NTNU
Norwegian University of
Science and Technology

Empirical analysis of supply and demand dynamics within a hydro based electricity market

Knut-Harald Bakke
Jon Ragnar Viggen

Project paper
Submission date: December 15 2011
Supervisor: Stein-Erik Fleten

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

Abstract

Due to the possibility of storage, a simultaneous treatment of supply and demand is suitable in hydro dominated electricity markets. In this paper we reassess the predictive ability of a fundamental market model originally developed in the mid-1990s for the Norwegian electricity sector. Albeit improved when considering only one state for price change, we find that the fit and short-term performance of this model is poorer when applied with newer data from a Norwegian pricing area. As the model hinges on an assumption about physical constraints being non-binding, model extensions serve little purpose. Error correction models are applied in order to obtain a benchmark for the fundamental model.

Contents

Acknowledgements

We would like to thank our supervisor, Stein-Erik Fleten, for guidance and feedback. In addition, senior researcher Terje Skjerpen at SSB and advisor Finn Erik Pettersen from NVE have provided useful advice in the aspects of modeling. Finally, we would like to thank Andreas Sylte and Siren Ørnes in Statkraft for sharing their practitioner's point of view.

1 Introduction

In deregulated electricity markets the actual operation of the production units depends on the decisions of firms whose goals are to maximize profits. Hence, electricity firms are exposed to significantly higher risks and their need for suitable decision-support models has greatly increased since the time when production was determined by state based centralized procedures (Ventosa, Baillo, Ramos, & Rivier, 2005). According to Weron (2006) accurate load models reduce costs by improving economic load dispatching and unit commitment, but due to high volatility of electricity prices the market risk related to trading is also considerable. Weron & Misiorek (2005) stress that high volatility forces wholesale consumers and producers to hedge not just against volume risk but also against price movements. In order to hedge efficiently, it is thus important to study and then accurately model the dynamics of the electricity price. Forecasts emanating from these models will in turn support bidding strategies that will contribute to the maximization of profits. In addition, regulatory agencies are dependent on such models in order to monitor market behavior.

In recent years various techniques have been applied to electricity price modeling. Weron (2006) classifies these as fundamental, stochastic and non-parametric models[1]. An alternative classification is made in Davison et al (2002) where electricity prices are modeled either through a top-down or a bottom-up approach. The latter approach refers to modeling of the electricity systems itself, whereas the former considers historic data of prices in order to directly infer aspects related to price trends and volatility. In the literature, many of the approaches considered are hybrid solutions. Depending on the purpose of the price model, the time horizon may vary from an hourly basis to several decades into the future. Motivated by a bottom-up philosophy, Johnsen (2001) develops a fundamental model of a hydro based electricity system. The focus of Johnsen (2001) is on short-term adjustments in the market.

The Norwegian electricity sector is characterized by the dominance of hydro generation and the extensive use of electricity for heating purposes (Pettersen & Holstad, 2011). These characteristics make weather conditions very important in forecasting demand and supply. Due to the considerable storage capacity, hydro generators can deal with fluctuations in precipitation levels by storing water in wet periods to generate in dry periods. But in a typically cold and dry period, electricity demand will be higher and generation will be lower than in a normal year resulting in higher prices. Conversely, for a typically wet and warm period, electricity generation will be higher and demand lower compared to a normal year, leading to lower prices (Førsund, 2007). Therefore the Norwegian market is also characterized by volatility in electricity prices. This price volatility between wet and dry years is reduced through the ability to store water and the market integration with other Nordic and European countries (Johnsen & Willumsen, 2009).

Johnsen (2001) made one of the first empirical electricity market studies with a simultaneous treatment of demand and supply. This simultaneity is especially important in the hydro based Norwegian electricity sector where storage of hydropower is possible. In this paper we replicate the modeling approach of Johnsen (2001) with an up-to-date dataset from a Norwegian pricing area.

[1] Further classifications are *cost-based models* that simulate the operation of generating units aiming to satisfy demand at minimum cost, and *game-theoretic approaches* such as cost-based models with strategic bidding or agent-based models. The *modeling horizon* is usually classified as long term (e.g. investment profitability analysis), medium term (e.g. derivatives pricing) or short term (e.g. bidding for spot electricity in a power exchange).

Then we evaluate the model, its limitations and possible extensions. We find that the fit and short-term performance of the model is poorer when applied to the Norwegian electricity market as of today. Since the model is based on an assumption about physical constraints being non-binding, we cannot incorporate multiple producers in an adequate manner. Furthermore, we find that the time-dependent effect of unexpected change in weather parameters is still present, and that the model does not manage the inclusion of several planning periods very well.

In section 2, Literature review, we will go through the assumptions underlying Johnsen's model from 2001. We will further comment on these and verify the derivation of his market model. In addition, we consider more recent and related literature. The dataset is presented in section 3 with a focus on describing characteristics in a hydro based system. In section 4, Methodology, we explain the steps taken in estimating the model. The results are given in section 5 followed by a discussion in section 6. Finally, we conclude on the model in section 7.

2 Literature Review

In a fundamental modeling approach, price dynamics are described by modeling the impact of important physical and economic factors on the commodity price (Skantze, Gubina, & Ilic, 2000). As such, fundamental modeling is strongly related to a bottom-up approach of transforming data and theory into simulated prices, see Davison et al (2002) and Fleten & Lemming (2003). In a bottom-up model, theory is based on the structure of the system, typically demand and supply, and the data is related to physical conditions, such as load and hydro inflows.

2.1 The fundamental approach of Johnsen (2001)

Johnsen (2001) analyses the joint determination of electricity demand and price in the competitive Norwegian electricity market in the time period from 1994 to 1996. He considers variables that are assumed to be crucial for demand and supply in a hydro based system, and investigates whether demand is affected by changes in the price or not. Furthermore, the impact of inflow and demand shocks on the price is considered. In order to explain the observed price and demand figures, Johnsen (2001) applies rational expectations and a seasonal approach.

Hydro power producers face a stochastic dynamic programming problem because the production or use of water in a particular week, and thereby the storage for next week, is decided under uncertainty about next week's inflow and demand conditions (Johnsen, 2001). The arbitrage condition for competitive storage states that today's price depends on the discounted present future price less storage costs. From this equilibrium condition Johnsen (2001) derives an expression where the change in the actual price from one week to the next depends on deviations between actual and expected weather conditions. The weather conditions in question are related to inflow, snowfall and temperature. Due to the non-negativity constraint on storage, the effect of deviations on the day-ahead price change[2] becomes stronger the closer one is to the next snow-melting period. Lucia and Schwartz (2002) find that this seasonal behavior of electricity prices has implications for derivatives pricing. They suggest that a simple sinusoidal function is adequate in order to capture the regular pattern of the futures and forward curve directly implied by the seasonal behavior of the electricity spot price.

[2] Since the data in the paper has a weekly time resolution, the notion of day-ahead price is here somewhat imprecise.

2.1.1 Assumptions

The stored amount of water is an important variable in the producer's optimization problem (Johnsen, 2001). Water and snow budgets thus constitute two fundamental pieces in the market model. The producer is assumed to observe the inflow (which consists of both rainfall and snow melting), temperature and snowfall of the week before the production decision is made. In the final market model it is assumed that overflow will not occur.

At the outset, Johnsen (2001) identifies a number of physical constraints relating to the management of hydro reservoirs, e.g. upper and lower filling constraints. However, in the derivation of his results, Johnsen (2001) assumes the constraints to be non-binding. This is partly supported by his data. Where no evidence is found in the dataset to support this simplification, it serves the purpose of making the derivation of his simultaneous equations possible.

Johnsen (2001) focuses on short-term price movements and uses an annual planning period consisting of 52 weeks, where the final week of the period, T_1, is the last week before the snow melting starts. In reality, the start of the snow melting period is random. Nonetheless, this implies that the planning period runs from week 19 in year t to week 18 in year $t+1$. Within this time period, price changes are random due to weather conditions.

The individual producers take the day-ahead price as given, while at the market level the price is endogenous. Johnsen (2001) therefore reformulate the problem as a social planner's problem. In other words, there is only one producer in the system. Risk neutrality is further assumed, and the weekly interest rate is assumed to be zero. For a given week, t, the social planner's problem is thus to select a water storage level that maximizes the flow of expected future welfare. The welfare is measured as the area under the demand curve in each of the future periods plus the terminal value of water in the reservoir and the snow volume which in the future will melt and flow into the reservoirs (Johnsen, 2001).

Johnsen (2001) assumes the expected weather – inflow, temperature and snowfall – for an arbitrary week to be the same and independent of which week the expectations are made. The intuition is that one week's weather will not change the expectations about the weather in the next week. This means that the expected value of a variable is related to a 'normal value' that is based on historical data.

The planning period is split in a snow-melting period of 15 weeks, and the rest of the year. Deviations between actual and expected inflow are assumed not to impact the day-ahead price in the snow melting period, since snow volume deviations are already discounted into the price when snowfalls were observed. A dummy variable is thus introduced to account for this relation.

Finally, throughout the snow melting period the inflow to the hydropower plants may reach very high levels, and Johnsen (2001) allows for a very low price in weeks when total inflow exceeds a critical level. This leads to three additional *state equations* for the price change. The first state considers a situation where there is a sequence of weeks with high, non-storable inflow. Consequently there will be no price change. The second state applies to a week where high, non-storable inflow starts to occur, but did not occur in the previous week, leading to a sudden price fall. In the third state the case is just the opposite, with high inflow not occurring in a week, but did so in the previous week. Thus a sudden price jump is expected. For ordinary weeks, the price change is described by the interior and time-dependent solution.

2.1.2 Fundamental dynamics

The fundamental dynamic of Johnsen's (2001) market model is that price changes occur when the actual weather values deviate from their expected, or normal, values. The magnitude of the influence, however, is time dependent. The effect of an unexpected high inflow is therefore much larger towards the end of the planning period than it is in the start of the period. If inflow, temperature and snowfall are equal to their expected values, the price will remain unchanged.

2.1.3 Econometric specification

A thorough understanding of the mathematical derivation of Johnsen's market model is relatively hard to acquire from just reading the original paper. Therefore, we have derived and verified his results in Appendix A. The results of the derivation are presented in the following subsections.

The demand side

Based on the assumption of rational hydro power producers, a simple demand equation is applied. The intuition is that high inflow, which increases the supply, will lead to lower prices, which increases demand. Likewise, the producers know that cold weather results in higher demand and thus a higher price. Short term change in demand is specified as a linear auto-regressive distributed-lag function

(1) $$\Delta y = \alpha_0 + \alpha_1 \Delta p_j + \alpha_2 p_{j-1} + \alpha_3 y_{j-1} + \alpha_4 \Delta w_j + \alpha_5 w_{j-1} + \alpha_6 \Delta \tau_j + \alpha_7 \tau_{j-1} + \varepsilon_j,\ j = t,\ldots,T_1$$

where α_i are unknown coefficients, w is a vector of exogenous explanatory variables such as day length, τ is temperature measured as heating degree days, and ϵ is the error term, $\epsilon_j \sim N(0,\sigma^2_\epsilon)$.

The supply side

Backward induction and maximization of the terminal value gives the terminal condition

(2) $$r_{T_1} = \beta_0 p_{T_1} + \beta_1 s_{T_1} + \beta_2$$

where β_i are unknown coefficients, r_{T1} is the water reservoir filling measured in energy units, s_{T1} is snow volume measured in energy units, and p_{T1} is price at time T_1.

When the shadow prices for the physical constraints are zero, the arbitrage condition for competitive storage gives

(3) $$p_t = E_t p_{t+1} = E_t p_{t+2} = \ldots = E_t p_{T_1}$$

which implies no arbitrage from storing more or less water from today to future periods. Reformulating the above equation gives

(4) $$\Delta p_t = p_t - p_{t-1} = E_t p_{T_1} - E_{t-1} p_{T_1}$$

Assuming no overflow in the water and snow budgets give

(5) $$r_{T_1} = r_t + \sum_{j=t+1}^{T_1} (I_j - y_j)$$

and

(6) $$s_{T_1} = s_t + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})$$

respectively. *I* denotes inflow, I^s denotes snowfall, *y* is production of hydropower in energy units, and ψ is a variable for snow melting.

Combining the equations above yields the following equation for the short term change in price

(7) $$\Delta p_t = \frac{1}{(T_1 - t - 1)\alpha_2 - \alpha_3 \beta_0} \begin{Bmatrix} (1-(1-\alpha_3)L)\left[(I_t - EI_t) - \beta_1(I_t^s - EI_t^s)\right] \\ -(\alpha_6 - (\alpha_7 - \alpha_6)L)(\tau_t - E\tau_t) - \varepsilon_t \end{Bmatrix} + u_t, u_t \sim N(0, \sigma_u^2)$$

where *L* indicates the lag operator.

The system of equations

By including a dummy variable for Christmas and Easter holidays and considering only day length in the vector for exogenous variables in the demand equation, and by including the state equations and a dummy variable for the snow melting period in the supply equation, the following system of equations is obtained

$$\Delta y = \alpha_0 + \alpha_1 \Delta p_t + \alpha_2 p_{t-1} + \alpha_3 y_{t-1} + \alpha_4 \Delta d_t + \alpha_5 d_{t-1} + Dummy_{Holiday} + \alpha_6 \Delta \tau_t + \alpha_7 \tau_{t-1} + \varepsilon_t$$

$\Delta p_t = 0,$ *if I_t and I_{t-1}> K*

$\Delta p_t = \partial_0,$ *if I_t> K or*

$\Delta p_t = \partial_1,$ *if I_{t-1}> K or*

$$\Delta p_t = \frac{1}{(T_1 - t - 1)\alpha_2 - \alpha_3 \beta_0} \begin{Bmatrix} (1-(1-\alpha_3)L) \\ \left[Dummy_{Snowmelting}(I_t - EI_t) - \beta_1(I_t^s - EI_t^s)\right] \\ -(\alpha_6 - (\alpha_7 - \alpha_6)L)(\tau_t - E\tau_t)_t \end{Bmatrix} + \omega_t,$$

$$\omega_t = -\frac{1}{(T_1 - t - 1)\alpha_2 - \alpha_3 \beta_0} \varepsilon_t + u_t$$

Problems with the error terms

As is evident from the system of equations above, there are some problems with the error terms. First, the variance in the price equation is time-dependent and increases throughout the season, i.e. the variance is heteroscedastic. Second, the error terms in the two equations are not independent. Johnsen (2001) finds no satisfactory way to deal with these problems in the error terms.

2.1.4 An alternative model - ECM

In order to evaluate the above model's out-of-sample predictive ability, Johnsen (2001) compares it with error correction models (ECM) for both price and demand. The snow melting period is treated as in the fundamental model. The main difference is the modeling of the price movements. In the fundamental price model, the coefficients change over the season, and actual generation is not an explanatory variable in the price. Furthermore, the price and demand equations have common coefficients in the fundamental model. In the ECM alternative, however, generation is included as an explanatory variable in the price equation, coefficients are unchanged throughout the year and there are no cross-equation parameter restrictions (Johnsen, 2001). The price quantity endogeneity is still recognized and the following error correction models are suggested in the paper:

$$\Delta y = \alpha_0 + \alpha_1 \Delta p_t + \alpha_2 p_{t-1} + \alpha_3 y_{t-1} + \alpha_4 \Delta d_t + \alpha_5 d_{t-1} + \alpha_6 \Delta \tau_t + \alpha_7 \tau_{t-1} + Dummy_{Holiday} + \varepsilon_t$$

$\Delta p_t = 0,$ *if I_t and I_{t-1}> K*

$\Delta p_t = \partial_0,$ *if I_t> K or*

$\Delta p_t = \partial_1,$ *if I_{t-1}> K or*

$$\Delta p_t = \beta_0 + \beta_1 \Delta y_t + \beta_2 y_{t-1} + \beta_3 p_{t-1} + \beta_4 Dummy_{Snowmelting} \Delta(I_t - EI_t) + \beta_5 (I_{t-1} - EI_{t-1})$$
$$+ \beta_6 \Delta(I_t^s - EI_{t\,t}^s) + \beta_7 (I_{t-1}^s - EI_{t-1}^s) + \beta_8 \Delta(\tau_t - E\tau_t) + \beta_9 (\tau_{t-1} - E\tau_{t-1}) + Dummy_{Holiday} + \omega_t$$

Johnsen (2001) implicitly assumes that the individual time series are integrated of first order, I(1). So by first-differencing and through the existence of some cointegrating vector of coefficients, a stationary linear combination of them are formed, see Engle & Granger (1987) and Greene (2002).

2.2 Related literature

The fundamental approach of Johnsen (2001) has a bottom-up treatment of time-dependency in the electricity price equation. As such, it is unparalleled in the literature. During the last decade, however, several papers have been published on fundamental models in hydro based electricity markets. Among these are Vehviläinen & Pyykkönen (2005), and Tipping & Read (2010).

Vehviläinen & Pyykkönen (2005) presents a stochastic factor based approach to mid-term modeling of spot prices in the Nordic electricity market. The fundamentals affecting the spot price are thus modeled as stochastic factors that follow statistical processes. As in Johnsen (2001), hydro production depends on the water level in the reservoirs, and hydro inflow results mainly from precipitation and the melting of snow-pack. Generation of the snow-pack is further directed by precipitation and temperature. Temperature is what governs the demand for electricity the most in Vehviläinen & Pyykkönen (2005). In contrast to Johnsen (2001), the market equilibrium is approximated by assuming that the electricity demand is inelastic.

Tipping & Read (2010) claim that bottom-up analysis is suitable for modeling structural change and hydro variation, but such models must make assumptions about fundamental system data and rational optimizing behavior that leave significant unexplained price volatility. Hence, they describe a technique for fitting a hybrid model, in which they estimate parameters for a simplified bottom-up model of participant behavior from market data, together with a stochastic process describing residual price volatility. The fitted model is then used to simulate market behavior in the hydro dominated New Zealand electricity market, as fundamental parameters vary.

Vehviläinen & Pyykkönen (2005) estimate a simplified Marginal Water Value (MWV) curve from Nordic market data on a monthly basis, and in addition to a piecewise linear supply curve for thermal power, they use this curve to calculate the level and price of both hydro and thermal power required to meet system load. The New Zealand model of Tipping & Read (2010) uses a similar MWV curve representation as its bottom-up model, but instead of assuming that the water value is to depend on the deviation from a constant storage level, they use a relative storage level. Even though

Vehviläinen & Pyykkönen (2005) form a simulation model, they do not employ a complementary stochastic process as in Tipping & Read (2010). This complementary stochastic process is not included in Vehviläinen & Pyykkönen (2005) due to their bottom-up model's ability to explain monthly prices in a market with significant storage.

Recent work relating to the demand-supply simultaneity in the hydro based Norwegian electricity sector, tend to focus on ECMs. This is the case for the electricity market models applied by both NVE and Statistics Norway, compare Johnsen & Willumsen (2010) and Holstad & Pettersen (2011). ECMs are relatively straight-forward to specify, more data driven and not as mathematically complex as a fundamental model.

3 Data

The data sample in Johnsen (2001) is from 1994 to 1995 when the Norwegian market for electric power could be considered to be isolated, and the market was therefore easily identified through a dataset. Since the mid-1990s there have been major changes in the market conditions affecting the definition of a pure 'Norwegian market' for electricity. During the years after the making of Johnsen's article several other countries have joined the Nord Pool spot market, making the system price from Nord Pool inaccurate for describing a Norwegian price for electricity. In addition to the market integration with other Nordic countries, there has been a subdivision of the Norwegian price into different price areas. Today this increased level of complexity makes it hard to capture and model a solely Norwegian market.

To handle this complexity in a coherent manner, we will use data from one of the price areas in Norway. By using the price area NO1 from 2003 to 2008 with an out-of-sample period in 2009, we are trying to model the fundamental dynamics addressed in Johnsen (2001).

Johnsen (2001) used a time span of two years (1994-1995) with an out-of-sample period of one year (1996). In this replication we will use a larger time span, covering a period of 5 years (2003-2008) and an out-of-sample period of one year (2009). The starting point of the dataset is chosen to exclude the extremely high and volatile prices during the winter of 2002-2003. Throughout our sample period the price areas in Norway are considered to be geographically stable and constant. From 2010 and forth, however, Statnett has made several changes to the price areas, making it difficult to collect data representative for the NO1 area. This instability in the price areas is the reason for ending the sample period in 2008 and using 2009 as the out-of-sample period.

3.1 Collection of data

Both political and market change in the electricity sector force us to look at a broad range of data providers to ensure that we are able to isolate and mimic the fundamental dynamics Johnsen (2001) is addressing. The main data sources are Statkraft, the Norwegian Water Resources and Energy Directorate (NVE) and Nord Pool spot. In addition, data on weather variables is collected from public databases like SeNorge.no and the Norwegian Meteorological Institute through met.no and eklima.no. There is little information available on snowfall for the price area used in the replication, and data representation of snowfall is thus generated from modeling. We determine the snowfall by looking at precipitation combined with temperature from the same measuring stations in the NO1 area. When the temperature is below 0 degrees Celsius the precipitation is assumed to be snowfall.

This is a simplification and ideally snowfall should be simulated according to the HBV hydrology model (Bergström, 1992), but such an approach goes beyond the scope of this paper.

3.2 Seasonability and deviations from normal

During the estimation period, the day-ahead price has been fluctuating between a top level of 76.4 Euro/MWh in week 34 of year 2006 and a bottom level of 3.8 Euro/MWh in week 34 of year 2007. The price spike in 2006 was due to high prices of fossil fuels and a strained hydrological situation with little snow in the winter, a dry summer and problems with the Swedish nuclear power plants[3]. In the out-of-sample period we see that the prices have been in the interval from 65.8 Euro/MWh in week 39 in 2008 to 10.0 Euro/MWh in week 19 in 2008. High prices of fossil fuels, increasing CO_2 prices and reduced capacity in the Swedish nuclear power plants are important factors that explain the peak in 2008[4]. As the explanations of the price spikes suggest, there are several factors which are not included in Johnsen's model that also affect price. The spot prices used for estimation purposes in area NO1 are collected from Nord Pool.

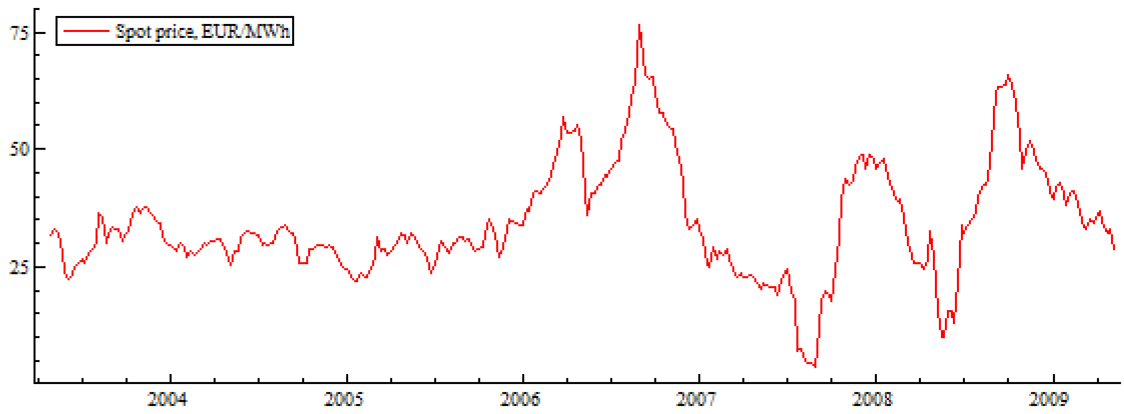

Figure 1: Spot price NO1

The demand shows seasonal patterns and is significantly lower in the summer season than during the winter season as the graph below illustrates. The demand is from the area NO1, and the data is provided by Statkraft. We can see deviations from the normal pattern of demand in the early winter of 2007. During this period the temperatures were very high compared to the normal.

[3] E-CO, Third quarter results 2006: http://www.e-co.no/?module=Articles;action=Article.publicShow;ID=771
[4] Statistics Norway, 2008: http://www.ssb.no/elkraftpris/arkiv/art-2009-01-08-01.html

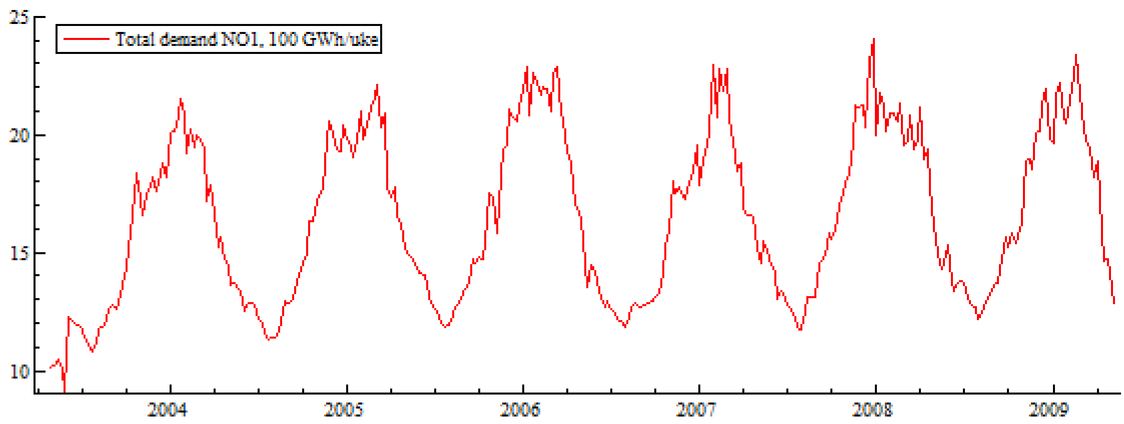

Figure 2: Demand NO1

The actual inflow is calculated as the sum of production in the NO1 area and the change in reservoir level from one period to the other in the same area. This calculation assumes that all generation emanates from hydro reservoirs, and it is a realistic assumption for the price area in question. The actual inflow was less than normal for most of the period from 2003 to the end of 2004. In 2005 and 2006 the inflow oscillates around the median level, while we observe an actual inflow higher than normal from the snow melting starts in 2007 and throughout 2008. High inflow persists in the out-of-sample period 2008-2009. Even during the years of 2003 and 2004 we see that the spikes in inflow are higher than the normal. This is due to the fact that the beginning of the snow melting period varies from year to year. Since the peak inflow occurs in different weeks each year, and not in a particular week as the normal value suggests, inflow will tend to deviate substantially from normal around this period of the year. Johnsen's assumption about the snow melting period starting in week 19 every year will therefore explain some of the observed deviations of actual inflow from the normal values.

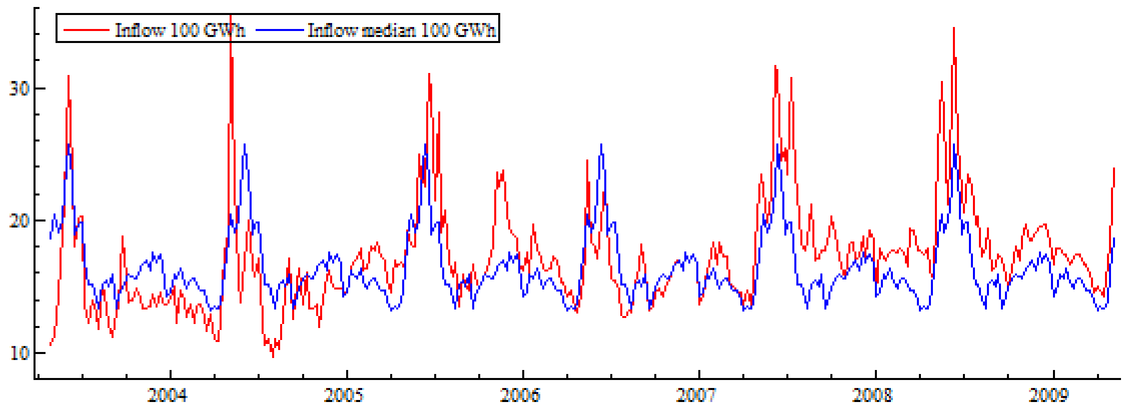

Figure 3: Inflow NO1

Temperature is measured in heating degree-days. Heating degree-days are calculated as the number of days times the number of degrees below 17 degrees Celsius. From the data we can see that the winters of 2005 and 2008 were warmer than normal, whereas the winter of 2006 was colder than normal. In the beginning of the winter of 2007, we have a short period where the temperatures were high compared to the normal. The other winters in the sample period are fluctuating around the normal value. In the out-of-sample period we have a small period in the end where we experience a few colder weeks than normal, otherwise the values are close to normal.

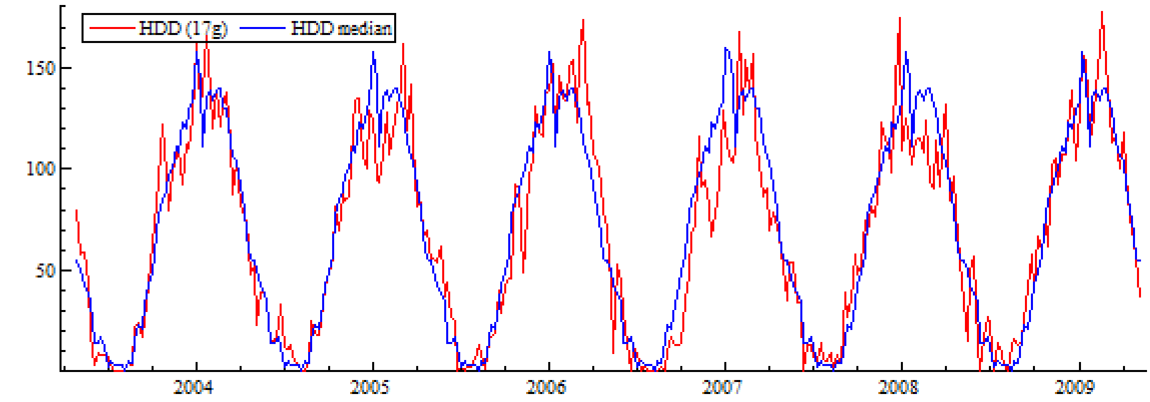

Figure 4: Heating degree-days NO1

The reservoir level has been deviating quite a lot from the normal through the sample period. In 2003 and 2004 the reservoir content was lower than normal. During the autumn of 2005 the reservoir reached a relatively high level, but was significantly reduced throughout 2006. It reached another peak in 2007. The out-of-sample period in general experiences higher reservoir content than normal.

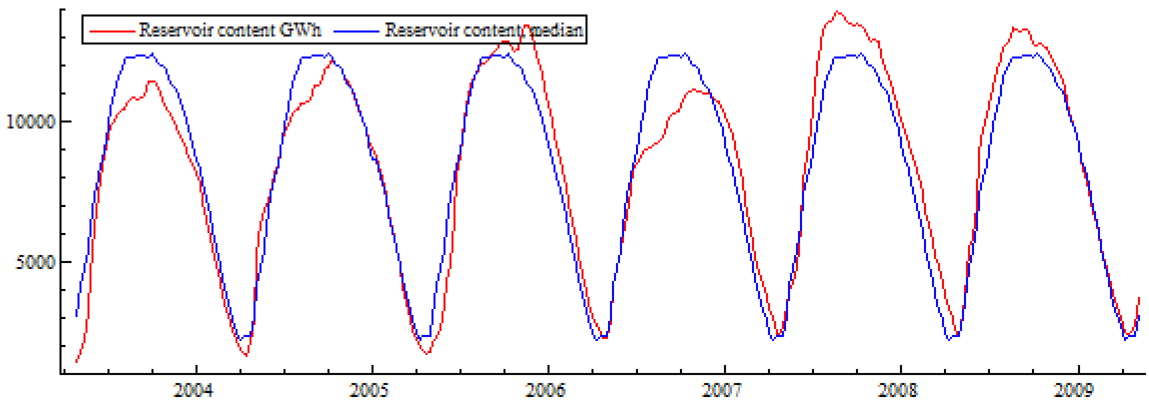

Figure 5: Reservoir content NO1

The snowfall data is obtained by looking at precipitation and temperatures from the three measurement stations Ringebu, Hol and Nissedal in the NO1 area. When the temperatures are below zero the precipitation is assumed to fall as snow. The average numbers from these stations are used to represent the deviation from normal snowfall, measured in mm water equivalents, in Figure 6.

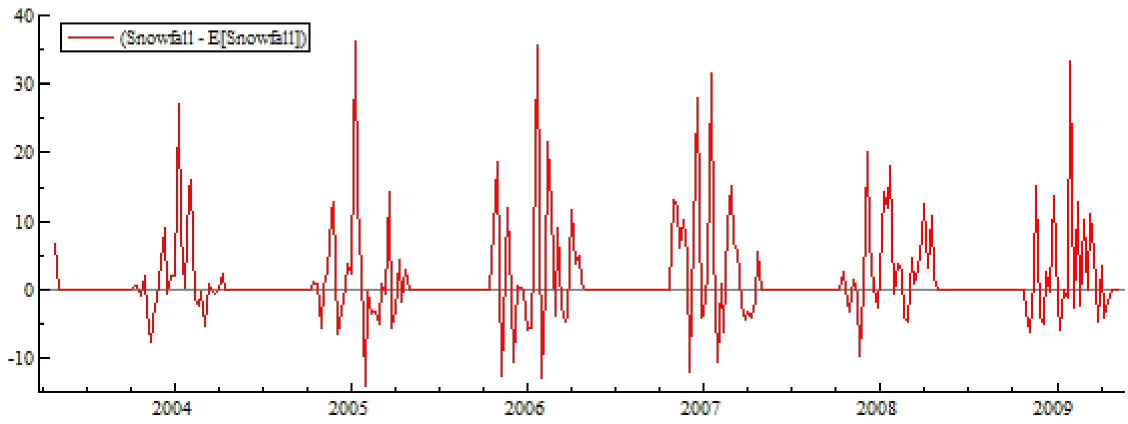

Figure 6: Snowfall NO1

The day length variable exhibits a deterministic and sinusoidal trend. Around mid-summer the day length is at its highest level. The day length variable is based on data from Blindern, Oslo, and is thus representative for the normal length of days in the NO1 area.

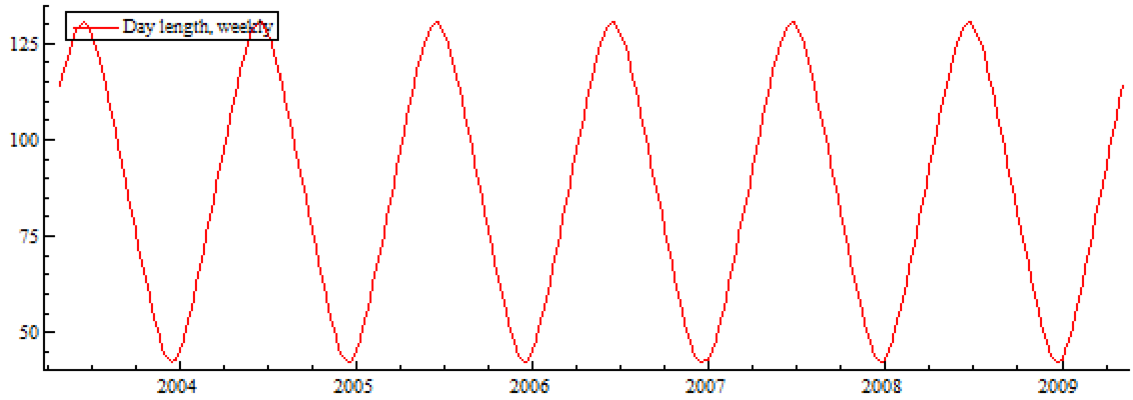

Figure 7: Normal day length NO1

4 Methodology

4.1 Fundamental approach

It is not entirely clear how Johnsen (2001) has estimated the simultaneous equations. This view is shared by other practitioners and researchers, including Pettersen (2011) and Skjerpen (2011). Due to the uncertainty concerning Johnsen's methodology, we try two different approaches to solving the simultaneous equations[5]. The first approach considers multiple equation dynamic modeling using constrained full information maximum likelihood, see Doornik & Hendry (2009). However, we encounter numerical problems when estimating the constrained system, and the regression does not

[5] We use the statistical software packages PCGive and EViews to estimate the system of equations.

converge even with a limited number of constraints[6]. According to Skjerpen (2011) the system should be solved as being non-linear in the parameters. The second approach estimates the parameters accordingly using full information maximum likelihood.

The estimation technique full information maximum likelihood (FIML) estimates the likelihood function under the assumption that the contemporaneous errors have a joint normal distribution. Provided that the likelihood function is correctly specified, FIML is fully efficient. FIML is suitable for cross-equation parameter restrictions, hence the 'constrained' denotation is used.

We formulate the simultaneous equations directly, including the interdependencies among the coefficients. The time dependent factor runs in cycles, from week 1 to week 52, through the planning periods. Changes in demand and price are defined as endogenous variables, and the state equations for price change are dealt with using dummy variables.

In essence, we make use of three dummy variables to handle the state equations for the price change. The first dummy is unity if inflow exceeds the critical level, K, in the current week t, but not in week $t-1$. The second dummy is unity in weeks where inflow exceeded K in the previous week, but not this week. Both of these dummies are multiplied by unknown coefficients, δ_0 and δ_1, respectively. Finally, a third dummy is multiplied by the time-dependent expression for price change, i.e. the dummy is unity whenever inflow for the current and previous week is below K. Otherwise the price change equation is zero as in the first state equation.

Just as in Johnsen (2001), we use a grid search procedure to find the value of critical inflow K that best fits the data. The procedure is discussion in Appendix B.

4.2 Error Correction Model approach

In the ECM alternative from Johnsen (2001) all variables are treated as if they are non-stationary. Note that this is an implicit assumption of his model, i.e. Johnsen never explicitly comments on it in his article from 2001. The assumption can be tested by studying the graphs of the variables, or, more formally, perform Dickey Fuller testing. However, we accept the assumption that all variables are integrated of first order, I(1). This means that the variables will be stationary when differencing them once. By differencing the non-stationary variables, we will be able to estimate the short term relationship between the variables. However, information on the long term relationship of the level values will be lost in the differencing process.

An ECM enables both short and long term relationships to be present in the model. A requirement is that the variables are cointegrated, meaning that there exists an equilibrium relationship that the integrated variables will tend towards in the long run. We can test for cointegration by using Johansen's maximum eigenvalue and trace tests (Greene, 2002). In addition to identifying the cointegrating relationship, Johansen's tests can check the model's assumptions about exogeneity. As the purpose of the paper at hand is to replicate the models of Johnsen (2001) with an up-to-date

[6] Based on the analysis of parameter stability from Johnsen (2001), we interpret the last week in the sample as being the value of t in the price change equation. We thus get a constant term instead of the time-dependent factor in the equation. This makes sense, as the final values in the parameter stability graphs correspond to the estimated coefficient values in Johnsen's article. Hence, we can treat the system as linear and constrain the coefficients through algebraic expressions. We essentially get six equations with six unknowns for the constrained coefficients.

dataset, we rely on the assumptions that Johnsen makes based on his understanding of the market. Hence, we do not test for cointegration.

As the ECMs are linear in the parameters, multiple-equation dynamic modeling is suitable (Doornik & Hendry, 2009). We specify changes in demand and price as the endogeneous variables, and solve the system using FIML. The dummy variable for snow melting is multiplied with the change in the deviation from normal inflow from the previous week to the present week. Thus we set the deviation from normal inflow to zero in the snow melting periods directly in the dataset. Single equation misspecification tests are then performed.

4.3 Specification of model extension

The extension focuses on including a coal fired producer in the system to see what effect the flexibility of another energy source would have on the fundamental dynamics of the electricity price. Coal, gas and CO_2 prices are affecting the price and demand for different sources of electricity, so it is interesting to see how a coal fired producer will influence the estimations. The total demand will then be covered by the sum of production from the hydro and the coal fired producer. To capture the essence of this extension an additional constraint is added for the physical maximum of coal fired production, denoted $\bar{x}$.

(8) $$x_j \leq \bar{x},\ j = t, \ldots$$

where x_t is the amount of energy produced from the coal fired power plant at time t. Following the derivation in Appendix A, we introduce the shadow price for coal production, λ_9, when maximizing the value of the reservoir content. Johnsen's original paper assumes all the constraints in the system to be non-binding, and this simplification is necessary to arrive at equation 3 which is the foundation for deriving the final model. By implementing the coal fired producer, the shadow price λ_9 would also have to equal zero. There is no logical reason for such an assumption, and continuing with this approach would require a completely new model to be established. Hence, a simplified heuristic approach is done in an attempt to include the effects from including a coal fired producer.

The coal fired producer is assumed to make his production decision based on the price of electricity and coal at time *t-1*, p_{t-1} and c_{t-1}, respectively[7]. Whenever the electricity price is lower than the coal price, the coal fired production is set to zero by utilizing a dummy variable, $dummy_{coal}$. Furthermore, it is assumed that coal fired production will constitute a maximum of γ % of the total power production in the system. Consequently the production from hydro power can be expressed as

(9) $$y_{hydro.producer} = y - y_{coal.fired.producer} = y\left(1 - \gamma \frac{(p_{t-1} - c_{t-1})}{p_{t-1}} dummy_{coal}\right)$$

Equation 9 enters the model as y_j in the expression for the terminal level of the hydro reservoir, see equation 5. The modified expression for the price change equation after completing the calculations according to Appendix A is

[7] Using the *p* and *c* values at time *t-1* is a simplification. Ideally the values at time *t* would be used, but this would lead to a polynomial expression of price change not suitable for regression purposes.

$$\Delta p = \frac{1}{(T_1 - t - 1)\alpha_2 \left(1 - \gamma \frac{(p_{t-1} - c_{t-1})}{p_{t-1}} dummy_{coal}\right) - \beta_0 \alpha_3} \tag{10}$$

$$\left[\begin{array}{l} \left(1-(1+\alpha_3)L\right)\left\{(I_t^w - EI_t^w) - \beta_1(I_t^s - EI_t^s)\right\} \\ -\left(1-\gamma \frac{(p_{t-1}-c_{t-1})}{p_{t-1}} dummy_{coal}\right)\left(\alpha_6 + (\alpha_7 - \alpha_6)L\right)\left\{\tau_t - E\tau_t\right\} - \varepsilon_t \end{array} \right] + u_t$$

Solving the model with the modified expression for the price change thus include the effect of a coal fired producer into the original hydro based system.

5 Results and interpretation

5.1 Fundamental approach

When including the state equations for price change as in the original model by Johnsen (2001), we get the estimated coefficients and their estimated standard deviation as shown in Table 2. The estimated equations explain approximately 57.2 % and 7.5 % of the observed variation in the first difference of demand and price, respectively, see Table 1. Plots of actual and fitted values, as well as scaled residuals, are shown in Appendix C.

Table 1: Summary statistics with state equations

Dependent variable	Mean	R^2	Sum of squared residuals	Standard error of regression
Δy_t	0.019	0.572	102.36	0.637
Δp_t	-0.061	0.075	1849.43	2.704

As is evident from the sum of squared residuals and the scaled residuals plots, there are some very large residuals in the equations, especially for the price change. Despite the poor fit of the model, there may be interesting information on the fundamental dynamics from the estimated coefficients.

Table 2: The estimated coefficients with state equations

Coefficient	Estimated value	Standard error	t-statistic	Probability
α_0	3.841	0.465	8.267	0.000
α_1	-0.143	0.052	-2.764	0.006
α_2	-0.027	0.003	-0.979	0.328
α_3	-0.210	0.022	-9.394	0.000
α_4	0.011	0.016	0.703	0.482
α_5	-0.011	0.003	-4.199	0.000
α_6	0.041	0.002	20.619	0.000
α_7	0.010	0.002	3.929	0.001
$Dummy_{Holiday}$	-1.019	0.103	-9.862	0.000
β_0	-12.238	2.710	-4.516	0.000
β_1	-0.065	0.038	-1.712	0.087
δ_0	-2.203	0.655	-3.362	0.008
δ_1	1.636	1.511	1.082	0.279

The variables of the two equations from the fundamental model have mathematical signs that appear to be reasonable. The negative sign of α_1 indicates that an increase in price decreases the demand. Similarly the sign of the coefficient for lagged price, α_2, is in line with the results of Johnsen (2001) and Johnsen & Willumsen (2010), but its insignificance is troublesome for the time-dependency in the price change equation. The coefficient for last week's demand, α_3, is negative as expected. Consistent with the results from Johnsen (2001), the coefficient for change in day length is insignificant, whereas the level value, d_{t-1}, is not. The coefficient for lagged day length, α_5, has a negative sign. The significance of the coefficient for temperature level, α_7, is not recognized in Johnsen (2001) while the coefficient for temperature change, α_6, is significant and positive. However, the significance of the temperature level is supported in both Johnsen & Willumsen (2010) and Holstad & Pettersen (2011). During Christmas and Easter holidays, demand is significantly reduced. Both β_0 and β_1 are negative, and they thus have the right sign for the interpretation of equation 2. β_1 is nearly significant, compare Johnsen (2001). The signs of δ_0 and δ_1 are intuitively appealing, as a sudden high inflow in one week and a sudden absence of high inflow in another, will lead to a price fall and rise, respectively.

To test the model we have carried out a dynamic simulation using the observed price and demand values for weeks 18 and 19 in 2008 as starting values. For the rest of the simulation the calculation of the endogenous variables of change in price and demand are used in combination with the real values of the exogenous variables to see how well the model performs. As Johnsen (2001) notes in his original paper, the additional year in the out-of-sample period may represent a too large challenge to this short term model, which does not handle the connection between subsequent planning periods very well.

As we can see from Table 2, the parameter α_2 is insignificant, but it has the right sign and by including the parameter we get a better out-out-sample performance of the model. The dynamic simulation of demand reproduces the observed demand pattern quite well throughout the whole estimation period. We note that the model overestimates the demand in the winter of 2004, while it underestimates the demand in the winter of 2008.

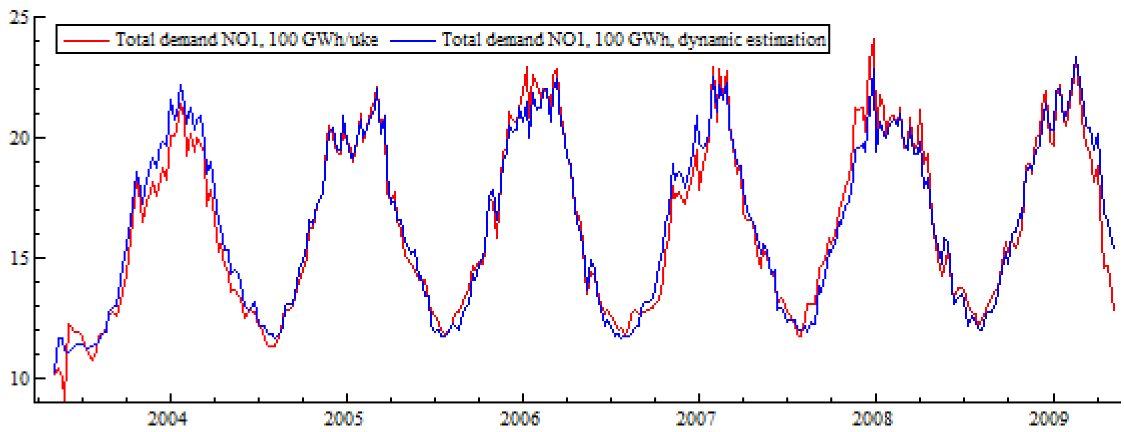

Figure 8: Dynamic simulation of demand with state equations for price change

In the price change equation the parameters α_2, β_1 and δ_1 are insignificant, but they have the right signs and improve the out-of-sample performance of the model. Consequently, they are included in the dynamic simulations. As we see from the simulations, several drops in the price are captured by the model, albeit with much lower amplitude. The price drops captured include the one in the autumn of 2003, the spring of 2004, the autumn of 2005 and the spring of 2006. In the out-of-sample period the model also detects the drop in the summer of 2008. When it comes to the price spikes, the model is not able to simulate any of them. The two spikes in 2006 and the spikes in the winter of 2008 and 2009 are not reflected in the simulation at all.

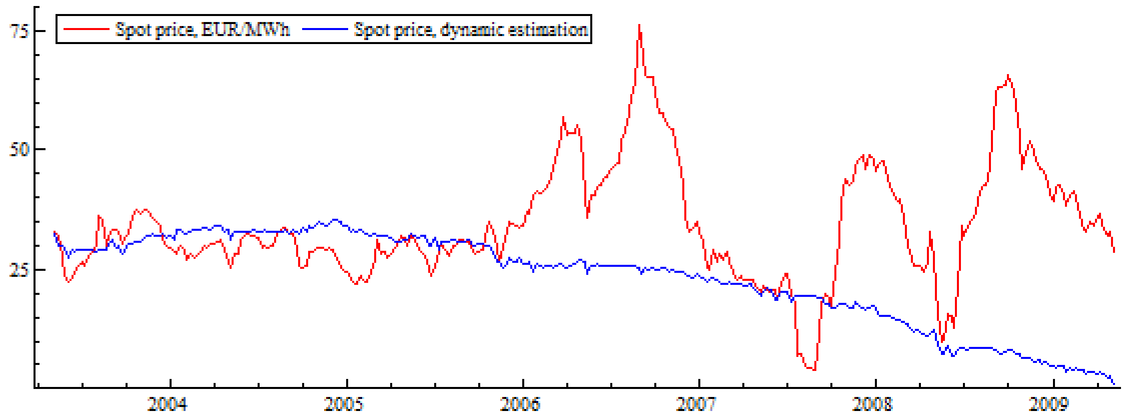

Figure 9: Dynamic simulation of price with state equations for price change

In general, the fit of the price change is poor when the model is replicated with up-to-date data. We also observe a negative trend in the model, where the price tends to move towards zero in the end of the sample. The main reason for this trend is the inclusion of the state equations in the price change, compare section 2.1.1. By including the state equations, the effect of price spikes caused by high non-storable inflow, is minimized for the time-dependent model. In the short term, this is not a problem and the simplification can be accepted. This is the case in Johnsen (2001) where only two planning periods are included in the estimation, and the negative trend can therefore not be observed. However, when the estimates are a result of looking at several planning periods, this simplification does not behave in the intended way. By construction, each δ-value will appear only once for each time where price spikes due to high, non-storable inflow, occur. As the values of δ_0 and δ_1 are -2.203 and 1.636 respectively, this will in the long term give a negative net influence on the price, leading to a negative trend in the simulated price. The longer the period included in the estimation, the closer the simulated price will move towards zero, which clearly is not the intention of including the state equations. Our grid search procedure for the K-value, see Appendix B, shows that the best fit for demand is obtained when the *K* is infinite, which basically means that the state equations of the price change are ignored.

To compare with ECMs that includes both short and long term dynamics, we continue our estimations with *K* equal to infinity to exclude the negative long term trend caused by the state equations.

5.1.1 Removal of the state equations for price change

Exclusion of the state equations for price change gives the estimated coefficients and their estimated standard deviation as shown in Table 4. The estimated equations explain approximately 62.7 % and 6.1 % of the observed variation in the change of demand and price, respectively, compare Table 3. Plots of actual and fitted values, as well as scaled residuals, are shown in Appendix C.

Table 3: Summary statistics without state equations

Dependent variable	Mean	R^2	Sum of squared residuals	Standard error of regression
Δy_t	0.019	0.627	89.32	0.595
Δp_t	-0.061	0.061	1878.36	2.714

Compared to the previous case, where the state equations were included, the demand change equation now achieves a better fit with the data, whereas the price change equation fits the data marginally worse. The residuals remain very high.

Table 4: The estimated coefficients without state equations

Coefficient	Estimated value	Standard error	t-statistic	Probability
α_0	3.878	0.471	8.225	0.000
α_1	-0.122	0.061	-1.995	0.046
α_2	-0.029	0.003	-0.977	0.328
α_3	-0.211	0.022	-9.631	0.000
α_4	0.013	0.016	0.781	0.435
α_5	-0.011	0.003	-4.103	0.000
α_6	0.041	0.002	20.009	0.000
α_7	0.010	0.002	3.912	0.001
$Dummy_{Holiday}$	-1.018	0.106	-9.613	0.000
β_0	-14.482	3.389	-4.273	0.000
β_1	-0.0833	0.050	-1.662	0.097

In general, the removal of the state equations for price change does not alter our interpretation of mathematical sign and significance. The most prominent difference is perhaps that the coefficient for price change in the equation for demand change, α_1, is now significant only on a 5 % significance level compared to the 1 % significance level obtained with the inclusion of state equations.

To test the performance of the model with infinite *K*, we carry out a dynamic simulation by applying starting values for the price and demand in week 18 and 19 in 2008. As can be seen in Table 4 the coefficient α_2, representing the lagged value of price, is still insignificant. Again, it has the right sign and we get a better out-of-sample performance for the model by including the parameter. When we compare the dynamic simulations with *K* equal to 26 and *K* equal to infinity, we recognize the same general patterns in the estimation period. If we look closely however, we can identify some differences, especially in the periods where we have spikes and falls in demand. During the first two years the simulated values follow the actual values pretty closely, while in the early spring of 2005 we start to see minor differences between the two alternatives. From this point of time and onwards, we see that for every spike and fall we have small differences between the two simulations, where the simulation with K equal to infinity gives us the best result.

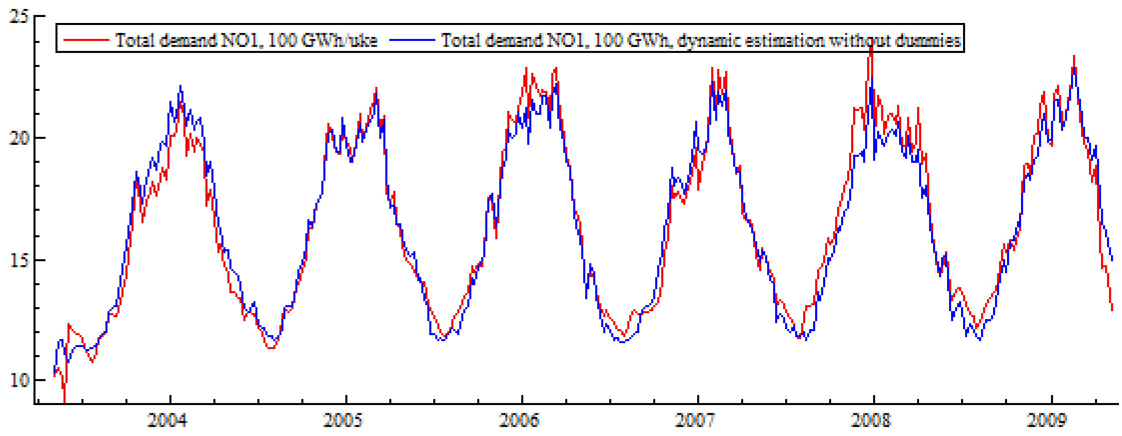

Figure 10: Dynamic simulation of demand without state equations for price change

For the price equation the parameters α_2 and β_1 are still insignificant, but have the right signs and they improve the out-of-sample performance of the model. Therefore we continue to include them in the dynamic simulations. Compared to the simulation with *K* equal to 26 we have a slightly poorer performance of the price change. While the *K* equal 26 alternative captures more drops in the price than the alternative with *K* equal to infinity, we have now managed to remove the negative trend, which ultimately will make the model more appropriate for comparing with an ECM. In the simulation with *K* equal to infinity we see that the model only identifies a few drops in the price, including the price drop in the spring of 2004, the autumn of 2005 and the spring of 2008.

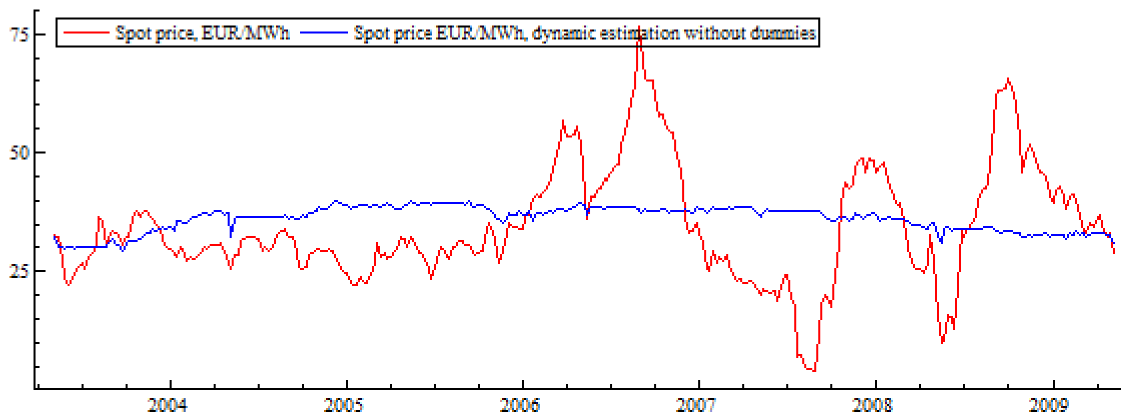

Figure 11: Dynamic simulation of price without state equations for price change

5.1.2 Problems with the error terms

Both single misspecifications tests and graphs of scaled residuals, see Appendix C, indicate autocorrelation problems in the price change equation, i.e. the model's residual in one period appear to be correlated with residuals in the previous periods. In the presence of autocorrelation, the estimators are still unbiased, but the variances, and hence the t-values of the estimates, will be affected. In addition, there are problems with conditional heteroscedasticity in the equation for price change. As pointed out in section 2.1.3, autoregression conditional heteroscedasticity (ARCH) is inherent in the model for price change and the evidence is therefore not surprising. The ARCH effect will not bias the estimates, but will cause the estimators to be inefficient. Consequently, the conclusions we draw from the estimation may be deceptive.

5.2 Error Correction Model approach

When estimating the system applying error correction models, the goodness of fit for the demand and price change equations are 57.9 % and 22.0 %, respectively. The summary statistics in Table 5 show that there still are very high residuals. This is also evident from the plots in Appendix C.

Table 5: Summary statistics for ECM alternative

Dependent variable	Mean	R^2	Sum of squared residuals	Standard error of regression
Δy_t	0.019	0.579	100.76	0.632
Δp_t	-0.061	0.220	1559.29	2.497

More coefficients are to be dealt with in the estimation of the ECMs than in the fundamental model. The upside is that their interpretation is rather obvious as there no longer are cross-equation relationships among the coefficients. The results from the estimation are given in Table 6.

Table 6: The estimated coefficients for ECM alternative

Variable	Coefficient	Standard error	Probability
Change in demand Δy_t			
Constant	3.231	0.719	0.000
Δp_t	-0.154	0.084	0.069
p_{t-1}	-0.007	0.004	0.079
y_{t-1}	-0.230	0.036	0.000
Δd_t	-0.025	0.029	0.382
d_{t-1}	-0.003	0.005	0.552
$\Delta \tau_t$	0.051	0.006	0.000
τ_{t-1}	0.016	0.005	0.002
$Dummy_{Holiday}$	-1.097	0.459	0.017
Change in price Δp_t			
Constant	0.461	0.947	0.626
Δy_t	0.839	0.336	0.013
y_{t-1}	0.023	0.054	0.674
p_{t-1}	-0.022	0.011	0.045
$\Delta(I_t\text{-E } I_t)$	-0.267	0.064	0.000
$(I_{t-1}\text{-E } I_{t-1})$	-0.043	0.044	0.330
$\Delta(I_t^S\text{-E } I_t^S)$	-0.027	0.023	0.245
$(I_{t-1}^S\text{-E } I_{t-1}^S)$	-0.015	0.030	0.611
$\Delta(\tau_t\text{-E } \tau_t)$	0.037	0.015	0.016
$(\tau_{t-1}\text{-E } \tau_{t-1})$	0.046	0.011	0.000
$Dummy_{Holiday}$	0.438	2.661	0.869

There are discrepancies in the mathematical sign and significance between some of the coefficients in the ECM models in Johnsen (2001) and the ones reproduced above. For the demand change equation the mathematical signs of the coefficients are in line with Johnsen's results. Surprisingly, the coefficient for the level of day length, d_{t-1}, is insignificant in our estimation. The values for change and level of price in the demand change equation are significant on a 10 % level. In the price change equation the lagged value of demand, y_{t-1}, has the opposite sign of Johnsen (2001), and is insignificant in our estimation. The level value of deviation in inflow has the right sign, but is not significant. Both the change and level value of deviation in snowfall have the wrong signs, but they are insignificant in the above estimation. The constant and the dummy variable for holidays are insignificant in the price change equation, but they do have the right signs.

In the dynamic simulations of the ECMs[8], the parameter for the level of day length in the demand change equation, d_{t-1}, is included because it gives a better out-of-sample result for the model. The values for change and level of price in the demand change equation are also included in the simulation. The differences between demand in the ECM and the fundamental model are found in the early summer of 2004 where the ECM performs a little worse. We see the same pattern in the winter of 2006, while the ECM performs marginally better than the fundamental model in the summer of 2007 and 2008.

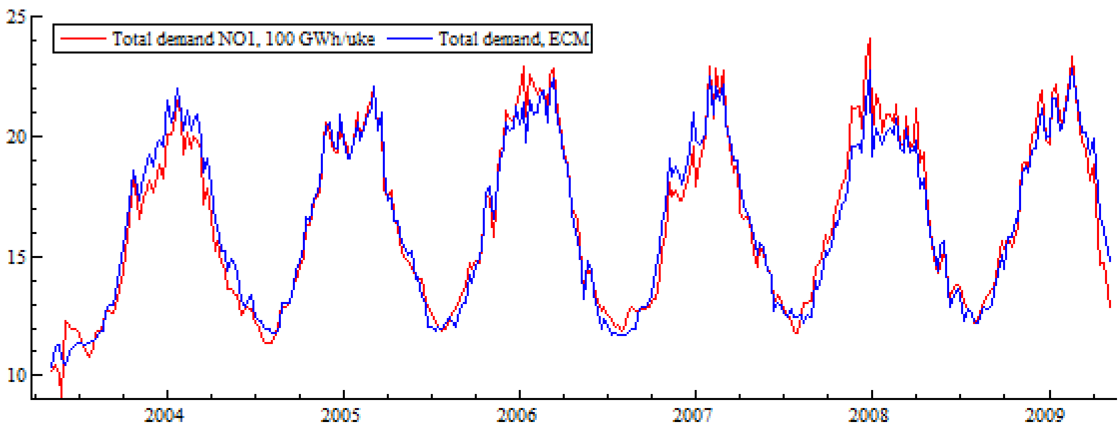

Figure 12: Dynamic simulation of demand using ECM

In the price change equation the lagged value of demand, y_{t-1}, is included along with the lagged value of inflow and the dummy variable for holidays as well as the constant. These parameters improve the out-of-sample performance of the model. However, the change and lagged values of deviation in snowfall are excluded. Compared with the fundamental model, the ECM performs a lot better for the price, compare Figure 13.The ECM simulates the variations to a larger extent, and exhibits a shape that follow the actual spot price pattern in a much better way. Even though the ECM cannot capture the spikes in an adequate manner, smaller increases in the spot price like the one in the winter of 2003, are identified. During the winter of 2006, 2008 and 2009 we see that the actual price increased heavily, and to some extent the ECM manages to mimic the increasing trend during these periods. In the same way as the fundamental model, there is the tendency that the ECM captures falls in the price more accurately than the spikes. In general we can conclude that the model mostly underestimates the amplitude of changes in the price.

[8] The dynamic simulation of the fundamental equations is straightforward, as change in demand does not appear in the price change equation. For the ECMs the change in demand is present in the equation for price change, and vice versa. The dynamic simulation is therefore done by rearranging the equations and solving them through matrix algebra before drawing the graphs.

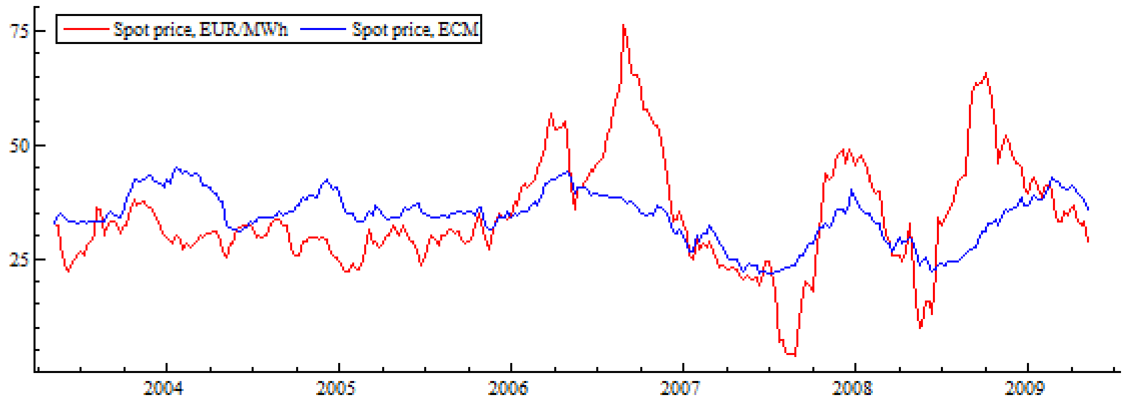

Figure 13: Dynamic simulation of price using ECM

5.3 Extended model

When adding a coal fired producer to the price dynamics in the fundamental model, the highest fit achieved by adjusting the *γ* variable in equation 10, is 0.61 and 0.06 for the change in demand and price, respectively. In other words, the goodness of fit is not improved with the suggested extension to the fundamental model, compare Table 7.

Table 7: Summary statistics for extended model

Dependent variable	Mean	R^2	Sum of squared residuals	Standard error of regression
Δy_t	0.019	0.610	94.21	0.611
Δp_t	-0.061	0.055	1890.51	2.723

6 Discussion

6.1 The fundamental model

When replicated with the dataset from 2003 to 2008, the fit of the fundamental model is relatively poor, especially when it comes to the price change equation. In Table 8 the values for R^2 and the standard error of regression achieved in the replication are compared with the values found in Johnsen (2001).

Table 8: Fit of the fundamental model - replication versus original

	R^2		Standard error of regression	
	Replication	Johnsen (2001)	Replication	Johnsen (2001)
Δy_t	0.63	0.90	0.60	0.45
Δp_t	0.06	0.35	2.71	1.54

The mean square errors (MSE)[9] from dynamic simulation of the fundamental model are given in Table 9. When comparing the MSE from the simulation of 2008-2009 with the results in Johnsen (2001), it is evident that the price model performed far better in the mid-1990s. Interestingly, the

[9] Here we have computed the MSE values in accordance to Johnsen (2001), i.e. MSE is the sum of squared differences between simulated and actual values, divided by the total number of observations. The model with the smallest MSE is generally interpreted as best explaining the variability in the observations.

demand MSE of 3.35 obtained in the replication is lower than in Johnsen (2001). This may appear contradictory to the previous observation that the fit of the demand change is poorer in the replication. However, Taylor (2007) emphasizes that the forecasting method that has the least MSE not always is the method with the highest value of R^2.

Table 9: Mean square errors from dynamic simulation of fundamental model

	Demand (y_t)		**Price (p_t)**	
	Replication	Johnsen (2001)	Replication	Johnsen (2001)
Within sample	0.54	0.76	145.54	7.53
Out of sample	3.35	4.65	960.19	16.99

6.1.1 Time dependent effect of unexpected high inflow

The time dependent effect of unexpected high inflow is reflected by the factor

(11)
$$\frac{1}{(T_1 - t - 1)\alpha_2 - \alpha_3\beta_0}$$

appearing in the price change equation in the fundamental model. The interpretation of this factor is that unexpected high inflow early in the winter season has a small effect on the price since the price fall affects the demand throughout the whole winter. Depending on the time of the year, the price effect will vary in size. As we approach the end of the planning period, an unexpected high inflow will have a greater impact on the price compared with an equal inflow in the beginning of the planning period. This price effect can be shown by simulating unexpected inflows of 100 GWh throughout the winter season, compare Figure 14. Conceptually, this is the same effect as is observed in Johnsen (2001).

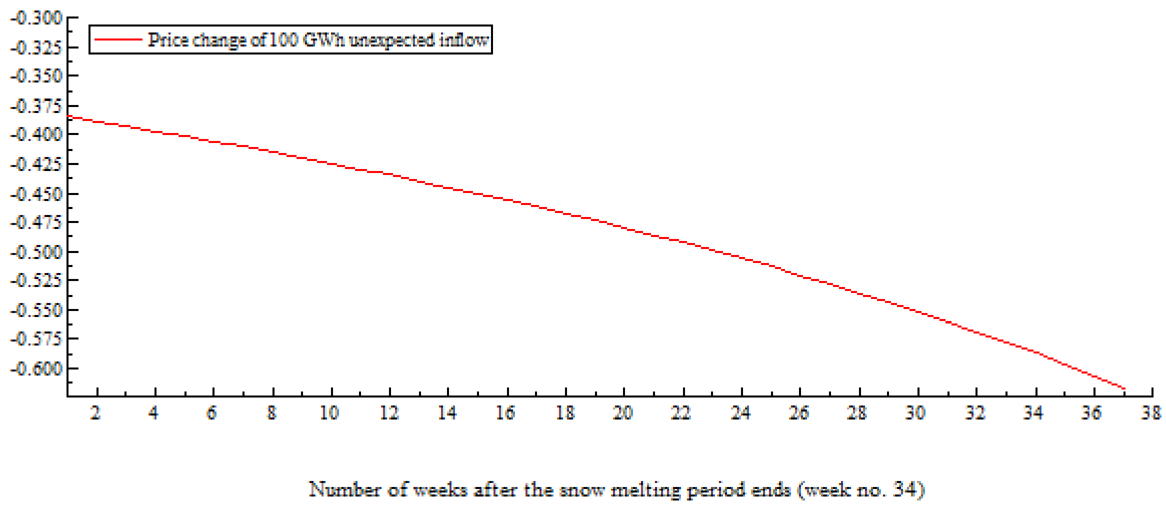

Figure 14: The effect of 100 GWh of unexpected inflow at different times during the planning period

6.1.2 Limitations of the model

In general, bottom-up models incorporate as much information as possible relating to both sides of the demand-supply equation. Consequently, such models are both data and computationally intensive. Even though bottom-up models have high explanatory ability in the context of cause and effect, there is still a range of factors that drive the real-world market volatility that is not captured by such models (Tipping & Read, 2010). In constructing the fundamental model, we make specific assumptions about economic and physical relationships in the marketplace. The price dynamics generated by the models are therefore sensitive to violations of these assumptions. Thus there exists a significant modeling risk in the application of the fundamental approach (Skantze, Gubina, & Ilic, 2000).

General limitations of fundamental and bottom-up models, may partially explain the relatively poor fit of the fundamental model with newer data. In the fundamental model of Johnsen (2001) specific assumptions were made about the relationships in the market, compare section 2.1.1. When applying this model to the current electricity market, it comes as no surprise that the price change equation has relatively little explanatory power. First of all, the assumption about all electricity generation coming from hydro power is an overly simplified approach to the electricity supply side as of today. Even though the NO1 pricing area is dominated by hydro production, market integration has led to increased influence from condensing power plants and the price of coal, gas, nuclear power and CO_2. Transmission constraints between areas further impact the price movements in the respective areas. Secondly, literature suggests that even with extensive bottom-up models, electricity price dynamics is not adequately modeled. Tipping & Read (2010) claim that forecasting performance and modeling application can be improved by combining the bottom-up model with a complementary stochastic factor accounting for the price volatility not captured in the model. Such hybrid models are becoming increasingly popular, see Weron (2006).

A specific limitation to the fundamental model is the poor ability to mimic price change between subsequent planning periods. Johnsen already pointed this out in his paper from 2001, but it first becomes apparent when multiple planning periods are actually included in the estimation, which is exactly the case in our replication. Albeit improved when removing the state equations for price change, the inclusion of more than two planning periods in the estimation reduces the model's fit and predictive ability. Perhaps the most pronounced indicator of this phenomenon is the high absolute value of the estimated coefficient β_0 that results when multiple planning periods are included. A relatively high β_0 will strongly reduce the movements in the price caused by deviations from normal in the fundamental variables, such as inflow, snowfall and temperature. Hence the modest price variations in the dynamic simulation, compare Figure 11, can be explained.

The significance of α_1 in the demand change equation suggests that there is price elasticity of demand. According to Vehviläinen & Pyykkönen (2005) inelastic demand is the norm for the Nordic electricity market in all but the most extreme cases. On the other hand, Holstad & Pettersen (2011) find that if the spot price for Norway increases by 1 % from one month to another, the general electricity consumption falls by 0.05 %. However, they find support in the data that the price elasticity has become lower in absolute value in the last part of their data period, spanning from 1996 to 2010. Since the market is treated as isolated in the fundamental model of Johnsen (2001), the significance of price elasticity found in our replication can to a certain extent be explained by the absence of flow between pricing areas as an explanatory variable in the model. In other words, the support in our data of price elasticity can be attributed to the fact there in reality is flow between interconnected price areas.

As mention earlier, the insignificance of α_2 is troublesome as it is the very coefficient that is multiplied by the time-dependent term in the price change equation, see equation 11. Because it, according to theory, cannot be rejected that α_2 is equal to zero, we may not want to include the coefficient in our model. Hence the time-dependent factor disappears. We find that the out-of-sample ability is better upon inclusion of α_2, but its insignificance is nevertheless a problem with our replicated model.

6.1.3 Limitations in the data

There are limitations in our dataset. The relatively poor performance of the fundamental model can thus partially be explained by shortcomings in the time series for the NO1 pricing area. Generalizing weather variables for an entire price area is a rather complex task, and simplifications are inevitably made. Above all, the modeling of the time series for snowfall is greatly simplified. Due to the excessive workload required to appropriately simulate the snowfall, we made several shortcuts in order to arrive at a reasonable time series, compare section 3.2. It is our hope that the time series calculated for snowfall indeed is proportional to usable inflow. The insignificance of the snowfall parameter in the ECM alternative, suggest that this is not the case.

The derivation of inflow from changes in reservoir levels and production in the NO1 price area assumes no overflow and that all generation comes from one hydro reservoir. Conveniently this is in line with the rationales for Johnsen's fundamental model, but the assumption of one hydro reservoir is deemed realistic due to the preponderance of hydro generation in the NO1 area. What is problematic about this derivation, however, is the fact that inflow will be correlated with production. Thus inflow could partly be considered an endogenous variable in the model[10]. This is unfortunate as inflow is declared an exogenous variable. The ramifications of this relationship are hard to trace in the model, and the consequences of it therefore remain dulled. As no overflow is assumed, water spillage is not accounted for in the fundamental model. This is another methodological weakness since inflow generally is larger than production.

When considering the actual values for spot price in the NO1 area, the all-time high and low in our sample appear in week 34 in subsequent years, that is, 2006 and 2007, respectively. According to the fundamental model, this should imply that there were large deviations from normal in the weather variables at these instances. But no such deviation can be identified from the corresponding time series of weather variables. In fact, very few of the explanatory variables seem to indicate any of the price spikes or falls within the sample. Thus there is no wonder why the model's fit is poor. The fact remains that the spot price in a Norwegian pricing area cannot be explained by looking at data from the Norwegian electricity sector alone. Furthermore, this supports the findings of Tipping & Read (2010) about bottom-up models inability to adequately capture observed price dynamics within hydro dominated electricity markets.

6.2 The ECM alternative

Although not as good as in Johnsen (2001), the price change equation achieves a significantly better fit with the data in the ECM alternative, see Table 10, than in the fundamental approach. In contrast to Johnsen's findings – where the fit remains equivalently good in both the fundamental and the ECM approach – we find that the goodness of fit of the demand change equation is slightly lower in the ECM alternative.

[10] Johnsen (2001) does not comment on inflow as a potential endogenous variable in his original paper, but this should be considered in future work based on his model.

Table 10: Fit of the ECM alternative - replication versus original

	R^2		Standard error of regression	
	Replication	Johnsen (2001)	Replication	Johnsen (2001)
Δy_t	0.58	0.90	0.63	0.45
Δp_t	0.22	0.45	2.50	1.49

The MSEs from dynamic simulation of the ECMs are given in Table 11. Not surprisingly, the ECM for price in Johnsen (2001) performs better than the ECM based on the up-to-date data. The replicated ECM for demand however, seems to perform better than in Johnsen (2001). This is similar to what we observe for the fundamental model, compare section 6.1. A potential explanation to this finding is that we have utilized a larger data sample for the estimation of demand dynamics, than Johnsen (2001) did.

Table 11: Mean square errors from dynamic simulation of ECM alternative

	Demand (y_t)		Price (p_t)	
	Replication	Johnsen (2001)	Replication	Johnsen (2001)
Within sample	0.49	0.89	110.35	8.51
Out of sample	2.94	16.34	915.86	118.08

6.2.1 Limitations of the model

The limitations of the ECMs concern the absence of formal testing of the time series and the relationships between these. We have replicated the approach of Johnsen (2001) regardless of whether the implicit assumptions of the models were valid or not, see section 2.1.4 on stationarity and cointegration. The focus of Johnsen (2001) was to provide a benchmark to the fundamental model. Therefore, the ECMs are formulated to resemble the same market dynamics as in the fundamental model. The same amount of effort is clearly not put into the specification of the ECMs as in the fundamental model. A more consistent treatment of ECMs is given in Johnsen & Willumsen (2010) where different lags for the regressors are included. The use of the natural logarithms in Johnsen & Willumsen (2010) is convenient as the estimated coefficients can be interpreted directly as the elasticity of the parameters in question.

Single equation misspecification tests suggest that the distribution of the residuals is not normal, compare Appendix E. The most efficient estimation method in large samples is FIML, but in order for the FIML estimation to be consistent, the error terms must be normally distributed. The results from the Bera-Jarque tests imply that the FIML technique is not supported. Hence two or three steps least squares (2SLS or 3SLS) should have been considered for estimation purposes. This is in line with the findings of Holstad & Pettersen (2011). An alternative way of getting the residuals more normally distributed is to use the natural logarithms of the parameters.

6.3 Implications from model extension

In an attempt to make the model more robust with regards to today's market conditions, we have extended Johnsen's original model to include production from a coal fired producer as a fundamental explanatory variable. This implementation still recognizes the fundamental dynamics within a hydro dominated electricity sector. From Table 12 we can see that our heuristic approach does not improve the fit of model.

Table 12: Fit of the extended versus fundamental model

	R^2		Standard error of regression	
	Extended	Fundamental	Extended	Fundamental
Δy_t	0.61	0.63	0.611	0.595
Δp_t	0.06	0.06	2.723	2.714

An alternative to the heuristic simplification would be to establish a completely new model in order to incorporate the dynamics of other energy sources. Our calculations show that the incorporation of other energy sources cannot be done without violating the fundamental assumptions introduced by Johnsen (2001), see section 4.3. A new model could also have handled gas and CO_2 prices as these factors will influence the electricity price, compare Burger et al (2007).

Limited transfer capacity between price areas can lead to sudden spikes or falls in the electricity price, and the flow between interconnected price areas will largely depend on the relative prices in these areas. Modeling flow between areas is a rather involved task in the context of Johnsen's work. It would have forced the model to treat flow as an additional endogenous variable, and this would have led to further complexity and more parameters to be fitted in the model. These implications are clearly not desirable if the purpose is to improve the explanatory power of the model. Therefore the system – represented by price area NO1 - is still considered to be isolated with no import or export of electricity. Nonetheless, an extension including flow could have been an interesting subject for further research.

Alternatively, the model could introduce another hydro power reservoir to the system in order to model the market participants in a more realistic way. In the case of an additional hydro power producer, the added producer is typically assumed not to base his production decision on the price dynamics, but rather on the normal level of the hydro reservoirs in the Nordic electricity market. But this extension would violate Johnsen's assumption regarding the shadow prices, compare appendix A, and is therefore not given further considerations.

7 Conclusion

We have replicated the work of Johnsen (2001) with an up-to-date dataset to see if his model for price and demand dynamics is suitable under present market conditions. In the replication we estimate a model for changes in weekly electricity demand and price for the Norwegian pricing area, NO1. Furthermore, we have investigated the possibilities of adding more producers to this hydro dominated system.

Table 13: Summary of fit achieved in fundamental, ECM and extended model

	R^2		
	Fundamental model	ECM approach	Extended model
Δy_t	0.63	0.58	0.61
Δp_t	0.06	0.22	0.06

The price change equation derived in Johnsen (2001) fits the up-to-date data poorly, compare Table 13. The main reason for the poor fit is the profound changes that have taken place in the Norwegian electricity market since the model was originally developed. In today's market, parameters like coal, gas and CO_2 prices, outage of nuclear power plants, flow between interconnected areas, and transmission constraints in the system are likely to affect the price movements. Consequently, short-term deviation from normal weather values will explain the observed price dynamics to a limited extent – even when the system is largely supplied by hydro power. However, it is not possible to adequately incorporate multiple producers in Johnsen's model without violating the key assumptions underlying his hydro based system.

We find that the time-dependent effect of unexpected change in weather parameters is still present. This finding suggests that the same unexpected change in a weather parameter will influence the amplitude of price change differently, depending on how close the change is to the snow melting period. Another finding worth noting is that Johnsen's model is strictly a short-term model. When employed to a larger time interval, the state equations in the original model induce a negative trend that biases the dynamic simulations of the model. The dynamic simulations thus improve upon removal of the state equations.

Finally, we have seen that the short-term performance of the demand model from Johnsen (2001) remains relatively good. This is partly due to the temperature dependency of general consumption in Norway. We find evidence of price elasticity of demand, but this can largely be attributed to the absence of flow between pricing areas as an explanatory variable in the demand equation.

8 Bibliography

Bergström, S. (1992). *The HBV model - its structure and applications.* Norrköping: SMHI Reports.

Burger, M., Graeber, B., & Schindlmayr, G. (2007). *Managing energy risk : an integrated view on power and other energy markets.* Chichester, NJ: John Wiley & Sons.

Davison, M., Anderson, C., Marcus, B., & Anderson, K. (2002). Development of a hybrid model for electrical power spot prices. *IEEE Transactions on Power Systems Vol(17)*, 257-264.

Doornik, J., & Hendry, D. (2009). *PcGive 13 Volume II Modelling Dynamic Systems.* London: Timberlake Consultants Press.

Engle, R., & Granger, C. (1987). Co-Integration and Error Correction: Representation, Estimation and Testing. *Econometrica Vol(55)*, 251-276.

Fleten, S.-E., & Lemming, J. (2003). Constructing forward price curves in electricity markets. *Energy Economics Vol(25)*, 409-424.

Førsund, F. (2007). *Hydropower Economics.* New York: Springer Science.

Greene, W. (2002). *Econometric Analysis.* New Jersey: Prentice Hall.

Johnsen, T. (2001). Demand, generation and price in the Norwegian market for electric power. *Energy Economics Vol(23)*, 227-251.

Johnsen, T., & Willumsen, M. (2009). *Electricity demand and temperature - an empirical methodology for temperature adjustment in Norway.* Retrieved 10 1, 2011, from Norwegian Water Resources and Energy Directorate: http://www.nve.no/Global/Publikasjoner/Publikasjoner%202010/Rapport%202010/rapport2010_01.pdf?epslanguage=no

Lucia, J., & Schwartz, E. (2002). Electricity prices and power derivatives: Evidence from the Nordic Power Exchange. *Review of derivatives research Vol(5)*, 5-50.

Pettersen, F. (2011, 10 5). Norwegian Water Resources and Energy Directorate. (K.-H. Bakke, Interviewer)

Pettersen, F., & Holstad, M. (2011). *Hvordan reagerer strømforbruket i alminnelig forsyning på endringer i spotpris? [In Norwegian].* Oslo: Statistics Norway.

Quantitative Micro Software. (2009). *EViews 7 User's Guide II.* Irvine: QMS.

Skantze, P., Gubina, A., & Ilic, M. (2000). *Bid-based stochastic model for electricity prices: the impact of fundamental drivers on market dynamics.* Massachusetts : Energy Laboratory Publications MIT EL 00-004, Massachusetts Institute of Technology.

Skjerpen, T. (2011, 10 7). Statistics Norway. (K.-H. Bakke, Interviewer)

Taylor, S. (2007). *Asset Price Dynamics, Volatility, and Prediction.* New Jersey: Princeton University Press.

Tipping, J., & Read, E. (2010). *Hybrid Bottom-Up/Top-Down Modeling.* Christchurch: Energy Modelling Research Group, University of Canterbury.

Vehviläinen, I., & Pyykkönen, T. (2005). Stochastic factor model for electricity spot. *Energy Economics Vol(27)*, 351– 367.

Ventosa, M., Baillo, A., Ramos, A., & Rivier, M. (2005). Electricity market modeling trends. *Energy Policy Vol(33)*, 897-913.

Weron, R. (2006). *Modeling and forecasting electricity loads and prices: a statistical approach.* Chichester, NJ: John Wiley & Sons.

Weron, R., & Misiorek, A. (2005). Forecasting spot electricity prices with time series. *The European Electricity Market EEM-05*, (pp. 133-141). Lodz.

Appendix A: Derivation of the simultaneous equations

The water reservoir budget is

(A1) $$r_j \leq r_{j-1} + I_j - y_j,\ j = t,...$$

where r is the water reservoir filling measured in energy units, and y is the production of hydropower in energy units. Overflow gives inequality in the above equation. Since inflow consists of both rainfall and snow melting, we have

(A2) $$I_j = I_j^W + \psi_j s_{j-1},\ j = t,...$$

where I^W is the inflow from rainfall in energy units, ψ is a variable for snow melting, and s is the snow volume measured in energy units.

The snow budget is then

(A3) $$s_j = s_{j-1} + I_j^s - \psi_j s_{j-1},\ j = t,...$$

Where I^s is the inflow of snow, or snowfall, in energy units.

The upper and lower reservoir filling constraints are given by

(A4) $$r_j \leq \overline{r},\ j = t,...$$

and

(A5) $$r_j \geq \underline{r},\ j = t,...$$

A constraint on weekly increase in reservoir level is expressed as

(A6) $$r_j - r_{j-1} \leq c,\ j = t,...$$

where c is the maximum weekly increase in the reservoir filling.

An upper level on weekly generation capacity is deduced from the maximum generation capacity, k, measured in energy units

(A7) $$y_j \leq \overline{k},\ j = t,...$$

The terminal value, V*, depends on water and snow volumes left at the end of the planning horizon

(A8) $$V^* = Ar_{T_1} + Bs_{T_1} + C(r_{T_1} + \gamma_s s_{T_1})^2$$

where A, B, C and γ_s are coefficients. At time t the complete value is

(A9) $$V_t = \sum_{j=t}^{T_1} E_t \int_0^{r_{j-1}+I_j-r_j} p(q)dq + E_t V^*$$

Here E_t indicates the expectation made in week t. It is evident from the above equation that no water overflow is assumed.

The storage of week T_1 is determined through backward induction and differentiation of

(A10) $$V_{T_1} = \int_0^{r_{T_1-1}+I_{T_1}-r_{T_1}} p(q)dq + Ar_{T_1} + Bs_{T_1} + C(r_{T_1} + \gamma_s s_{T_1})^2 - \lambda_{T_1,5}(r_{T_1} - \bar{r}) - \lambda_{T_1,6}(\underline{r} - r_{T_1}) - \lambda_{T_1,7}(r_{T_1} - r_{T_1-1} - c) - \lambda_{T_1,8}(r_{T_1-1} + I_{T_1} - r_{T_1} - \bar{k})$$

where λ_i is the Lagrange relaxation, or shadow prices, of the above constraints on the reservoir. Using the fundamental theorem and the core rule for derivation yields

(A11) $$\frac{\partial V_{T_1}}{\partial r_{T_1}} = -p(r_{T_1-1} + I_{T_1} - r_{T_1}) + A_{T_1} + 2C(r_{T_1} + \gamma_s s_{T_1}) - \lambda_{T_1,5} + \lambda_{T_1,6} - \lambda_{T_1,7} + \lambda_{T_1,8} = 0$$

If the constraints are non-binding , the first order condition for maximization is

(A12) $$p_{T_1} = A_{T_1} + 2C(r_{T_1} + \gamma_s s_{T_1})$$

which can be written as a terminal condition for the reservoir

(A13) $$r_{T_1} = \beta_0 p_{T_1} + \beta_1 s_{T_1} + \beta_2$$

For week T_1*-1* the storage is found by maximization of

(A14) $$V_{T_1-1} = \sum_{j=T_1-1}^{T_1} E_{T_1-1}\int_0^{r_{j-1}+I_j-r_j} p(q)dq + E_{T_1-1}V^* - \lambda_{T_1-1,5}(r_{T_1-1} - \bar{r}) - \lambda_{T_1-1,6}(\underline{r} - r_{T_1-1}) - \lambda_{T_1-1,7}(r_{T_1-1} - r_{T_1-2} - c) - \lambda_{T_1-1,8}(r_{T_1-2} + I_{T_1-1} - r_{T_1-1} - \bar{k})$$

with respect to r_{T1-1} and by using the fundamental theorem we the core rule for derivation we get

(A15) $$\frac{\partial V_{T_1-1}}{\partial r_{T_1-1}} = -p(r_{T_1-2} + I_{T_1-1} - r_{T_1-1}) + E_{T_1-1}\left[p(r_{T_1-1} + I_{T_1} - r_{T_1})\right] - \lambda_{T_1-1,5} + \lambda_{T_1-1,6} - \lambda_{T_1-1,7} + \lambda_{T_1-1,8} = 0$$

Generally we will get

(A16) $$p(r_{j-1} + I_j - r_{j-1}) = E_j\left[p(r_j + I_{j+1} - r_{j+1})\right] - \lambda_{j,5} + \lambda_{j,6} - \lambda_{j,7} + \lambda_{j,8}$$

and if the shadow prices are assumed to be zero, we have

(A17) $$p_t = E[p_{t+1}] = E[p_{t+2}] \ldots = E[p_{T_1}]$$

which implies

(A18) $$\Delta p = E_t[p_{T_1}] - E_{t-1}[p_{T_1}]$$

In the derivation of the price change equation the six following equations are combined

(A19) $$r_{T_1} = r_t + \sum_{j=t+1}^{T_1} (I_j - y_j)$$

(A20) $$I_t = I_t^W + \psi_t s_{t-1}$$

(A21) $$s_{T_1} = s_t + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})$$

(A22) $$r_{T_1} = \beta_0 p_{T_1} + \beta_1 s_{T_1} + \beta_2$$

(A23) $$p_t = E[p_{t+1}] = ... = E[p_{T_1}]$$

(A24) $$\Delta y = \alpha_0 + \alpha_1 \Delta p_t + \alpha_2 p_{t-1} + \alpha_3 y_{t-1} + \alpha_4 \Delta w_t + \alpha_5 w_{t-1} + \alpha_6 \Delta \tau_t + \alpha_7 \tau_{t-1} + \varepsilon_t$$

We insert A20 in A19

(A25) $$r_{T_1} = r_t + \sum_{j=t+1}^{T_1} (I_j^W + \psi_j s_{j-1} - y_j)$$

Equation A21 is then inserted into A22

(A26) $$r_{T_1} = \beta_0 p_{T_1} + \beta_1 (s_t + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})) + \beta_2$$

The two equations above, A25 and A26, are equalized

(A27) $$r_t + \sum_{j=t+1}^{T_1} (I_j^W + \psi_j s_{j-1} - y_j) = \beta_0 p_{T_1} + \beta_1 (s_t + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})) + \beta_2$$

We find an expression for y_j using

(A28) $$\Delta y_t = y_t - y_{t-1}$$

so that

(A29) $$y_t = y_{t-1} + \alpha_0 + \alpha_1 \Delta p_t + \alpha_2 p_{t-1} + \alpha_3 y_{t-1} + \alpha_4 (w_t - w_{t-1}) + \alpha_5 w_{t-1} + \alpha_6 (\tau_t - \tau_{t-1}) + \alpha_7 \tau_{t-1} + \varepsilon_t$$

The lag operator *L* is introduced

(A30) $$y_t = \alpha_0 + \alpha_1 \Delta p_t + \alpha_2 p_{t-1} + (1+\alpha_3) y_{t-1} + \alpha_4 w_t + (\alpha_5 - \alpha_4) w_{t-1} + \alpha_6 \tau_t + (\alpha_7 - \alpha_6) \tau_{t-1} + \varepsilon_t$$

and

(A31) $$y_t = \alpha_0 + \alpha_1 \Delta p_t + \alpha_2 p_{t-1} + (1+\alpha_3) L y_t + (\alpha_4 + (\alpha_5 - \alpha_4) L) w_t + (\alpha_6 + (\alpha_7 - \alpha_6) L) \tau_t + \varepsilon_t$$

Then we solve this for y_t

(A32) $y_t - (1+\alpha_3)Ly_t = \alpha_0 + \alpha_1 \Delta p_t + \alpha_2 p_{t-1} + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_t + (\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_t + \varepsilon_t$

and

(A33) $y_t(1-(1+\alpha_3)L) = \alpha_0 + \alpha_1 \Delta p_t + \alpha_2 p_{t-1} + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_t + (\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_t + \varepsilon_t$

We thus get

(A34) $y_t = \dfrac{1}{(1-(1+\alpha_3)L)}(\alpha_0 + \alpha_1 \Delta p_t + \alpha_2 p_{t-1} + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_t + (\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_t + \varepsilon_t)$

Now we can insert A34 in A27, and we add the expression for r_t to get the following

(A35)
$$
\begin{aligned}
& r_{t-1} + I_t^W + \psi_t s_{t-1} - y_t + \sum_{j=t+1}^{T_1} \left[\begin{aligned} & I_j^W + \psi_j s_{j-1} - \frac{1}{(1-(1+\alpha_3)L)}(\alpha_0 + \alpha_1 \Delta p_j + \alpha_2 p_{j-1} \\ & + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_j + (\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_j + \varepsilon_j) \end{aligned} \right] \\
& = \beta_0 p_{T_1} + \beta_1 (s_t + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})) + \beta_2
\end{aligned}
$$

We solve for p_t

(A36)
$$
\begin{aligned}
& \beta_0 p_{T_1} + \sum_{j=t}^{T_1} \frac{1}{(1-(1+\alpha_3)L)}(\alpha_1 \Delta p_j + \alpha_2 p_{j-1}) = r_{t-1} + I_t^W + \psi_t s_{t-1} \\
& - \frac{1}{(1-(1+\alpha_3)L)} \left\{ \begin{aligned} & (\alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_t \\ & + (\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_t + \varepsilon_t \end{aligned} \right\} \\
& + \sum_{j=t+1}^{T_1} \left[\begin{aligned} & I_j^W + \psi_j s_{j-1} - \frac{1}{(1-(1+\alpha_3)L)}(\alpha_0 + (\alpha_4 + \\ & (\alpha_5 - \alpha_4)L)w_j + (\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_j + \varepsilon_j) \end{aligned} \right] \\
& - \beta_1 (s_t + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})) - \beta_2
\end{aligned}
$$

We then multiply both sides with the denominator in the second term on the right hand side

(A37)
$$
\begin{aligned}
& \beta_0 p_{T_1}(1-(1+\alpha_3)L) + \sum_{j=t}^{T_1} (\alpha_1 \Delta p_j + \alpha_2 p_{j-1}) = (r_{t-1} + I_t^W + \psi_t s_{t-1})(1-(1+\alpha_3)L) \\
& - \left\{ \begin{aligned} & (\alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_t \\ & + (\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_t + \varepsilon_t \end{aligned} \right\} + \sum_{j=t+1}^{T_1} \left[\begin{aligned} & (1-(1+\alpha_3)L)(I_j^W + \psi_j s_{j-1}) - (\alpha_0 + (\alpha_4 + \\ & (\alpha_5 - \alpha_4)L)w_j + (\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_j + \varepsilon_j) \end{aligned} \right] \\
& - \beta_1 (1-(1+\alpha_3)L)(s_t + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})) - (1-(1+\alpha_3)L)\beta_2
\end{aligned}
$$

Then we expand the lag expression for price

$$\beta_0 \Delta p_{T_1} - \beta_0 \alpha_3 p_{T_1-1} + \sum_{j=t}^{T_1} (\alpha_1 \Delta p_j + \alpha_2 p_{j-1}) = (r_{t-1} + I_t^W + \psi_t s_{t-1})(1-(1+\alpha_3)L)$$

$$\text{(A38)} \quad -\begin{Bmatrix} (\alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_t \\ +(\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_t + \varepsilon_t \end{Bmatrix} + \sum_{j=t+1}^{T_1} \begin{bmatrix} (1-(1+\alpha_3)L)(I_j^W + \psi_j s_{j-1}) - (\alpha_0 + (\alpha_4 + \\ (\alpha_5 - \alpha_4)L)w_j + (\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_j + \varepsilon_j) \end{bmatrix}$$

$$-\beta_1 (1-(1+\alpha_3)L)(s_t + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})) - (1-(1+\alpha_3)L)\beta_2$$

Now we take the expectation on both sides of A38

$$E_t[\beta_0 \Delta p_{T_1} - \beta_0 \alpha_3 p_{T_1-1} + (T_1 - t - 1)(\alpha_1 \Delta p_{T_1} + \alpha_2 p_{T_1-1})] = E_t[(r_{t-1} + I_t^W + \psi_t s_{t-1})(1-(1+\alpha_3)L)$$

$$\text{(A39)} -\begin{Bmatrix} (\alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_t \\ +(\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_t + \varepsilon_t \end{Bmatrix} + \sum_{j=t+1}^{T_1} \begin{bmatrix} (1-(1+\alpha_3)L)(I_j^W + \psi_j s_{j-1}) - (\alpha_0 + (\alpha_4 + \\ (\alpha_5 - \alpha_4)L)w_j + (\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_j + \varepsilon_j) \end{bmatrix}$$

$$-\beta_1 (1-(1+\alpha_3)L)(s_t + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})) - (1-(1+\alpha_3)L)\beta_2]$$

Using equation A23 gives us

$$E_t[p_{T_1-1}](T_1 - t - 1)\alpha_2 - \beta_0 \alpha_3) = E_t[p_{T_1}](T_1 - t - 1)\alpha_2 - \beta_0 \alpha_3)$$

$$\text{(A40)} \quad = E_t \begin{bmatrix} (1-(1+\alpha_3)L) \begin{Bmatrix} r_{t-1} + I_t^W + \psi_t s_{t-1} + \sum_{j=t+1}^{T_1} (I_j^W + \psi_j s_{j-1}) \\ -\beta_1 (s_t + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})) - \beta_2 \end{Bmatrix} \\ -\begin{Bmatrix} (\alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_t \\ +(\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_t + \varepsilon_t \end{Bmatrix} - \sum_{j=t+1}^{T_1} \begin{bmatrix} \alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_j \\ +(\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_j + \varepsilon_j \end{bmatrix} \end{bmatrix}$$

If we also include A21 we get the following

$$E_t[p_{T_1}] = \frac{1}{(T_1 - t - 1)\alpha_2 - \beta_0 \alpha_3)}$$

$$\text{(A41)} \quad E_t \begin{bmatrix} (1-(1+\alpha_3)L) \begin{Bmatrix} r_{t-1} + I_t^W + \psi_t s_{t-1} + \sum_{j=t+1}^{T_1} (I_j^W + \psi_j s_{j-1}) \\ -\beta_1 (s_{t-1} + I_t^s - \psi_t s_{t-1} + \sum_{j=t+1}^{T_1} (I_j^s - \psi_j s_{j-1})) - \beta_2 \end{Bmatrix} \\ -\sum_{j=t}^{T_1} \begin{bmatrix} \alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_j \\ +(\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_j + \varepsilon_j \end{bmatrix} \end{bmatrix}$$

Finally, using that the expected value of the error term is zero, we arrive at

(A42)

$$E_t[p_{T_1}] = \frac{1}{(T_1 - t - 1)\alpha_2 - \beta_0\alpha_3}$$

$$\left[\begin{array}{l} (1-(1+\alpha_3)L)\left\{ \begin{array}{l} r_{t-1} + I_t^W + \psi_t s_{t-1} + \sum_{j=t+1}^{T_1}(E_t I_j^W + E_t \psi_j s_{j-1}) - \beta_2 \\ -\beta_1(s_{t-1} + I_t^s - \psi_t s_{t-1} + \sum_{j=t+1}^{T_1}(E_t I_j^s - E_t \psi_j s_{j-1})) \end{array} \right\} \\ -\left\{ \begin{array}{l} (\alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_t \\ +(\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_t + \varepsilon_t \end{array} \right\} - \sum_{j=t+1}^{T_1}\left[\begin{array}{l} \alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_j \\ +(\alpha_6 + (\alpha_7 - \alpha_6)L)E_t\tau_j \end{array} \right] \end{array} \right]$$

which is exactly the result of Johnsen (2001).

In a similar way we derive the expression

(A43)

$$E_{t-1}[p_{T_1}] = \frac{1}{(T_1 - t - 1)\alpha_2 - \beta_0\alpha_3}$$

$$\left[\begin{array}{l} (1-(1+\alpha_3)L)\left\{ \begin{array}{l} r_{t-1} + \sum_{j=t}^{T_1}(E_{t-1} I_j^W + E_{t-1} \psi_j s_{j-1}) - \beta_2 \\ -\beta_1(s_{t-1} + \sum_{j=t}^{T_1}(E_{t-1} I_j^s - E_{t-1} \psi_j s_{j-1})) \end{array} \right\} \\ -\sum_{j=t}^{T_1}\left[\alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_j + (\alpha_6 + (\alpha_7 - \alpha_6)L)E_{t-1}\tau_j \right] \end{array} \right]$$

From equation A18 we have that

(A44)

$$\Delta p = E_t[p_{T_1}] - E_{t-1}[p_{T_1}] = \frac{1}{(T_1 - t - 1)\alpha_2 - \beta_0\alpha_3}$$

$$\left[\begin{array}{l} (1-(1+\alpha_3)L)\left\{ \begin{array}{l} r_{t-1} + I_t^W + \psi_t s_{t-1} + \sum_{j=t+1}^{T_1}(E_t I_j^W + E_t \psi_j s_{j-1}) - \beta_2 \\ -\beta_1(s_{t-1} + I_t^s - \psi_t s_{t-1} + \sum_{j=t+1}^{T_1}(E_t I_j^s - E_t \psi_j s_{j-1})) - r_{t-1} \\ -\sum_{j=t}^{T_1}(E_{t-1} I_j^W + E_{t-1} \psi_j s_{j-1}) + \beta_2 + \beta_1(s_{t-1} + \sum_{j=t}^{T_1}(E_{t-1} I_j^s - E_{t-1} \psi_j s_{j-1})) \end{array} \right\} \\ -\left\{ \begin{array}{l} (\alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_t \\ +(\alpha_6 + (\alpha_7 - \alpha_6)L)\tau_t + \varepsilon_t \end{array} \right\} - \sum_{j=t+1}^{T_1}\left[\begin{array}{l} \alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_j \\ +(\alpha_6 + (\alpha_7 - \alpha_6)L)E_t\tau_j \end{array} \right] + \sum_{j=t}^{T_1}\left[\begin{array}{l} \alpha_0 + (\alpha_4 + (\alpha_5 - \alpha_4)L)w_j \\ +(\alpha_6 + (\alpha_7 - \alpha_6)L)E_{t-1}\tau_j \end{array} \right] \end{array} \right]$$

Note the term

(A45) $$(1+\beta_1)\left[\psi_t s_{t-1} - E\psi_t s_{t-1}\right] = 0$$

and that

(A46) $$E_t I_j^s = E_{t-1} I_j^s, E_t I_j^w = E_{t-1} I_j^w, E_t \tau_j = E_{t-1} \tau_j$$

so we get price change equal to

(A47) $$\Delta p = \frac{1}{(T_1 - t - 1)\alpha_2 - \beta_0 \alpha_3}\begin{bmatrix}(1-(1+\alpha_3)L)\left\{(I_t^w - EI_t^w) - \beta_1(I_t^s - EI_t^s)\right\} \\ -(\alpha_6 + (\alpha_7 - \alpha_6)L)(\tau_t - E\tau_t) - \varepsilon_t\end{bmatrix} + u_t$$

which the reader should recognize as the price change equation presented in Johnsen (2001).

Appendix B: Grid search procedure

Table 14: Grid search procedure

	Δy_t		Δp_t		Probability	
K-value	R^2	**SER**	R^2	**SER**	α_1	α_2
22.0	0.3993	0.7553	0.0513	2.7387	0.0096	0.2898
24.0	0.4967	0.6913	0.0659	2.7176	0.0033	0.3304
26.0	0.5723	0.6373	0.0754	2.7037	0.0057	0.3276
28.0	0.5722	0.6373	0.0754	2.7037	0.0057	0.3276
30.0	0.5669	0.6413	0.0758	2.7031	0.0067	0.3858
Infinite	0.6267	0.5953	0.0609	2.7141	0.0461	0.3283

As we can see from the table above, the fit of both demand and price change, tend to improve when the K-value increases. We note that α_1 is only significant on a 5 % level when *K* is infinite, whereas α_1 is significant on a 1 % level for lower values of *K*. However, both the grid search procedure above and the plot of the dynamic simulation, compare Figure 11, imply that the exclusion of the state equations for price change, is a valid undertaking.

Appendix C: Fitted values, cross-plots and scaled residuals

The plots of actual and fitted values of demand and price change in the fundamental model, are given in the figures below.

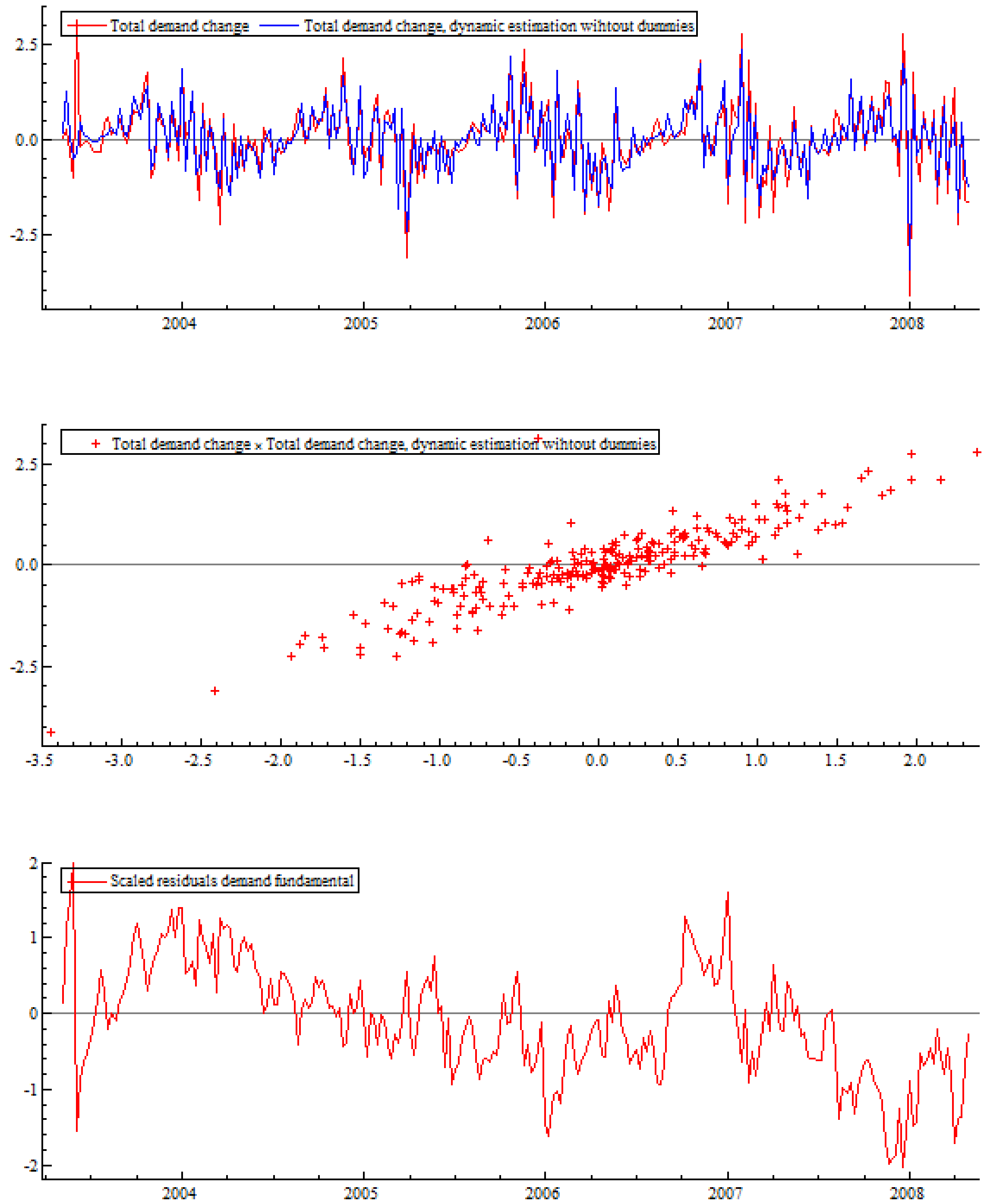

Figure 15: Plots of actual and fitted value of demand change - fundamental model

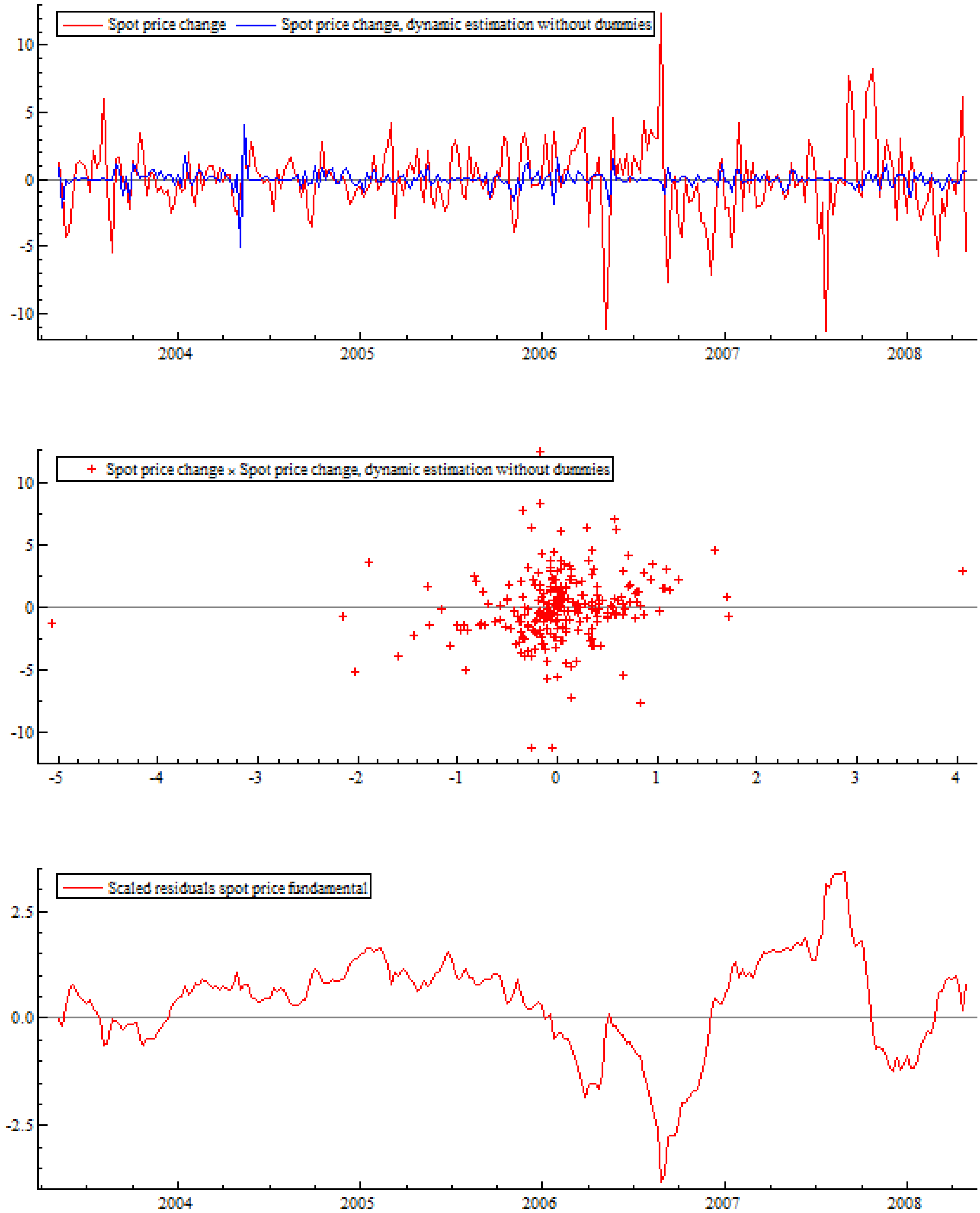

Figure 16: Plots of actual and fitted value of price change - fundamental model

The plots of actual and fitted values of demand and price change in the ECM approach, are given in the figures below.

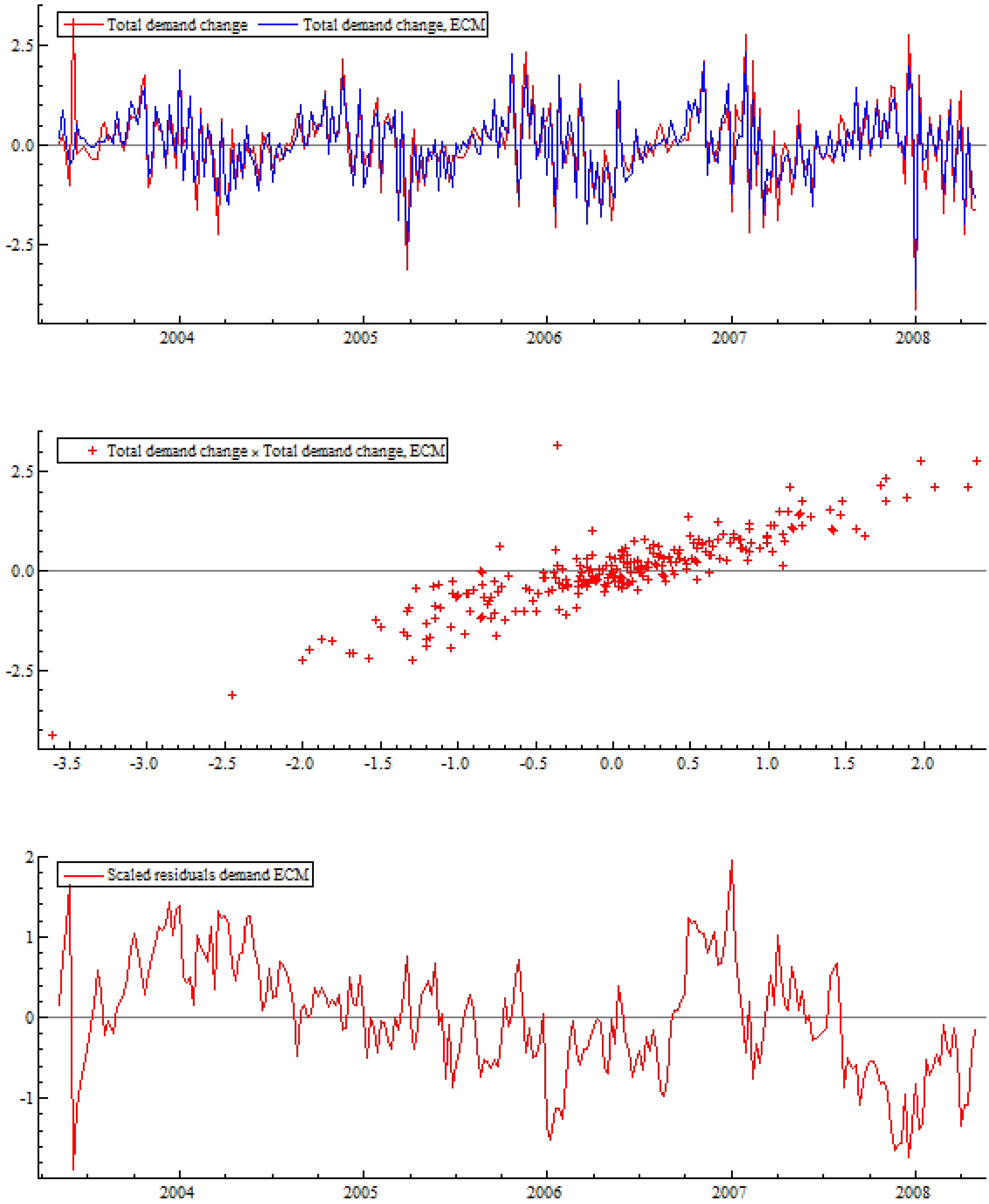

Figure 17: Plots of actual and fitted value of demand change - ECM approach

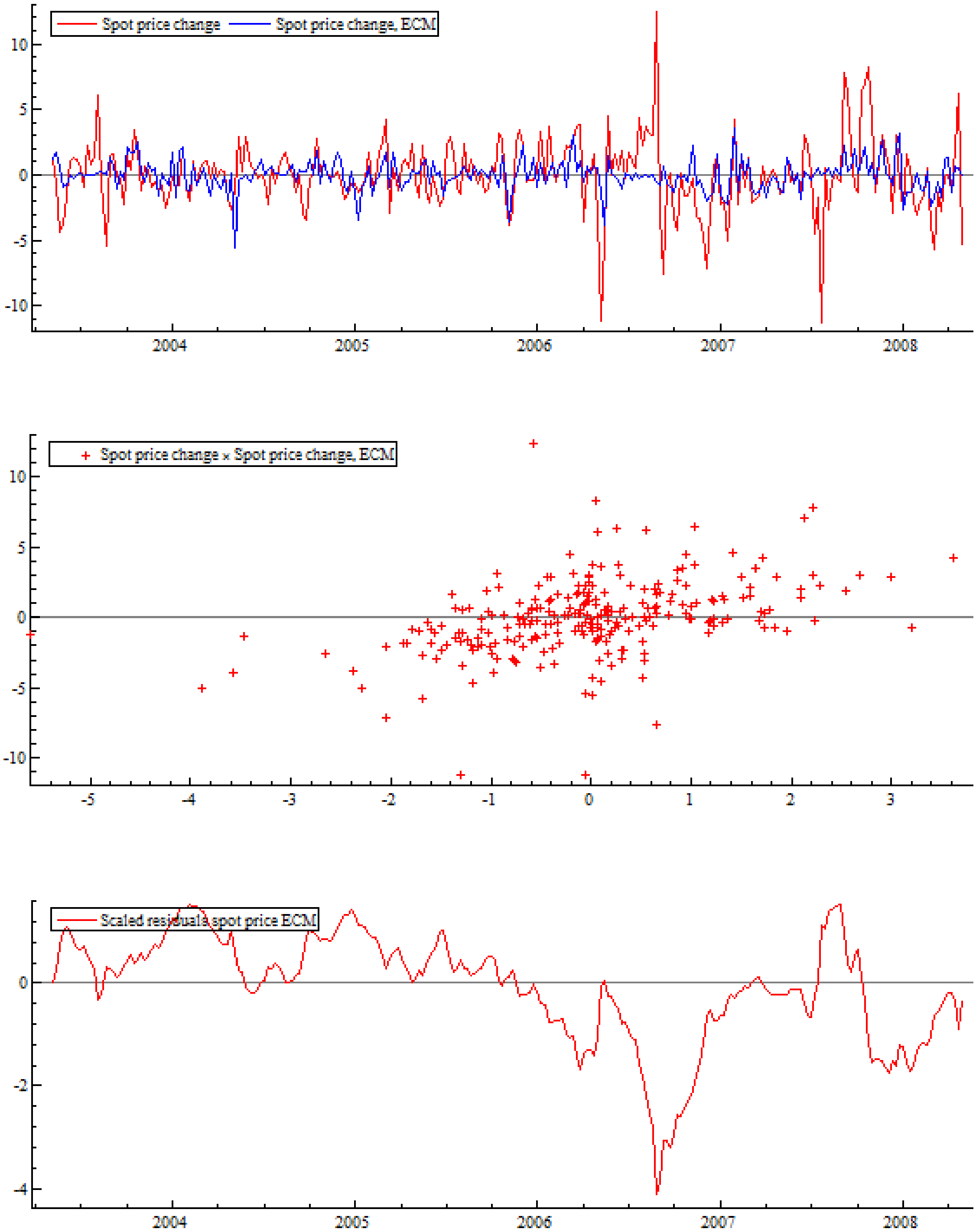

Figure 18: Plots of actual and fitted value of price change - ECM approach

Appendix D: Presentation of data for model extension

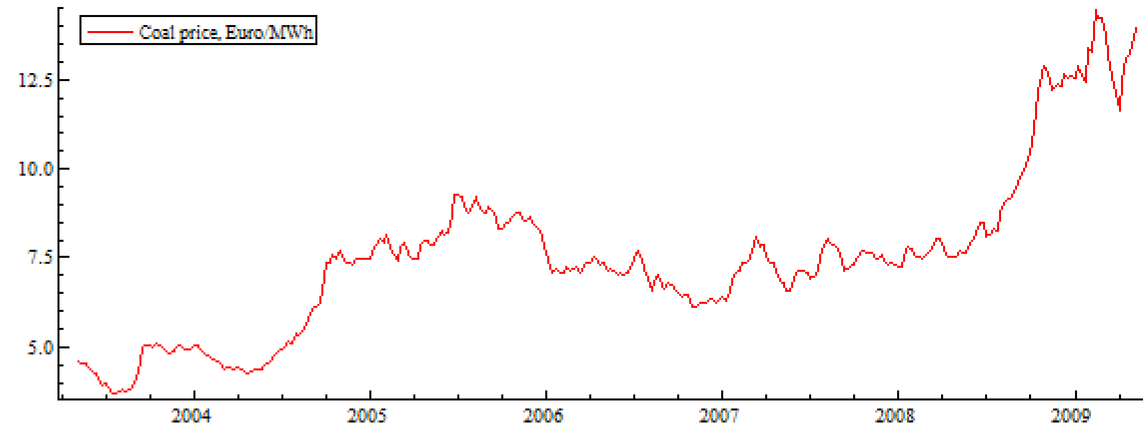

Figure 19: Coal price

Appendix E: Single equations misspecification tests

Table 15: Single equation misspecification test - ECM

Demand change equation		
Autocorrelation	$F_{AR\ 1\text{-}7}(7,240)$	1.45
Autoregression conditional heteroscededasticity	$F_{ARCH\ 1\text{-}7}(7,247)$	0.18
Normality	$X^2_{NORM}(2)$	100.35**
Heteroscedasticity	$F_{HET}(25,235)$	2.80**
Price change equation		
Autocorrelation	$F_{AR\ 1\text{-}7}(7,240)$	7.69**
Autoregression conditional heteroscededasticity	$F_{ARCH\ 1\text{-}7}(7,247)$	2.99**
Normality	$X^2_{NORM}(2)$	153.97**
Heteroscedasticity	$F_{HET}(25,235)$	1.40

*A significance level of 5 % is indicated by **.*