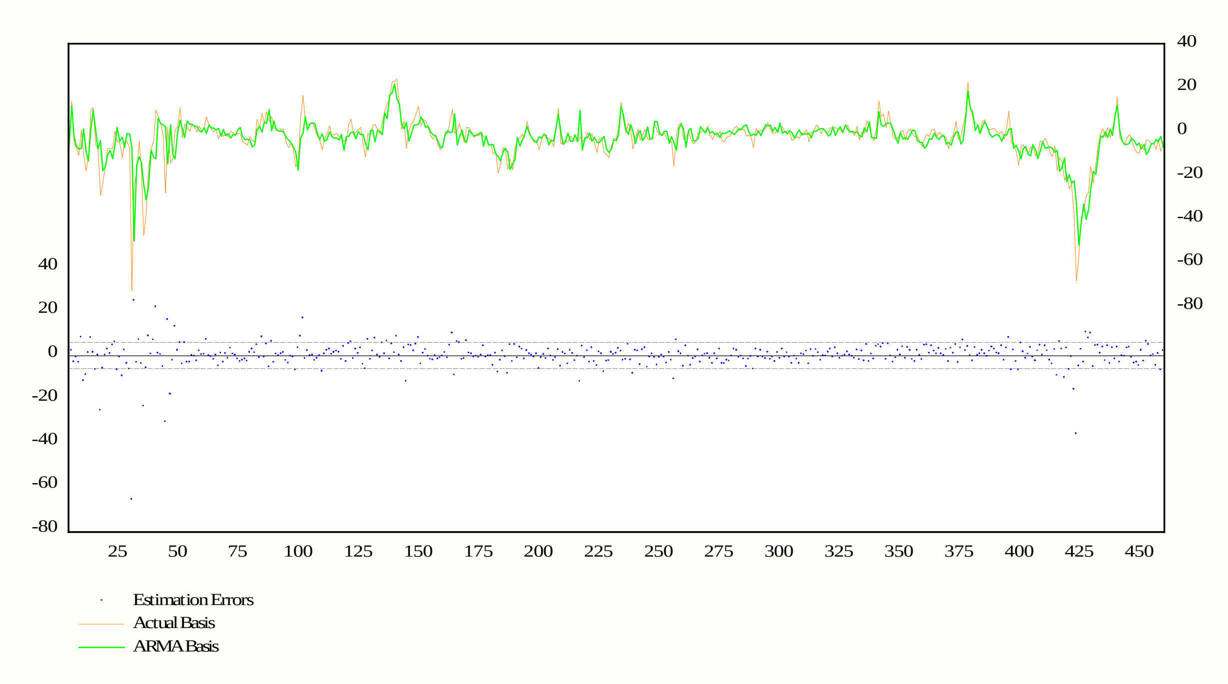
[19] White heteroskedasticity tests conducted on OLS residuals of all regressions considered in Table 1 reject the hypothesis of no heteroskedasticity with high significance in all cases.

[20] The fact that OLS lagged disturbances have an explanatory role for the dependent variable is tantamount to including lagged values of the independent variable in the regression. See the following note.

[21] More precisely, a general AR(1) model is specified as ($y_t = x_t'\beta + u_t$), with ($u_t = \gamma u_{t-1} + \varepsilon_t$). In these expressions, x_t and β are the explanatory variable and its parameter; u_{t-1} is a disturbance term (structural error) directly inserted in the estimation so that, by substitution, the general AR(1) specification becomes non linear [$y_t = \gamma y_{t-1} + (x_t - \gamma x_{t-1})'\beta + \varepsilon_t$]; γ is the parameter of the AR term; and ε_t is an innovation. Therefore, as explained in the text, innovations (ε_t in the equation just discussed), at the appropriate lag, represent the MA term $\eta(t)$ considered in [9.8].

spelled out in Section 9.3 (the second absolute highest among all tested specifications).[22] EViews estimates α and $\beta^{(\cdot)}$ by using non-linear regression techniques and the estimation output is summarized in Table 9.2.[23] As can be seen, the first auto-regressive term has the highest significance in the model (above the 99 per cent level), while all other terms are significant at least above the 95 per cent level (all AR and MA inverted roots are also comfortably within the unit root circle).

Figure 9.2 – Adjusted Basis vs. ARMA Modelled Residual Load



[22] Using the partial correlation statistics and reiterating the comparison between structured residuals (residuals computed after the sole role of $RL(t)$ on $-Y(t)$ has been considered) and residuals after a whole ARMA specification has been adapted to data, it is indeed possible to identify at least another (slightly) more significant ARMA specification, using higher order MA terms. In this case however, the estimated structure does not fully comply with the weekly pattern of trading followed in the EEX power market. It is therefore difficult to support a rational economic explanation for it, which is why this specification is not presented here.

[23] According to the software producer, EViews uses the Marquardt algorithm to estimate ARMA models. This is a modification of the Gauss-Newton approach. The latter converts the general non-linear model specification [$y_t = h(x_t, \beta) + \varepsilon_t$] into a linear one through subsequent derivations, and finds the true parameter for β using OLS error minimization. See Greene (2003). Here, such an initial non-linear specification is necessary because the application replaces lagged disturbance terms in [8] with the corresponding lagged values of the dependent and the independent variables (as explained in footnote 21). It thus re-expresses [9.8] in a non-linear form and then estimates its parameters.

Figure 9.2 above elucidates the graphic comparison between actual basis values, –*Y*(*t*), and fitted values obtained by using the right-hand side of [9.8] (except the error term). The graph contains two parts. The upper curves show how the model manages to replicate the trajectory followed by the actual basis. The fit is graphically good, although the model appears to cope with innovations with delay. The bottom part shows the plot of ARMA residuals. Note that they tend to increase when innovations are large (that is, on or around price spikes).[24]

Note that the Durbin-Watson statistic provided in Table 9.2 now has a value much closer to two. This suggests that, (1) serial correlation of residuals is relatively small after fitting the ARMA specification in [9.8], and (2) no major terms have been forgotten in the estimation. The other two serial correlation tests mentioned in the previous subsection also confirm this fact.

On the other hand, the White test does not reject the presence of heteroskedasticity among ARMA residuals.[25] EViews allows for improving ARMA estimations in the presence of such a drawback by supporting *robust estimation* through the White estimator—which is a heteroskedastic consistent estimator—for the same model specification. Unfortunately, this additional technique does not really improve estimation results and this suggests tackling the problem of heteroskedasticity in a more direct way. This is done with GARCH estimation at the end of this section.

§9. 6.3 – Generalized Estimation

In addition to heteroskedasticity, ARMA estimation residuals still reveal another drawback. The lower part of Figure 9.2 shows that numerous estimation errors have large values (which plot outside the jagged confidence lines). This gives significant kurtosis (39.86) to their distribution. Accordingly, the Jarque-Bera test—a test which controls the normality of the distribution of estimated residuals—applied to the whole set of *n* ARMA errors, rejects normality with high power.[26] From a graphic inspection of Figure 9.2, it seems that regressors in the ARMA model are partially unable to capture large positive price spikes when or before they occur. This inability generates large errors when power prices jump and may cause the regressors to co-vary with estimation residuals, thus violating the

---

[24] The skewness of ARMA residuals is negative. Errors therefore tend to be more negative than positive. This indicates that errors are larger and/or more numerous when the basis plummets, *i.e.*, when spot prices mark positive spikes.

[25] The statistics for this test are 5.089417 and 10.02073 for the F-statistic and the $N \times R^2$ value, respectively; with associated probabilities of 0.006520 and 0.006668 which confirm heteroskedasticity at the 99 per cent level. Heteroskedasticity given by the exogenous variable, *RL*(*t*), is significant both with respect to its direct and quadratic values.

[26] The Jarcque-Bera statistic, which is distributed as a χ^2, has, in this case, a value of 26,779.12 and rejects normality above the 99 per cent confidence level.

underlying assumption of exogenously chosen explanatory variables which accompanies all regression estimations.

Controlling whether there exists some covariance between each of the regressors in [9.8] and the ARMA disturbance error vector, actually confirms some lack of independence between them.[27] Hence, using an estimation method that generalizes the disturbance generating process in the variance-covariance matrix of the residuals (thus excluding normality) can possibly provide some improvement. But in order to do this, it is first necessary to identify a set of instrumental variables correlated to regressors in the original specification, but uncorrelated to the ARMA error vector (that is, variables which are orthogonal to errors).

EViews automatically uses lagged values of the original vectors of regressors in order to create a set of instrumental variables and to avoid the under-identification of a generalized estimation.[28] Therefore, it is sufficient to find one or more vectors of instrumental variables for the exogenous regressor in [9.8] so as to make the estimation possible. To do this, we use an algorithm to simulate vectors of values with zero covariance with ARMA errors and pre-defined covariance with regressors.[29] Once this is done, we introduce the appropriate instrumental variables into the estimation.

Eview supports a TSLS estimation for the same ARMA specification presented above. Unfortunately, TSLS estimation does not provide any significant improvement to the results presented in Table 9.2.

However, with a slight modification of the ARMA specification in [9.8] into the following AR model,

$$-Y(t)=\alpha+\beta^{1}RL(t)+\beta^{2}u(t-1)+\beta^{3}u(t-2)+\beta^{4}u(t-4)+\beta^{5}u(t-5)+\varepsilon(t) \quad [9.9]$$

[27] The perfect absence of correlation would mean that, for a set of ARMA regressors belonging to **X**, Cov[**X**,ε]=**0**, which is not the present case.

[28] An under-identified generalized model that makes use of instrumental variables is one in which the number of instrumental variable vectors is less than the number of parameters to be estimated (that is 5 parameters in equation [9.8]). Over-identification occurs instead when instrumental variable vectors are greater than the parameters to be estimated.

[29] Using the same estimation results presented in Table 2, the covariance of $RL(t)$ in [9.8] with the fitted basis is Cov[$RL(t)$,-$Y(t)$]=63.55. Specifically, we set an algorithm that, through randomization of log-values of $RL(t)$, finds j instrumental variable vectors, $\mathbf{IV}(t)=(IV_1(t), IV_2(t), \ldots, IV_j(t))$, for which Cov[$\mathbf{IV}(t)$,$\varepsilon(t)$]=0 is verified, and the covariance with $RL(t)$ is equal to Cov[$\mathbf{IV}(t)$,$RL(t)$]=$\theta\times$63.55, where θ assumes values between zero and two. Best weighted results between normality in residuals and goodness of fit in the estimation (R^2) are achieved with one vector of instrumental variables and θ set around unit values.

it is possible to estimate an over-identified GMM model.[30] Results for this estimation are presented in Table 9.3 below.

Table 9.3 – GMM Estimation Statistics

Statistic	α	β^1	β^2	β^3	β^4	β^4
Value	-32.57230	9.037821	0.551878	0.126734	0.155170	0.125565
t	-4.885180	4.682371	8.725179	2.411225	2.468386	2.583626
Prob. t	0.0000	0.0000	0.0000	0.0163	0.0140	0.0101
J-statistic	0.006479					
Adjusted R^2	0.574862					
Durbin-Watson	1.949849					

With this GMM estimation,[31] while the significance of regressors and the goodness of fit is not perceptibly lost (and serial correlation not introduced), some improvements in the normality of residuals are possible, after the AR specification in [9.9] is accommodated to the dataset. Their kurtosis diminishes to 34.01 and the Jarque-Bera test, while still rejecting normality, has a better statistic.[32] This is, however, a small amelioration that does not significantly change the ability of the model to replicate actual basis trajectories.

§ 9.6.4 – GARCH Estimation

Note that, so far, heteroskedasticity in residuals (detected in Sub-section 9.6.2) has not been directly tackled. However, since GARCH models provide the possibility to model the variance of residuals in the estimation, by leveraging on the linkage between the latter and lagged error information, it may be possible to better capture the local error variability generated by the concentrated occurrence of price spikes.[33] To do this, we begin by estimating an exponential GARCH(1,1) model (which we call EGARCH) with the same ARMA specification used in [9.8].[34]

The choice of an EGARCH(1,1) responds to the possibility of modelling both in an asymmetric and exponential way the effect of volatility on the conditional variance $\sigma^2(t)$. As previously specified (see footnote 12), in plain GARCH(1,1) models, the conditional

[30] Eviews does not support the estimation of MA terms in GMM estimations.

[31] Estimation is here performed with the automatic bandwidth selection of the weighting matrix for the disturbance generating process of the variance-covariance matrix (there are no particular assumptions to be made about it). Moments are determined following Andrews' autoregressive methodology.

[32] The χ^2 value of the Jarque-Bera test here goes down to 18,593.84 as compared to the value of 26,779.12 that was obtained using the ARMA specification in [9.8].

[33] See the discussion in Section 4.

[34] The numbers in parentheses indicate the lag order of the GARCH specification (see footnote 12). Other specifications with respect both to (1) the ARMA structure of regressors in [9.8] and (2) the lagged structure of the variance equation, provide slightly better results. Here, the choice of referring to the same specification adopted in [9.8] is nonetheless preferred to maintain the highest consistency across the different estimation approaches presented in this Section 6.

variance is a function of, (1) a constant, (2) the estimation of the conditional variance until the last observation before t, $\sigma^2(t-1)$, and (3) information about innovations in the previous period $\varepsilon(t-1)$. Given the presence of large spikes in electricity prices, it, therefore, makes sense to imagine that in this type of market, positive price innovations have different effects from negative innovations. An EGARCH(1,1) specification models the conditional variance in a logarithmic way, as described below,

$$\ln\left[\sigma^2(t)\right] = \omega + \gamma^1 \ln\left[\sigma^2(t-1)\right] + \gamma^2 \left|\frac{\varepsilon(t-1)}{\sigma(t-1)}\right| + \gamma^3 \frac{\varepsilon(t-1)}{\sigma(t-1)}. \qquad [9.10]$$

Here, if γ^3 is different from zero, the effect of an innovation is, therefore, asymmetric and exponential.

With an EGARCH estimation, improvements with respect to residual normality are excellent, without material loss in either the significance of regressors or in the goodness of fit. Unfortunately, using the same ARMA specification as in 9.8 introduces some serial correlation in residuals.

We circumvent this problem by re-specifying the lag structure of our ARMA model (within the maximum time window of one week of trading) as,

$$-Y(t) = \alpha + \beta^1 RL(t) + \sum_{i=1}^{4} \beta^{1+i} u(t-i) + \beta^6 \eta(t-1) + \beta^7 \eta(t-3) + \varepsilon(t). \qquad [9.11]$$

Table 9.4 presents EGARCH estimation statistics.[35]

Table 9.4 – EGARCH Estimation Statistics

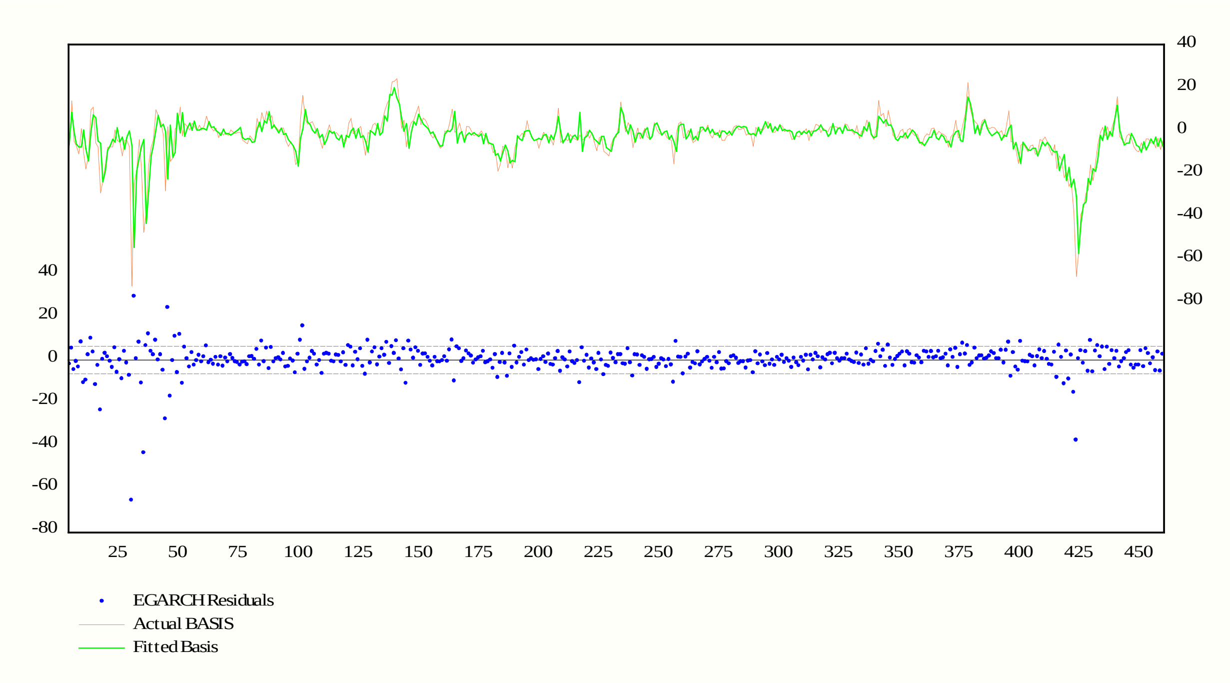
Mean Equation (ARMA Specification in [9.11])			
Statistic	Value	z	Prob. z
α	-17.78400	-5.678035	0.0000
β^1	4.921593	18.84899	0.0000
β^2	0.093774	1.864278	0.0623
β^3	0.371047	23.01242	0.0000
β^4	0.846726	56.00497	0.0000
β^5	-0.348952	-7.082831	0.0000
β^6	0.464834	238.6482	0.0000
β^7	-0.787020	-311.7142	0.0000

[35] In the estimation of [9.10], different distributions for errors can be assumed. EViews in fact estimates GARCH models by maximizing the likelihood function of error variance, given their distribution. Here we choose a generalized error distribution (GED) with a parameter of 1.5. In this way, we inform the estimation on the fat-tailed nature of our disturbances (that is, of the presence of price spikes).

Mean Equation (cont'd)				
F	69.09332			
Prob. F	0.000000			
Adjusted R^2	0.622618			
Durbin-Watson	1.900707			
Variance Equation (as of [9.10])				
Statistic	ω	γ^1	γ^2	γ^3
Value	-0.102128	0.315226	-0.126205	0.951777
z	-1.739489	5.563972	-3.615108	59.43451
Prob. z	0.0819	0.0000	0.0000	0.0000

Figure 9.3 provides a comparison between the actual and the modelled basis.

Figure 9.3 - Adjusted Basis vs. EGARCH Modelled Residual Load



With this modified lag structure residuals no longer present significant serial correlation and the normality of their distribution is greatly ameliorated. Table 9.5 specifically illustrates a comparison between the normality of the distribution of GMM and EGARCH residuals, which clearly supports the better performance of the latter estimation procedure.

Notice that Figure 9.3 also sketches the ability of the ARMA specification in [9.10] (estimated with the EGARGH method) to use a regressor on the right hand side in order to replicate the adjusted basis. Some visual improvement is detectable with respect to Figure 9.2, particularly in the ability of this approach to capture large basis swings.

Table 9.5 – Residual Distribution Statistics

Statistic	GMM	EGARCH
Mean	0.002801	-0.015712
Maximum	27.54331	3.204043
Minimum	-65.14843	-7.191287
Median	0.404030	0.019234
Std. Dev.	6.246448	1.042347
Skewness	-3.199416	-1.061752
Kurtosis	34.01387	8.670481
Jarque-Bera	18593.68	695.0812
Probability	0.000000	0.000000

§ 9.7 – Discussion of Results and Conclusions

In this Chapter, we tested the hypothesis that differences between forward and spot prices in an electricity marketplace—the German one—may be explained by leveraging on an interpretation of the storage theory through which the impossibility of directly observing power inventories is bypassed by the construction of a measure of retained power production capacity. Using daily residual load observations in the German power grid, it has been shown that, in our dataset, available residual capacity maintained to cope with unanticipated demand swings has a significant role in driving the power spot-forward price basis.

This result may possibly provide some grounds for two separate considerations. On the one hand, it may suggest that electricity is not altogether different from other tradable commodities. Certainly, non-storability in a direct fashion and the necessity to declare before time bids and offers for its exchange, give particular features to the trading of this secondary source of energy. However, the fact that a specific economic factor, residual production capacity, seems to replace the role of inventories in guiding the convenience of inter-temporal exchanges, may mean that power trading does not respond to a pricing rationale different from that of other industrial commodities.

This leads to a second observation. As the exchange of power in dedicated financial markets is still greatly undeveloped when compared to the trading of mature commodities, the existence of a non-heterodox explanation that possibly bears a functional relationship between term and spot power prices, might anticipate the ability of power markets to evolve towards greater completeness. Experience shows that, even in the presence of challenging financial innovations, traded asset prices tend to respond to an identifiable rationale, if minimum liquidity is present and information is available (McKinlay and Ramaswamy, 1988). Financial actors follow a learning process in their trading activities. Their ability to develop rational bidding behaviour and eliminate arbitrage opportunities that plague young markets, improves over time. Therefore, provided some basic transparency and liquidity are at work, power exchanges may not, in the end, be relegated to the realm of financial exoticism and might, perhaps, assume a greater role in giving enhanced public utility to the liberalization of electricity markets.

Bibliography

Bessembinder, H. and M. L. Lemmon, ‘Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets’, *Journal of Finance,* 57, 2 (2002).

Borestein, S., ‘The Trouble with Electricity Markets (and Some Solutions)’, *Program on Workable Energy Regulation Working Paper*, (2001).

Botterud, A., A. Bhattacharyya and I. Marija. ‘Futures and Spot Prices – An Analysis of the Scandinavian Electricity Market’, *Norwegian Research Council Working Paper,* (2003).

Brennan, M. The Price of Convenience and the Pricing of Commodity Contingent Claims. In Lund, D. and B. Oksendal (editors), Stochastic Models and Option Values. New York, New York: Elsevier. (1991).

Bresnahan, T. and P. Spiller, ‘Futures Market Backwardation under Risk Neutrality’, *Economic Inquiry,* 24 (July 1986).

Copeland, T. and F. Weston, *Financial Theory and Corporate Policy*, (Reading, Massachusetts: Addison-Wesley, 1992).

Deaton, A and G. Laroque, ‘Commodity Prices’, *The Review of Economic Studies,* 59, 1 (1992).

Escribano, A., J. Pena, and P. Villaplana, ‘Modeling Electricity Prices: International Evidence’, *Universidad Carlos III de Madrid Working Paper*, (2002).

Fama, E. and K.R. French, ‘Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage’, *The Journal of Business*, 60, 1 (1987).

Fama, E. and K.R. French, ‘Business Cycles and the Behavior of Metal Prices’, *Journal of Finance,* 43, 5 (1988).

Gjolberg, O. and T. Johnsen, ‘Electricity Futures: Inventories and Price Relationships at Nord Pool’, *Discussion Paper,* (2001).

Greene, W., *Econometric Analysis,* (Upper Saddle River, New Jersey: Prentice-Hall, 2003).

Hull, J., *Options, Futures, and Other Derivative Securities*, (Upper Saddle River, New Jersey: Prentice-Hall, Inc., 1993).

Kaldor, D., ‘Speculation and Economic Stability’, *Review of Economic Studies*, 7 (1939).

Keynes, J.M., *A Treatise on Money,* (London, United Kingdom: Macmillan, 1930).

McKinlay, C. and K. Ramaswamy, ‘Index-Futures Arbitrages and the Behavior of Stock Index Futures Prices’, *Review of Financial Studies*, 1, 2 (1988).

Mork, E., ‘The Dynamics of Risk Premiums in Nord Pool’s Futures Market’, *24th USAEE/IAEE North American Conference Proceedings*, (2004).

Ng, V. and C. Pirrong, ‘Fundamentals and Volatility: Storage, Spreads, and the Dynamics of Metal Prices’, *Journal of Business*, 67, 2 (1994).

Pindyck, R.S., ‘Inventories and the Short-Run Dynamics of Commodity Prices’, *The Rand Journal of Economics*, 25, 1 (1994).

Pirrong, C., ‘The Price of Power: The Valuation of Power and Weather Derivatives’, *Oklahoma State University Working Paper*, (2001).

Routledge, B, D. Seppi, and C. Spatt, ‘Equilibrium Forward Curves for Commodities’, *Journal of Finance*, 55, 3 (2000).

Telser, L., ‘Futures Trading and the Storage of Cotton and Wheat’, *Journal of Political Economy*, 66 (1958).

Woo, C., I. Horowitz, and K. Hoang, ‘Cross Hedging and Forward-Contract Pricing of Electricity’, *Energy Economics*, 23 (2001).

Working, H., ‘Theory of the Inverse Carrying Charge in Futures Markets’, *Journal of Farm Economics*, 30 (1948).

Working, H., ‘The Theory of the Price of Storage’, *American Economic Review*, 39 (1949).