NTNU
Norwegian University of Science and Technology

Fundamental risk analysis and VaR forecasts of the Nord Pool system price

Martin Lundby
Kristoffer Uppheim

Industrial Economics and Technology Management
Submission date: May 2011
Supervisor: Sjur Westgaard, IØT

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

Fundamental risk analysis and VaR forecasts of the Nord Pool system price

Martin Lundby and Kristoffer Uppheim

The Norwegian University of Science and Technology
Trondheim, Norway*

May 30, 2011

*This is our master thesis in *Risk Modeling of Energy Markets* for *The Institute of Industrial Economics and Technology Management* at *The Norwegian University of Science and Technology, NTNU.* Sjur Westgaard is our supervisor. Contact information; Martin Lundby, martinlu@stud.ntnu.no; Kristoffer Uppheim, uppheim@stud.ntnu.no.

Abstract

This paper compares the Value at Risk (VaR) forecasting performance of different quantile regression models to conventional GARCH specifications on the Nord Pool system price. The sample covers hourly data from 2005-2011. In order to identify significant explanatory variables, we use a linear quantile regression to characterize the effects of fundamental factors on the system price formations. From our analysis we are able to show how the sensitivity of the variables change over the range of price quantiles and detect how these sensitivities vary over the hours of the day. Our findings suggest that the demand forecast and the price volatility is the most important determinants of the price in the tails of the distribution. We use these variables in the further analysis and test the out-of-sample VaR performance of linear quantile regression, exponentially weighted quantile regression (EWQR) and conditional autoregressive value at risk (CAViaR) models on the system price. We extend the CAViaR models to account for asymmetrical response to returns and are innovative in including explanatory variables in the CAViaR specification. Our results show that the I-GARCHX CAViaR model with demand forecast as explanatory variable outperform the other models, and that CAViaR models in general perform well. The linear quantile regression with price volatility as explanatory variable also provides good results. The computational complexity of CAViaR models favors a linear quantile regression, so market participants have to make a tradeoff between the level of accuracy in the forecasts and the complexity of the model. Our findings are useful for producers, consumers and traders, as well as clearinghouses, as they provide an accurate measure of the price risk.

Keywords: Nord Pool; Value at Risk; Quantile Regression; EWQR; CAViaR

Contents

1 Introduction

The liberalization of the electricity markets has fundamentally changed the way power companies do business. Vertical integration characterizes the current structure, and retail sales has been separated from production. Due to fierce competition in the market, pressure has been put on improved operational efficiency and cost reduction. More competition has increased the price volatility and the use of complex contract portfolios, enhancing the need of an accurate assessment of the exposure to price risk. Value at Risk (VaR) is the most prominent measure of risk in the financial industry. VaR puts a single number on the potential change of an asset/portfolio value, over a determined time horizon. Increasing the accuracy of VaR forecasts is interesting for two reasons; it will improve risk management and reduce costs.

In this paper we compare the day-ahead out-of-sample forecasting performance of well-known parametric and semi-parametric VaR models, and our own extensions of these. Based on a thorough fundamental analysis we introduce an I-GARCHX CAViaR model with demand forecast as explanatory variable, and find that this model outperforms its peers. The CAViaR models generally perform well in modeling the VaR. A linear quantile regression with price volatility also provides good forecasts and has the advantage of not being as complex as the CAViaR models.

The explanatory variables are selected using linear quantile regression to characterize the effect of fundamental factors over the entire range of Nord Pool system price quantiles. We find that factors such as demand forecast and price volatility are important determinants of the price, with stronger relations in the tails than in the interior of the distribution. The relation between the price and demand forecast is especially strong in the tails of the intraday price, with opposite effects during peak periods and off-peak periods. The price volatility shows significant effects for all periods analyzed, however of smaller magnitude than the demand forecast. Other factors such as hydrological balance, lagged system prices, gas and coal price, wind production forecasts, temperature and inflow, all show significant effects of smaller magnitude.

In the out-of-sample analysis our parametric models, GARCH-N and skew-t GARCH-GJR, are included for benchmarking purposes. We refer to Bollerslev (1986) for a general GARCH introduction, and Glosten et al. (1993) for the GARCH-GJR extension. For an example of these models used on power markets, see Escribano et al. (2002).

The semi parametric models include both linear and nonlinear quantile regressions. Koenker and Basset (1978) first introduced the linear quantile regression. It models the quantile directly with no distributional assumptions, which is an advantage when the distribution is unknown or time varying. We implement two linear quantile regression models, one using demand forecast, and one using price volatility as explanatory variable. Quantile regression using an intercept and no explanatory variables gives an unconditional VaR estimate and we therefore exclude this from our analysis. The exponentially weighted quantile regression (EWQR) introduced by Taylor (2008) provides a conditional VaR estimate using only a constant as independent variable. We include this EWQR in our out-of-sample VaR models.

Manganelli and Engle (2004) prove that the performance measure applied in linear quantile regression is applicable to nonlinear quantile models. They propose a new nonlinear quantile regression specification; Conditional Autoregressive Value at Risk (CAViaR). We find that CAViaR specification is theoretically promising for two main reasons; like linear quantile regression it makes no distributional assumption, and its autoregressive nature respond well to volatility clustering. Manganelli and Engle (2004) provide four examples of specific CAViaR models. Out of these we have chosen to work with the Indirect GARCH (I-GARCH) CAViaR model. We extend this model to account for explanatory variables.[1] The motivation behind this are the findings of Contreras et al. (2003) and Garcia et al. (2005). They find that a GARCH model with demand as the explanatory variable outperforms general time series ARIMA models in day-ahead forecasting. Our fundamental analysis using linear quantile regression supports this conclusion. Therefore, we extend the I-GARCH CAViaR model to include demand forecast as explanatory variable. Based on Manganelli and Engle's asymmetric slope model we also extend the I-GARCH CAViaR specifications to respond asymmetrically to returns. We re-estimate the parameters for each run to make our results closer to what one might expect from real world implementation.

This paper is structured as follows: section 2 presents background on the Nord Pool market setting and section 3 describes the data set. In section 4 we provide an overview of the models used in this paper. The results where we describe the effect of fundamental factors on the price and compare the different out-of-sample models are found in section 5. In section 6 we present our conclusions. The appendix at the end of the paper includes additional theory and results.

[1]We also extended the model to account for autoregressive (AR) time series (continuing the work of Kuester et al. (2006)). The motivation behind the AR extension is the highly autoregressive nature of power market prices. However, we experience a significant drop in performance and choose not to include these models in our results.

2 Background

2.1 The Nord Pool market setting

The liberalization of the Nordic energy market was initiated by the Norwegian government and the Energy Act of 1990. Following trends in Chile, UK and other European countries, the Energy Act formed the basis for the deregulation of the Nordic countries, opening the electricity sector to competition. The new structure was characterized by the unbundling of the previously vertically integrated activities such as generation, transmission and retail sales. Market participants, in a larger extend than earlier, recognized the value in the retail end of the supply chain.

Established in 1993, Nord Pool started as an electricity pool covering the Norwegian market only. In 1996 it became the world's first multinational exchange for electricity contracts, as Sweden was included in the market. In the succeeding years, Finland joined the market in 1998, Western Denmark in 1999 and Eastern Denmark in 2004. In recent years Nord Pool has been launched in Germany and Estonia. Today, Nord Pool is the largest electricity market in the world and most of the consumption of electricity in the Nordic countries is traded through the exchange.

Energy source	Denmark	Finland	Norway	Sweden	Sum	Share [%]
Wind power	6.7	0.3	1.0	2.5	10.5	2.8
Other renewables	2.4	8.2	0.0	11.1	21.7	5.9
Fossil fuels	25.3	24.9	3.5	4.8	58.5	15.8
Nuclear power	0.0	22.6	0.0	50.0	72.6	19.6
Hydro power	0.0	12.6	128.3	65.3	206.2	55.7
Non-identifiable	0.0	0.6	0.0	0.0	0.7	0.2
Total production	34.5	69.2	132.8	133.7	370.2	100.0

Table 1: The Nord Pool electricity production by country and energy source. Numbers in TWh per year, except right column. Source: www.nordpoolspot.com.

The total production in the Nordic area during 2009 was 370 TWh. Nord Pool had a turnover of 288 TWh, a market share of 77%, representing a value of EUR 10.8 billion. Table 1 shows the production from various energy sources across the Nordic countries. Hydropower alone is 56% of the total production. Nuclear power is 20% of total production while fossil fuels and renewables are 16% and 9% respectively.

Nord Pool organizes three different markets, the day-ahead market (Elspot), the cross border intraday market (Elbas)[2] and the financial market. The real time or balancing market is organized by the transmission system operators (TSO) for short term upward or downward regulation. Both demand-side and supply-side bids are posted in the balancing market, stating prices and volumes. These prices are known to the market first after the hour of delivery and are highly volatile. Positions in the balancing market therefore carries high risk.

Elspot is the spot market where hourly power contracts trade daily for physical delivery in the next 24-hour period. This is a non-mandatory day-ahead market. The demand and supply curve is aggregated from bids and offers made by market participants and the intersection point between these curves sets the system price. The system price is the price the market is willing to pay for electricity for the given hour with no capacity constrains (bottlenecks) in the market. Market participants place their bids and offers each morning for the different hours the following day. Trading ends 12:00 (noon), called gate closure, and prices are published 14:00 in the afternoon. The system price is also the reference price in settlements at Nord Pool's financial market.

In order to handle bottleneck situations, Nord Pool is geographically divided into bidding areas. Bottlenecks between these areas are managed using pricing mechanisms in the spot market. As a result prices must be adjusted and zonal prices are calculated besides the system price. Internal bottlenecks is managed within the bidding area and handled by the respective TSO.

Nord Pool's financial market is a commercial center for trading of financial contracts. In 2009 the financial market had a turnover of 1218 TWh. The high turnover compared to the physical market underlines the importance of derivatives as a hedging tool in electricity markets and the need of an accurate assessment of price risk.

[2]Elbas is a continuous cross border intraday market that trades from the time where the day-ahead prices are published until one hour prior to delivery. This market fills the gap between the day-ahead market and the balancing market and allows the market participants to adjust their market exposure. Elbas gives the participants access to the whole Nordic, German and Estonian market and reduces the risk in the balancing market. It also provides opportunity for power consumers to sell back power bought in the day-ahead market. Elbas is a relatively small marked with a turnover of 2.4 TWh compared to Elspot's 286 TWh in 2009.

3 Data

3.1 Variable description

We obtain Nord Pool spot prices from Nord Pool Spot's historical data reports.[3] It includes 24 hourly data points of the system price for each day, seven days a week. The prices are stated in Euro per MWh and cover the period January 1st 2005 to January 23rd 2011, totaling 53136 data points over 2214 days.[4] This is a satisfactory sample size for both in- and out-of-sample analysis of the outer 1% and 99% quantiles.

Nord Pool uses the arithmetic average of the hourly prices in a day as the reference price in the cash-settlement calculations of derivative contracts in the financial market. Due to this we calculate the arithmetic average of all intraday prices and refer to this price as the daily average price. We also examine the intraday effects by analyzing three periods of the daily 24, namely period 4, 9 and 18. As figure 1 shows, these periods represent off-peak (period 4), super-peak (period 9) and peak time (period 18). We have chosen these periods as we want to observe how the effects of the fundamental factors vary over the different periods of the day, expecting them to be clearest in the extremes, and test the forecasting models on time series with different characteristics. We will refer to these time series as period 4, period 9 and period 18 respectively. This leaves us with four time series to study.

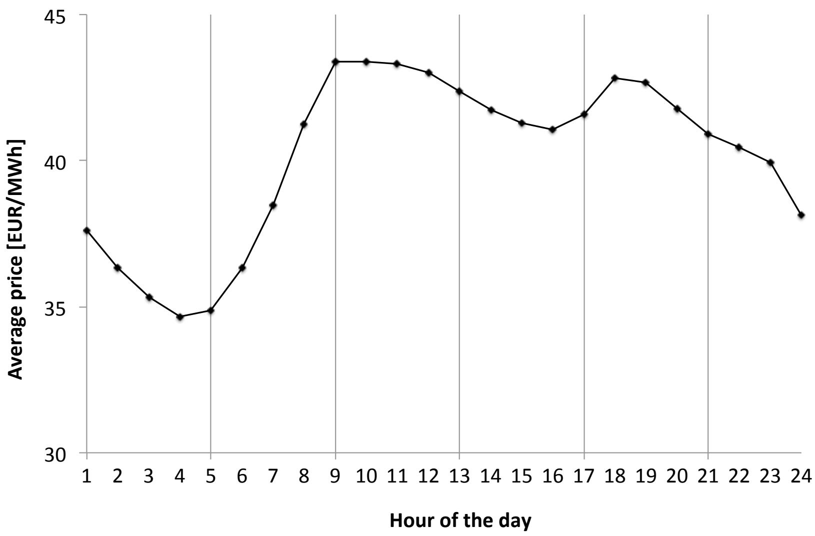

Figure 1: Intraday average prices for the whole sample period (2005-2011).

[3]Source: www.nordpoolspot.com.

[4]We have chosen not to use earlier data as Eastern Denmark was included in 2004, potentially changing the price structure.

In addition to the spot price, the total intraday data set includes actual demand and wind production forecast. For factors such as gas price, coal price, hydrological balance, temperature- and inflow forecasts, we have daily data. All price data in other currencies are converted to Euro using the respective exchange rate for the day in question. In cases of missing data or extreme outliers we use linear interpolation. For gas and coal prices this is done for all weekends.

We calculate hourly demand forecasts using feed-forward neural networks[5] with the following inputs; temperature forecast, one-day (24 hour) lagged demand, last days average demand, seven-day (168 hour) lagged demand, day of week [1:7], and weekend/holiday dummy. The parameters are found using 1000 in-sample data, and we do not re-estimate the model. This is a simplified version of the forecasting model proposed by Malki et al. (2004), a well-known method for short term load forecasting. For theory on neural networks, and its use in load forecasting we refer to Liu et al. (1996) and Weron (2006), as this is beyond the scope of this paper.

Another explanatory variable in the analysis is the price volatility. This is calculated separately for each of the periods and for the daily average price, using a skew-t GARCH-GJR(1,1) model on the price returns. GARCH modeling is well known and tested, leading us to choose this fairly simple and accurate estimate of volatility. We refer to section 7.1 in the appendix for a more detailed description of all the variables used in this paper.

The Box-Cox power test (Box and Cox, 1964) indicates a variance stabilization transformation by taking the natural logarithm of the spot prices. This is performed on all prices and variables, making it possible to interpret the coefficients in the regression as elasticities. See appendix section 7.2 for results from the Box-Cox test.

3.2 Descriptive statistics

The descriptive statistics in table 2 reveals that for both price, log price, change in price and log returns we have skewed time series with high sample kurtosis and volatility. It is significant difference in skewness, kurtosis and price range between the four series analyzed. This indicates quite erratic behavior of the system price. The changes are large both within the day, with an average difference of EUR 11 per MWh between period 4 and 9, and over the entire period with a difference between lowest and highest price of close to EUR 300 per MWh.

[5]The network is a two-layer feed-forward network with 20 sigmoid hidden neurons and linear output neurons trained with the Levenberg-Marquardt backpropagation algorithm in Matlab.

	Descriptive Statistics								
	No. Obs.	Mean	Median	Minimum	Maximum	St. Dev.	Skewness	Kurtosis	Jerque-Bera
Period 4									
P_t	2215	34.66	33.11	1.00	81.63	13.98	0.49	3.38	102
$P_t - P_{t-1}$	2214	0.02	0.04	-31.52	31.07	3.88	-0.35	20.37	27867
$\ln P_t$	2215	3.45	3.50	0.00	4.40	0.49	-1.59	8.33	3554
$\ln P_t - \ln P_{t-1}$	2214	0.00	0.00	-3.31	2.62	0.23	-1.94	53.81	239390
Period 9									
P_t	2215	43.39	41.03	2.78	330.03	17.86	3.49	42.58	149055
$P_t - P_{t-1}$	2214	0.02	-0.50	-254.24	228.30	12.77	-0.55	151.71	2039331
$\ln P_t$	2215	3.70	3.71	1.02	5.70	0.39	-0.44	5.93	865
$\ln P_t - \ln P_{t-1}$	2214	0.00	-0.01	-2.39	2.44	0.23	0.65	22.22	34201
Period 18									
P_t	2215	42.83	40.67	9.80	181.83	16.89	1.79	10.15	5894
$P_t - P_{t-1}$	2214	0.02	-0.26	-76.07	115.33	8.21	1.41	50.18	205960
$\ln P_t$	2215	3.69	3.71	2.28	5.21	0.37	0.04	3.57	30
$\ln P_t - \ln P_{t-1}$	2214	0.00	-0.01	-0.91	1.08	0.14	0.97	18.06	21272
Daily Average									
P_t	2215	40.11	38.00	8.80	134.80	14.13	0.92	4.79	609
$P_t - P_{t-1}$	2214	0.02	-0.25	-47.57	52.79	3.87	0.99	40.90	132830
$\ln P_t$	2215	3.63	3.64	2.17	4.90	0.36	-0.30	3.58	65
$\ln P_t - \ln P_{t-1}$	2214	0.00	-0.01	-0.71	0.84	0.10	1.03	16.56	17340

The Jerque-Bera test statistic is a goodness-of-fit measure of departure from normality based on the sample kurtosis and skewness. The statistic has an asymptotic chi-squared distribution with two degrees of freedom and the null hypothesis is that the data are from a normal distribution. Critical value at 5% is 5.99.;

Table 2: Descriptive statistics for period 4, 9, 18 and the daily average price.

	Autocorrelation coefficients of lag												
	1	2	3	4	5	6	7	14	21	28	35	ADF	Q(5)
Period 4													
P_t	0.961	0.948	0.935	0.923	0.912	0.900	0.892	0.818	0.758	0.707	0.649	-4.02	9960
$P_t - P_{t-1}$	-0.324	-0.005	-0.018	-0.014	0.017	-0.055	0.104	0.049	0.048	0.103	0.071	-34.04	235
$\ln P_t$	0.887	0.852	0.824	0.807	0.788	0.770	0.769	0.675	0.608	0.560	0.513	-6.88	7662
$\ln P_t - \ln P_{t-1}$	-0.342	-0.036	-0.047	0.006	-0.001	-0.077	0.091	0.115	0.071	0.091	0.066	-37.79	267
Period 9													
P_t	0.744	0.650	0.666	0.642	0.603	0.636	0.704	0.655	0.647	0.556	0.515	-9.56	4857
$P_t - P_{t-1}$	-0.316	-0.216	0.079	0.028	-0.138	-0.070	0.295	0.270	0.360	0.227	0.223	-41.85	382
$\ln P_t$	0.818	0.725	0.721	0.706	0.694	0.750	0.842	0.797	0.752	0.696	0.661	-8.71	5958
$\ln P_t - \ln P_{t-1}$	-0.245	-0.243	0.028	-0.009	-0.184	-0.103	0.542	0.528	0.514	0.457	0.437	-40.62	340
Period 18													
P_t	0.882	0.819	0.805	0.798	0.786	0.802	0.815	0.770	0.743	0.657	0.612	-7.19	7405
$P_t - P_{t-1}$	-0.234	-0.207	-0.030	0.020	-0.121	0.018	0.209	0.215	0.301	0.134	0.166	-40.62	251
$\ln P_t$	0.931	0.889	0.868	0.860	0.862	0.877	0.892	0.851	0.808	0.752	0.717	-6.08	8599
$\ln P_t - \ln P_{t-1}$	-0.203	-0.144	-0.100	-0.064	-0.097	-0.005	0.353	0.359	0.363	0.260	0.283	-38.93	189
Daily Average													
P_t	0.963	0.932	0.921	0.913	0.906	0.912	0.919	0.863	0.816	0.755	0.700	-4.45	9487
$P_t - P_{t-1}$	-0.095	-0.261	-0.029	-0.031	-0.161	-0.015	0.389	0.369	0.407	0.324	0.302	-36.10	232
$\ln P_t$	0.959	0.927	0.913	0.903	0.901	0.913	0.926	0.871	0.820	0.772	0.728	-4.84	9362
$\ln P_t - \ln P_{t-1}$	-0.113	-0.216	-0.060	-0.086	-0.161	-0.025	0.502	0.489	0.476	0.408	0.387	-36.58	214

ADF is the augmented Dickey-Fuller test for unit root with a constant and no trend including two augmented lags. Critical value for stationarity at 5% level is -2.86.; *Q(5)* is the Ljung-Box portmanteau Q-statistic for autocorrelation up to lag 5. The asymptotic distribution of Q(k) is chi-squared with k degrees of freedom. For k=5 the critical value at 5% level is 11.07 and the null hypothesis is that the data are random.

Table 3: Autocorrelation coefficients for period 4, 9, 18 and the daily average price.

As we can see from the standard deviation in table 2, the spot prices are highly volatile. The standard deviation of the daily average log-returns is 0.10. This translates to an annualized volatility of 195%.[6] The same value for the super-peak period 9 is 443%. We observe a clear difference between the periods analyzed, with period 4 and 9 being the most volatile.

The skew coefficients vary across the time series and we observe both positive and negative coefficients for price change, log returns and log prices. All skew coefficients for the spot price are positive. This effect is as anticipated for electricity markets and reveals that extreme price outliers occur on the upside of the average. We also observe positive skew for the log-return in period 9 and 18, and negative skew for the log-return in period 4. This implies that extreme absolute returns are more likely to be positive than negative in the peak periods 9 and 18, with the opposite effect in the off-peak period 4.

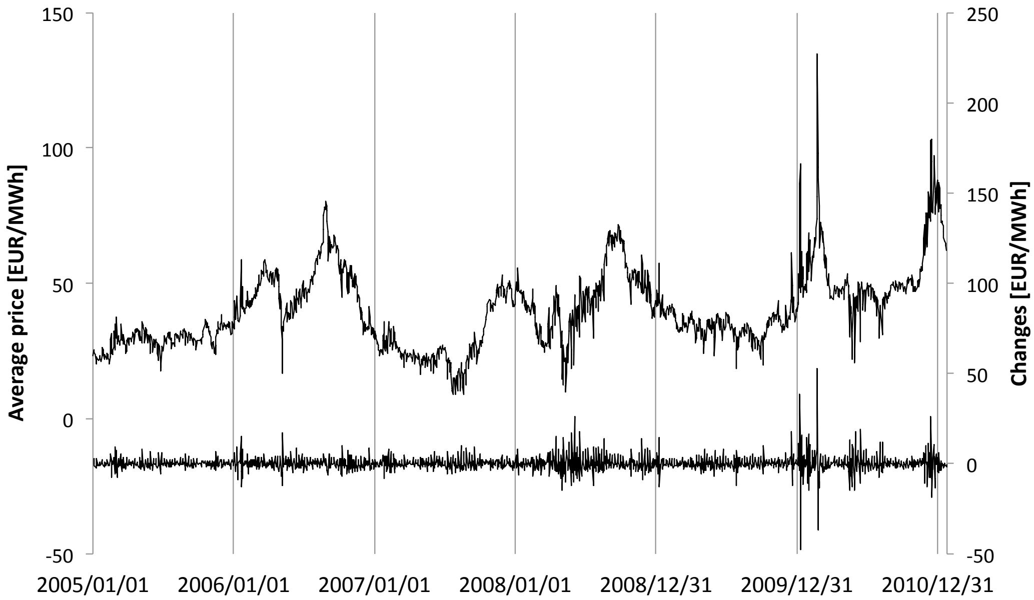

Figure 2: Daily average system price. The figure shows the average daily price (top) and changes in average price (bottom) for the whole period (2005-2011).

Extreme prices are common in the Nord Pool spot market. This is found in the kurtosis coefficients, as all values are way above three, which is the value for a normal distribution. Non-normality is also confirmed by the Jerque-Bera statistic. The kurtosis for the daily average price is 4.8, while the same value for the daily average log-return is 16.6. This indicates that spikes and jumps are present in the spot prices as extreme

[6]This is obtained by multiplying the standard deviation with the square root of 365.

positive or negative returns have high probability of occurrence. Lucia and Schwartz (2002) find that most of these extreme events occur during the cold season. The largest spike in price is observed on February 22nd 2010 with an increase of 228 EUR/MWh from the previous day. This was during one of many extreme cold periods during the winter of 2009/2010. Most of the larger jumps occur due to temporary shocks in the demand, frequently linked to sudden and pronounced changes in temperature (Lucia and Schwartz, 2002). Figure 2 shows that prices generally are higher during winter and that the largest spikes also occur during this season.

Another observation from figure 2 is the volatility clustering in the price changes. We clearly observe periods of large changes followed by periods of small changes. The large changes occur during winter and spring and are results of shocks in demand and uncertainty related to snow melting and precipitation.

To test for stationarity we use the augmented Dickey-Fuller test for unit root (see table 3). It shows that all series are stationary for two lags at the 5% significance level. From this we conclude that both prices and returns are stationary.

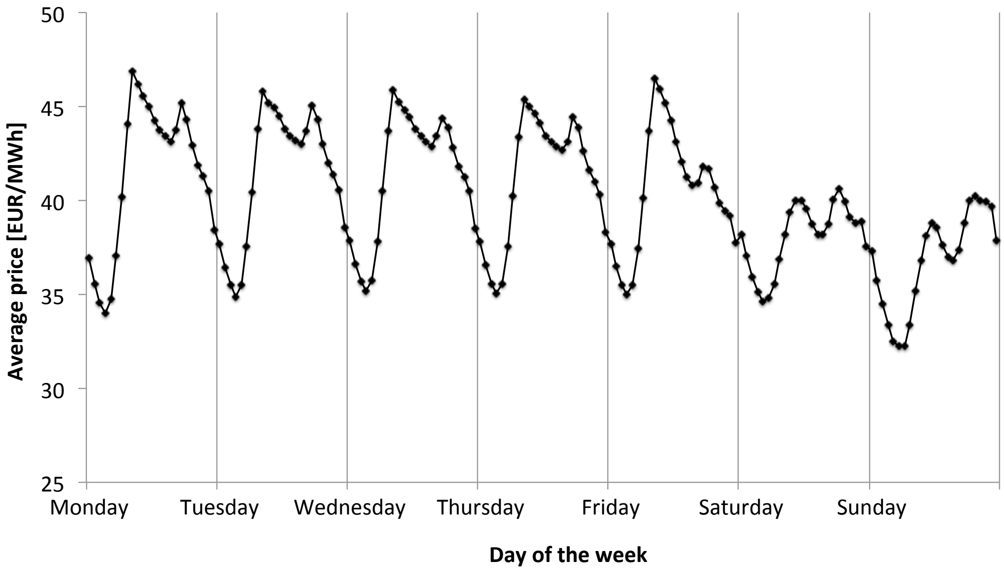

Figure 3: Hourly average price patterns. The figure shows the average intraday prices throughout the week for the whole period (2005-2011).

The autocorrelation coefficients and the Q-statistic in table 3 clearly indicate that there is severe serial correlation in the data. This is as expected, especially for the prices. For the changes in price the test statistics are generally lower, but all are well above the 5% critical value, so we conclude that the price returns are also serially correlated. These features of the spot price display signs of predictability. All autocorrelation coefficients of lag one are positive for the price (log-price) and negative for the change in price (log-return). For lags multiple of seven, both the price (log-price) and change in price (log-returns) have a positive coefficient. This is due to consistent intra-day and intra-week price patterns mainly determined by business activity (figure 3). Note the significant lower prices during weekends. Yearly price patterns (figure 2) are also present in the time series due to seasonal variations in temperature and reservoir inflow.

The descriptive statistics reveals quite erratic behavior of the system price with a time varying and complex distribution. This motivates an analysis of advanced time series techniques such as linear and nonlinear quantile regressions.

4 Models

In this section we present the models applied in the paper. The first model is used for variable selection, and examines how explanatory variables influence the spot price over its range of quantiles. For this we make use of a linear quantile regression. Secondly we describe the models used to forecast VaR out-of-sample. This includes linear quantile regression models with different explanatory variables, one EWQR model, and different CAViaR specifications. We compare the performance of these models to two GARCH models; the GARCH-N(1,1) and the skew-t GARCH-GJR(1,1), both with ARMA(7,7) filtering. We refer to section 7.3 in the appendix for a more thorough description of the GARCH models implemented. To backtest the performance of the out-of-sample analysis we use the conditional and unconditional coverage test, and the dynamic quantile (DQ) test.

4.1 Linear quantile regression

Introduced by Koenker and Basset (1978) quantile regression seeks to estimate conditional quantile functions. Its asymptotic behavior is not as favorable as least squares, so it is solved as a minimization problem. This was known to be time consuming, leading to a lack of interest in the method. However, the linear programming technique suggested by Koenker and D'Orey (1987), and the interior point method for linear programming by Koenker and Park (1996) and Koenker et al. (1997) has made it comparable to least squares in computation. Hence, the quantile regression approach is competitive in VaR estimation.

The variable selection model is a linear quantile regression with a set of explanatory variables, defined as $VaR_t(\theta) = \mathbf{x}'_t\boldsymbol{\beta}(\theta)$. VaR_t is the quantile value, $\mathbf{x}_t$ is a vector of explanatory variables and $\boldsymbol{\beta}(\theta)$ is a vector of parameters for different values of θ, $0 < \theta < 1$. Equation (1) presents the model with explanatory variables. See appendix 7.1 for variable description.

$$\begin{aligned} VaR_t(\theta) = \; & \beta_0(\theta) + \beta_1(\theta) Price_{t-1} + \beta_2(\theta) AvePrice_{t-7} + \beta_3(\theta) PriceVol \\ & + \beta_4(\theta) GasPrice + \beta_5(\theta) CoalPrice + \beta_6(\theta) DemandForecast \\ & + \beta_7(\theta) WindForecast + \beta_8(\theta) HydBal + \beta_9(\theta) DevTemp \\ & + \beta_{10}(\theta) DevInflow + \beta_{11}(\theta) Weekend \end{aligned} \tag{1}$$

The sample quantile $VaR_t(\theta)$ is defined as the solution to the minimization problem:

$$RQ = \min_{VaR_t \in \mathbb{R}} \left[\sum_{t \in \{t: y_t \geq VaR_t\}} \theta |y_t - VaR_t| + \sum_{t \in \{t: y_t \leq VaR_t\}} (1-\theta)|y_t - VaR_t| \right] \tag{2}$$

Where y_t is the price/return and $\theta = 1/2$ gives the least absolute error estimate, i.e. the regression median. We minimize the RQ criterion (2) on the natural logarithm of prices with the explanatory variables in (1) to characterize the sensitivities of the variable coefficients over the quantiles. The quantile regression is run in Stata and we include a regular OLS regression for reference purposes.

The out-of-sample analysis utilizes the same model to forecast VaR. The only difference from the first model is the number of explanatory variables and the use of log returns instead of log prices, making the methodology identical. We forecast using the betas found on data up to time t. This makes VaR_{t+1} estimation using quantile regression quite straightforward. We refer to the quantile regression models as Qreg.

$$V\hat{a}R_{t+1}(\theta) = \mathbf{x'}_{t+1} \hat{\boldsymbol{\beta}}(\theta) \tag{3}$$

Where $\hat{\boldsymbol{\beta}}(\theta)$ is found solving (2) with values ranging $[1:t]$.

4.2 Exponentially weighted quantile regression

Taylor (2008) introduced Exponentially Weighted Quantile Regression (EWQR), also referred to as discounted quantile regression. For a specified value of the weighting parameter, λ, the RQ criterion in (2) takes the form:

$$\min_{VaR_t \in \mathbb{R}} \left[\lambda^{T-t} \Bigg(\sum_{t \in \{t: y_t \geq VaR_t\}} \theta |y_t - VaR_t| + \sum_{t \in \{t: y_t \leq VaR_t\}} (1-\theta)|y_t - VaR_t| \Bigg) \right] \tag{4}$$

With $0 < \lambda < 1$ this method puts a higher weight on more recent data than data in the past. Taylor proves that this minimization problem has the same features as the linear quantile regression. Hence, the method of forecasting VaR_{t+1} will be equivalent to the linear quantile regression in equation (3). We will refer to this model as EWQR.

4.3 Conditional Autoregressive Value at Risk

Engle and Manganelli (2004) note that the quantile is tightly linked to the variance. As the variance has shown autoregressive properties, they propose a conditional autoregressive specification that models the VaR as an autoregressive process and call the model Conditional Autoregressive Value at Risk (CAViaR).

$$VaR_t(\boldsymbol{\beta}) = \beta_0 + \sum_{i=1}^{q} \beta_i VaR_{t-i}(\boldsymbol{\beta}) + \sum_{j=1}^{r} \beta_j l(\mathbf{x}_{t-j}), \tag{5}$$

Where $p = q + r + 1$ is the dimension of $\boldsymbol{\beta}$ and l is a function of a finite number of lagged values of observables.

They suggest four different CAViaR specifications. We use their Indirect-GARCH (I-GARCH) model (6),

$$VaR_t(\beta) = (\beta_0 + \beta_1 VaR_{t-1}^2(\beta) + \beta_2 y_{t-1}^2)^{1/2}, \tag{6}$$

and extend the model to account for asymmetrical response to returns. We call this model I-GJR-GARCH CAViaR (7).

$$VaR_t(\beta) = (\beta_0 + \beta_1 VaR_{t-1}^2(\beta) + \beta_2 y_{t-1}^2 + \beta_3 I_{(y_{t-1}<0)} y_{t-1}^2)^{1/2}, \tag{7}$$

The main innovation in this paper is the extension of the CAViaR model to respond to fundamental market factors. To include the significant relations linking explanatory variables to power market tail price behavior, we adjust the model to account for explanatory variables. We call these models I-GARCHX CAViaR and I-GJR-GARCHX CAViaR.

$$VaR_t(\beta) = (\beta_0 + \beta_1 (VaR_{t-1} - \beta_3 z_{t-1})^2 + \beta_2 y_{t-1}^2)^{1/2} + \beta_3 z_t, \tag{8}$$

$$VaR_t(\beta) = (\beta_0 + \beta_1 (VaR_{t-1} - \beta_4 z_{t-1})^2 + \beta_2 y_{t-1}^2 + \beta_3 I_{(y_{t-1}<0)} y_{t-1}^2)^{1/2} + \beta_4 z_t, \tag{9}$$

Where z_t is an explanatory variable known at time $(t-1)$. Note that the explanatory variable models the quantile and not the return. This is deliberate as we have found stronger relations to fundamental factors in the tails. If this is not the case, we recommend using a model where the explanatory variable is linked to the return. This is exemplified in section 7.4 in the appendix.

4.4 Backtesting models

"The coverage probability of a confidence interval is the proportion of the time that the interval contains the true value of interest."[7] Hence, for a θ tail of a distribution the coverage probability is θ. Kupiec's test for unconditional coverage is a likelihood ratio test where the null is an accurate forecast (Kupiec, 1995).

$$LR_{uc} = \frac{\pi_{exp}^{n_1}(1-\pi_{exp})^{n_0}}{\pi_{obs}^{n_1}(1-\pi_{obs})^{n_0}} \tag{10}$$

where π_{exp} is the expected proportion of exceedances, π_{obs} is the observed proportion of exceedances, n_1 is the observed number of exceedances and $n_0 = n - n_1$ where n is the sample size of the backtest. The asymptotic distribution of $-2ln(LR_{uc})$ is chi-squared with one degree of freedom (Alexander, 2008).

A flaw of the unconditional coverage test is that it does not punish a model for successive exceedances of the predicted VaR. A proper modeling of the VaR would result in the exceedances following a Bernoulli process. The conditional coverage test developed by Christoffersen (1998) is able to incorporate this by also testing for clustering of exceedances. Hence, the test combines a check for autocorrelation and unconditional coverage. The null is that the forecast is accurate and there is no clustering of exceedances.

$$LR_{cc} = \frac{\pi_{exp}^{n_1}(1-\pi_{exp})^{n_0}}{\pi_{01}^{n_{01}}(1-\pi_{01})^{n_{00}}\pi_{11}^{n_{11}}(1-\pi_{11})^{n_{10}}} \tag{11}$$

As before, n_1 is the observed number of exceedances and $n_0 = n - n_1$ is the number of 'good' returns. n_{00} is the number of times a good return is followed by another good return, n_{01} the number of times a good return is followed by an exceedance, n_{10} the number of times an exceedance is followed by a good return, and n_{11} the number of times an exceedance is followed by another exceedance. So, $n_1 = n_{11} + n_{01}$ and $n_0 = n_{10} + n_{00}$. Also,

$$\pi_{01} = \frac{n_{01}}{n_{00}+n_{01}} \quad \text{and} \quad \pi_{11} = \frac{n_{11}}{n_{10}+n_{11}} \tag{12}$$

The asymptotic distribution of $-2ln(LR_{cc})$ is chi-squared with two degrees of freedom (Alexander, 2008).

The conditional coverage test checks for clustering, but it only uses consecutive data points. In other words, it only tests the clustering of one lag. Engle and Manganelli (2004) propose another test, the dynamic quantile or DQ test:

$$Hit(y_t,\theta) \equiv Hit_{\theta t} \equiv I(y_t < -VaR_t) - \theta \tag{13}$$

[7] The Oxford Dictionary of Statistical Terms (2003).

They define a Hit function that takes the value $(1-\theta)$ every time y_t is less than VaR_t and $(-\theta)$ otherwise. For a good model the expected value of Hit is zero and Hit_t will be uncorrelated with any lag Hit_{t-k}, with the forecasted VaR_t and with any constant. If these conditions are satisfied, there will be no autocorrelation in the hits, no measurement error and there will be the correct fraction of exceedances. To test the independence of Hit_t we regress Hit_t on a constant and the lagged Hit_{t-k} up to $k = 4$:

$$Hit_t = \delta_0 + \delta_1 Hit_{t-1} + \delta_2 Hit_{t-2} + \delta_3 Hit_{t-3} + \delta_4 Hit_{t-4} + u_t \tag{14}$$

$$or \qquad Hit_t = X\delta + u_t \tag{15}$$

A good model should produce a sequence of unbiased and uncorrelated hits, so the regression coefficients should be zero. Hence, the following is true for u_t:

$$u_t = \begin{cases} -\theta & prob(1-\theta) \\ (1-\theta)/2 & prob(\theta) \end{cases} \tag{16}$$

To test the performance of the model we want to test the null hypothesis $H_0 : \delta = 0$. The asymptotic distribution of the OLS estimator under the null is defined as

$$\hat{\delta}_{OLS} = (X'X)^{-1}X'Hit \quad \sim \quad N\Big(0, \theta(1-\theta)(X'X)^{-1}\Big) \tag{17}$$

From this Engle and Manganelli (2004) derive the DQ test statistic:

$$DQ = \frac{\hat{\delta}'_{OLS} X'X \hat{\delta}_{OLS}}{\theta(1-\theta)} \tag{18}$$

The asymptotic distribution of DQ is chi-squared with six degrees of freedom.

5 Results

In order to find appropriate explanatory variables to use in the out-of-sample VaR forecasts, we perform an in-depth analysis of the effect of fundamental factors on the Nord Pool system price. Based on the variable selection we extend the quantile regressions and CAViaR specifications to include one variable with high explanatory power in the tails. Finally, we perform out-of-sample VaR forecasts and test the ability of the different models in making accurate assessments of the price risk.

5.1 Variable selection

We find that both the daily average price level and the intraday prices show strong correlation with the variables defined in the model. The effects vary over the price distribution, so a regular OLS regression will not be able to capture the full effects from the fundamental factors on the Nord Pool system price. We present the results from the fundamental analysis in figure 4 to 14 in this section. The solid black line in the figures is the value of the coefficient in the quantile regression, while the shaded grey area is the 100 reps bootstrapped 95% confidence interval for the coefficient. The dotted straight lines are the OLS coefficient and its 95% confidence interval.

5.1.1 Lagged prices

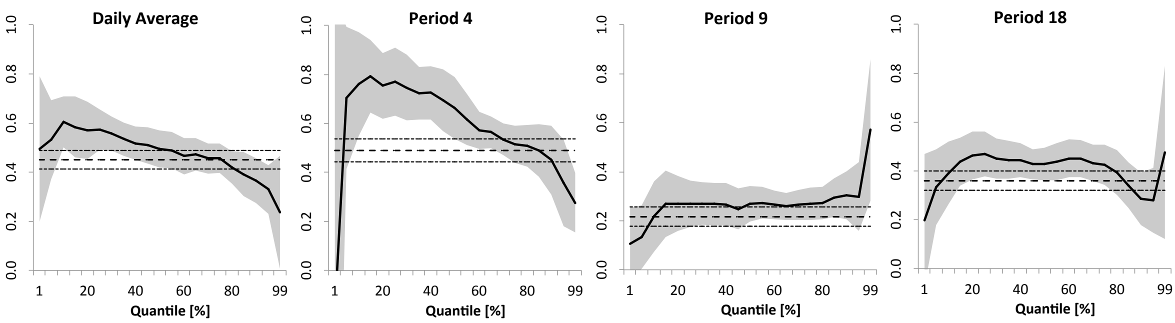

Figure 4: One-day (24 hour) lagged price variable - sensitivities over quantiles.

For the one-day lagged (figure 4) and seven-day average prices (figure 5) we find significant effects for most quantiles. The one-day lagged price decreases its explanatory power over the quantiles for the daily average price and for period 4. Period 9 and 18 show stable effects with some changes on the upper tail. The coefficients are positive for all periods with increasing uncertainty in the tails. The seven-day average price shows signs of decreased explanatory power over the quantiles with some positive effect in the upper tail of the daily average price and for period 18. All periods have increased effect in the lower tail and all values are positive. In general the seven-day lagged price shows large confidence intervals, indicating uncertainty in the estimates.

The results indicate that period 4 is closer linked to the last day's price and that period 9 and 18 is more linked to the average price of the last week. Period 9, being the most volatile data set with large deviation in prices, shows weaker correlation with lagged prices, especially in the upper tail of the price. This is logical as price shocks often are results of changes in other fundamental factors such as temperature and demand.

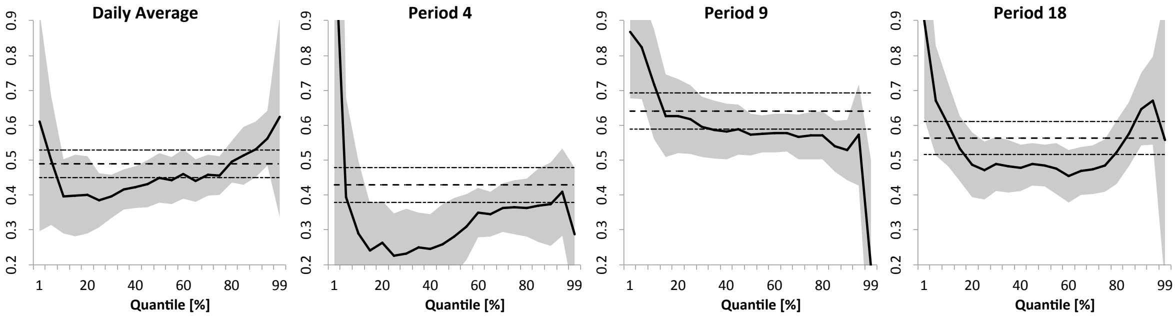

Figure 5: Seven-day (168 hours) average price variable - sensitivities over quantiles.

Lucia and Schwartz (2002) conclude that the Nord Pool system price tend to be mean reverting. This means that prices will continue to return to the long-term average, despite fluctuations above or below the average price. The further away from average the price gets, the higher is the probability for the price to move back towards the average. Our findings do not reject the mean reverting hypothesis as the lagged price coefficients are smaller than one in absolute value.

5.1.2 Price volatility

Figure 6 shows that the price volatility coefficient moves from negative to positive over the quantiles. The effects are significant with small confidence intervals. As expected, this indicates that increased volatility increases the absolute value of the VaR in both the upper and lower quantile. The effect has the highest magnitude in the lower tail of period 4 and in the upper tail of period 9. Period 18 shows close to symmetrical results, with signs of higher magnitude in the upper quantile.

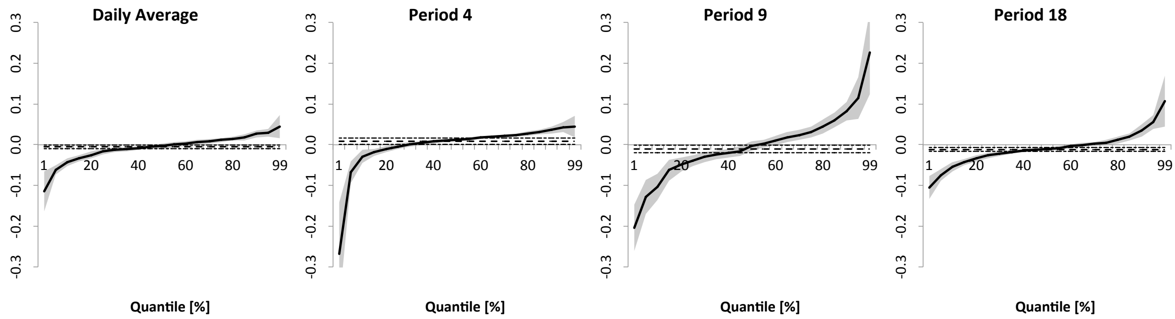

Figure 6: Price volatility variable from the skew-t GARCH-GJR - sensitivities over quantiles.

These asymmetries indicate a negative skew in period 4 and positive skew in period 9 and 18. In other words, the absolute value of the VaR is more sensitive to changes in volatility on the downside of period 4 and on the upside of period 9 and 18.

5.1.3 Gas and coal prices

Both gas prices (figure 7) and coal prices (figure 8) show increased coefficient value over the price quantiles. Period 4 has a negative effect in the lower tail and a positive effect in the upper tail, while period 9 only has the positive effect in the upper tail. Period 18 is stable over the quantiles with a value close to zero. The effect from gas and coal on the daily average price is smaller than for the intraday periods, but shows a negative effect in the lower tail and a positive effect in the upper tail. The coal price differs from the gas price in the upper tail of period 9 and 18 with decreased effect in this region. Both series show increased uncertainty in the outer quantiles with few significant values.

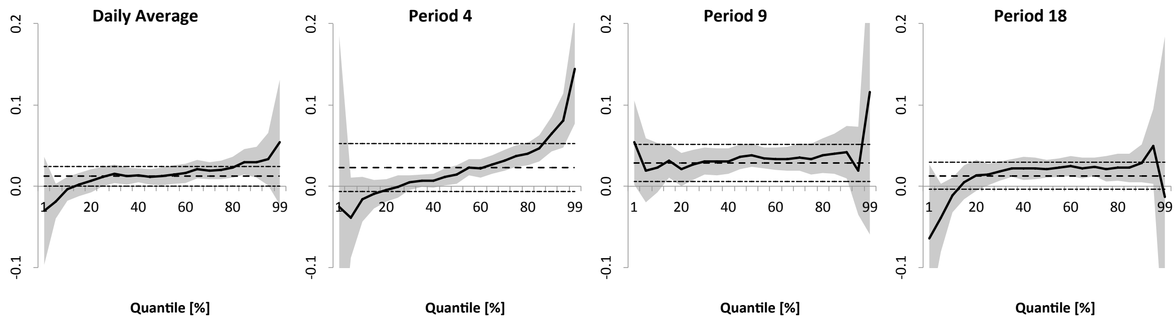

Figure 7: Gas price variable - sensitivities over quantiles.

The effect of gas price on the system price is small and close to zero for low prices and spikes when prices get high. This is because little or no energy is produced from gas during base load conditions, but kicks in at some level of demand. We observe that the price is more sensitive to gas price in period 9 than period 18. This is probably because most gas plants that are ramped up to cover the demand in period 9 are kept running until period 18.

Denmark and Finland have some coal production mainly used to cover the base load. However, expensive coal power is also imported from Europe during peak periods. This explains why the coal price shows some of the same effects as the gas price. The smaller effects in the upper tail of period 9 and 18, compared to gas, is a result of less flexibility in up- and downward regulation of coal fired power plants.

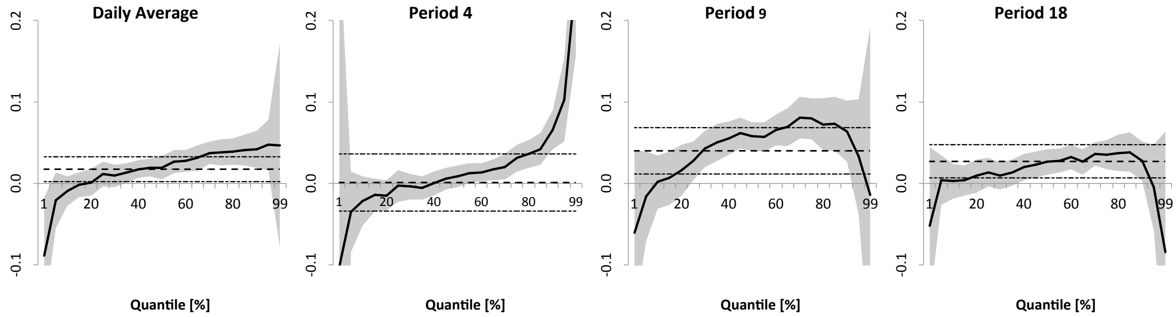

Figure 8: Coal price variable - sensitivities over quantiles.

5.1.4 Demand forecast

Figure 9 shows that the demand forecast has significant estimates and high coefficient values, with some uncertainty in the lower quantiles. The off-peak period 4 shows a decreasing trend over the quantiles. The coefficient has a negative slope and flats out in the upper tail. For period 9 and 18 we observe the opposite effect, with a positive and steep increasing slope in the upper quantile. Hence, the demand forecast is an important determinant of the price in the upper tail during peak periods and in the lower tail during off-peak periods of the day. For the daily average price we observe some positive effects in both tails, but of smaller magnitude than in the intraday periods.

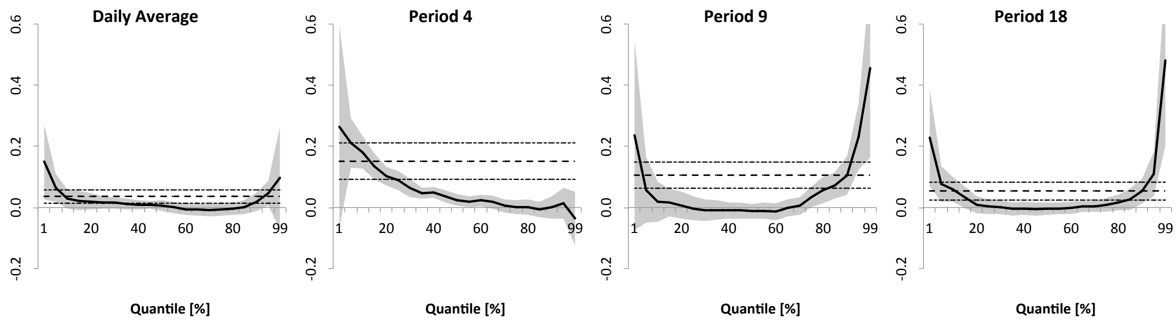

Figure 9: Demand forecast variable - sensitivities over quantiles.

In the low price, low demand period 4, we observe a significant reduction in downside risk associated with an increase in the demand forecast, and vice versa. This stems from the fact that a drop in demand in this period can bring the production down to base load were the marginal cost is close to zero, and therefore decrease the price significantly. For period 9 and 18, where demand is already high, a spike in demand will bring the demand closer to its capacity and increase the probability of shortage in the system. This is the nature of the inelastic supply curve in electricity markets. The effect is smaller for the daily average price as the average demand forecast will be further off base load and the capacity than the forecasts within the day.

5.1.5 Wind production forecasts

Figure 10 shows that wind production forecast has a negative effect on the price and that this effect increases in the tails. The estimates carry some uncertainty in the extremes with wide confidence intervals. For period 4 the effect is largest in the lower tail. Period 9 and 18 has an increased effect in the upper tail, and the daily average effects are similar.

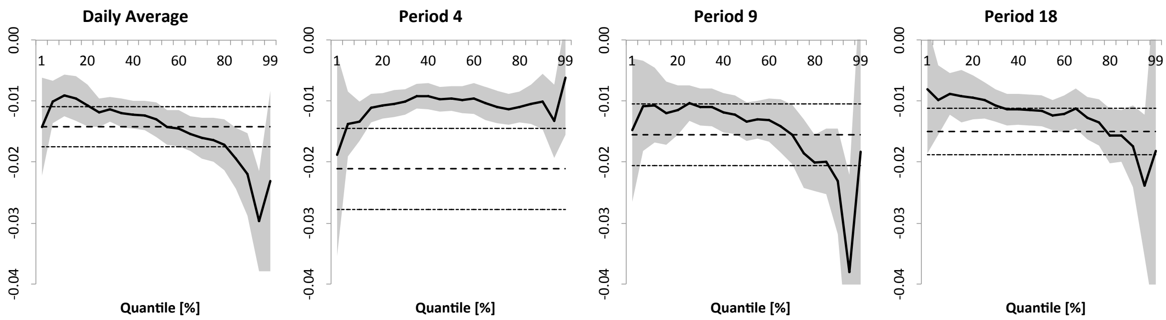

Figure 10: Wind production forecast variable - sensitivities over quantiles.

It is not possible to regulate the production upwards from wind power plants, as the plants produces only when the wind blows. It is therefore neither considered base- or peak-load. Still, forecasts of high wind production will lower the prices, especially during peak prices. Wind power is therefore important in deflecting peaks and reducing extreme spikes in the Nord Pool market.

The magnitude of the wind coefficient is small. Observations in the market indicate that this effect is larger than the figure shows.[8] This could be due to the increased use of wind power over the period analyzed in this paper. In recent years and in the future the effect from wind production can be larger than the effects found here.

5.1.6 Hydrological balance

The hydrological balance shows a negative and decreasing slope over the price quantiles for all periods analyzed in figure 11. The effect is most significant for period 4 and smaller for period 9 and 18. In both tails the estimates are generally not significant. The high percentage of electricity produced from hydropower plants in Nord Pool makes

[8]Observations made by TrønderEnergi.

this factor a determinant of the price. Low hydrological balance will result in high electricity prices, as the risk of future scarcity in energy supply increases. The results indicate that the price is more sensitive to hydrological balance in period 4 than in period 9 and 18, with a clear increase in magnitude in the upper tail of period 4.

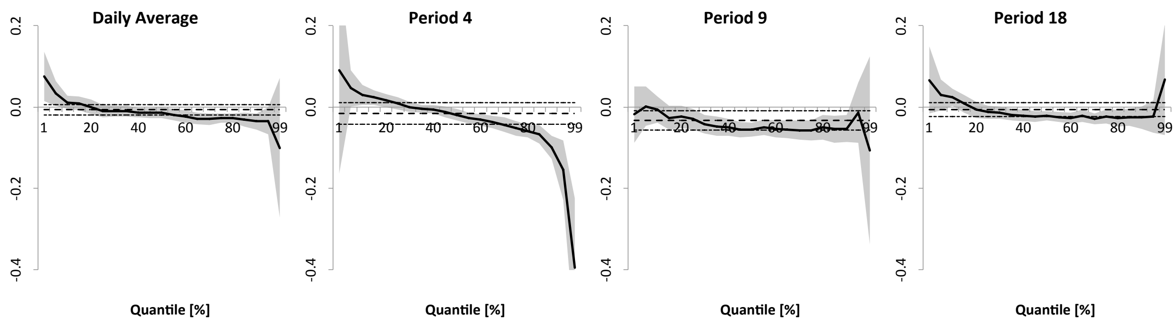

Figure 11: Hydrological balance variable - sensitivities over quantiles.

The regression median is slightly negative for all periods and for the daily average price, indicating that an increase in the hydrological balance will put downward pressure on the price. This is as expected. The form of the curve for the daily average price and period 4, indicate less risk with higher values of the hydrological balance, both on the up- and downside. Reduced upside risk combined with higher hydrological balance is logical. We also observe this in period 9. The reduced downside risk however, we believe originates from the price curve turning more and more elastic with increasing hydrological balance in an already low demand hour of the day. Another explanation is the low water values in this range of prices, providing producers with incentives to postpone their production. This will have a positive effect on the price.

5.1.7 Deviation in temperature and inflow

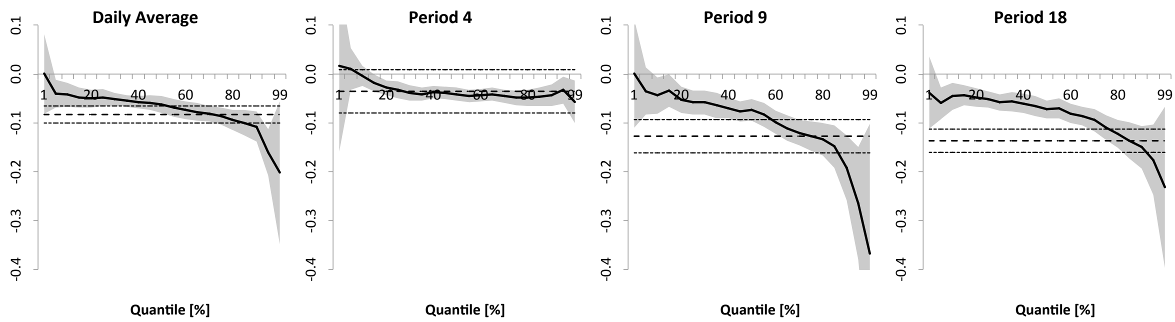

Figure 12: Temperature deviation variable - sensitivities over quantiles.

The deviation between forecasted and normal temperature (figure 12) shows a consistent negative value with a negative spike in the upper tail of period 9 and 18. Most values are significant with some uncertainties in the 1% and 99% quantiles. When forecasted temperature is higher than normal this will lower the price and vice versa. This is because a large part of the electricity in the Nordic market is used for heating, and a higher temperature indicates less demand. In period 4 we observe a similar effect as we did for the hydrological balance: an increase in the temperature will put downward pressure on the price and make the price curve more elastic, i.e. reducing upside risk.

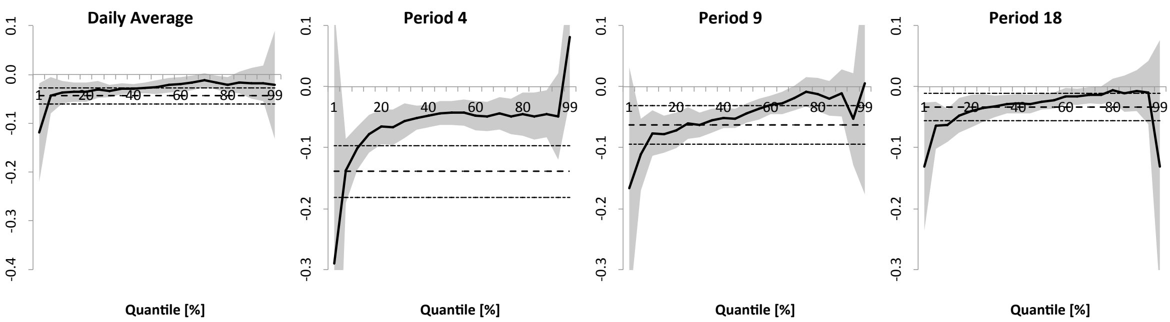

Figure 13: Inflow deviation variable - sensitivities over quantiles.

The negative effect is also present for the deviation between forecasted and normal inflow (figure 13); however, with large confidence intervals in the tails. When inflow is lower than normal this will increase the price and vice versa.

5.1.8 Weekend effects

Figure 14 presents a clear negative effect on the median price during weekends. The effect is larger in the lower tail, indicating more downside risk entering weekends. Lower industry activity during weekends is one explanation of this effect.

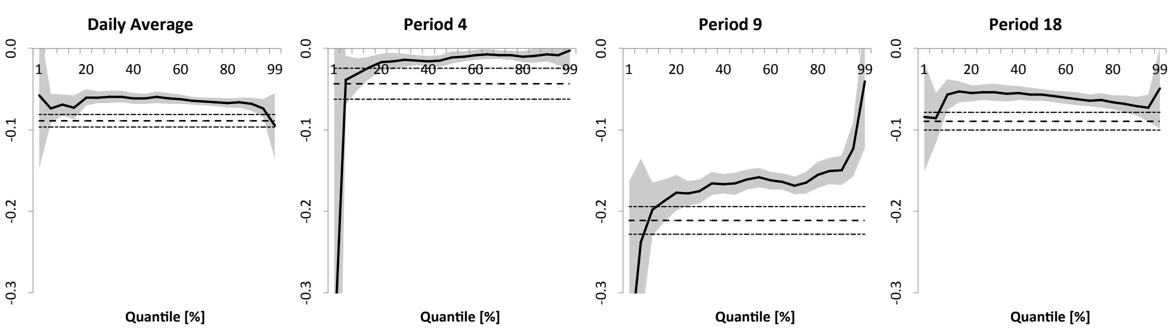

Figure 14: Weekend effects variable - sensitivities over quantiles.

5.1.9 Main findings from the variable selection

The above figures and description reveal a mean reverting system price with strong intraday correlation to fundamental market factors. For the daily average price the correlation to fundamental factors is weaker. Demand forecast is the most significant factor with the highest explanatory power in the tails of the intraday periods. The price volatility shows significant effects for both the daily average price and the three periods, but of smaller magnitude than the demand forecast. Other factors such as lagged system prices, gas and coal price, hydrological balance, wind production forecasts, and temperature and inflow deviation, all show some significant effects of smaller magnitude than the demand forecast and price volatility.

Moving to the out-of-sample analysis, our findings from the variable selection suggest that the demand forecast is the best factor to use in models including explanatory variables. Price volatility is another factor that can provide good out-of-sample forecasts.

5.2 Out-of-sample Value at Risk analysis

To perform out-of-sample analysis on the models presented in section 4, we use a sample of 2214 daily prices from Nord Pool for hours 4, 9 and 18, as well as the average daily price. We compute the log returns, and use these in all our models. We use a fixed starting point increasing window with initial size of 1000 in-sample days, and estimate 1%, 5%, 95% and 99% VaR 1-day ahead. This leaves us with 1213 forecasts for each time series, model and quantile, with their corresponding realized returns. An event is flagged if the realized return exceeds the given VaR forecast, and the events are used in our coverage, conditional coverage and DQ tests. All models are run in Matlab, and parameters for all models are re-estimated in every step.[9]

The GARCH model parameters are found using maximum likelihood, and the VaR forecast is produced from the GARCH parameters and the last day's return.

Based on our findings in the variable selection we test two linear quantile regressions, one with demand forecast as explanatory variable (Qreg DF) and one with price volatility as explanatory variable (Qreg Vol). Both models include a constant (intercept). For the Qreg DF model, the parameters are found using the in-sample window. The forecast is produced using the demand forecast for the next day and the parameters. Qreg Vol first estimates the GARCH model parameters using maximum likelihood, and then

[9]Except for λ in the EWQR model.

uses the GARCH volatility as explanatory variable in linear quantile regression. We forecast volatility from the GARCH model parameters and the last day's return, and use the forecasted volatility with the beta from the quantile regression to forecast the next day's VaR.

The EWQR model is different from the other models in that it uses a moving window of 250 in-sample days, and the only independent variable is a constant. Optimal λ is estimated using the sample spanning from day 250-1000. RQ criterion for λ from 0.7 to 1 with step size .002 is found, and we pick the one yielding the lowest RQ. In some cases the λ is radically lower than the others and is set to the median of the rest. We do not re-estimate λ weights (these are included in section 7.5 in the appendix). The day-ahead VaR forecast from the EWQR is the result of the regression on the last 250 day's log returns.

The CAViaR models' parameters are found using the in-sample window. Day-ahead VaR is forecasted using the estimated parameters, the last day's estimated VaR, the last day's return, and in the case of the I-GARCHX CAViaR, the demand forecast. We set the starting VaR to the empirical quantile of the first 300 observations, and estimate the parameters using the following optimization routine: i) n random uniform(0,1) distributed vectors are generated. We calculate in-sample RQ criterion using each of the vectors as model parameters, and select the m vectors with lowest RQ as initial values for the optimization routine.[10] ii) We run the simplex algorithm (once for every m initial conditions), feed the resulting parameters to the quasi-Newton algorithm, and return the results back into the simplex algorithm. We repeat this procedure until it converges.[11] The vector yielding the lowest RQ criterion is finally selected. As this is very time-consuming, i) is only run for every 100 samples. For all samples in between, the optimal parameters found in the last step is used as initial conditions. As a result, following sudden shocks in the data, the CAViaR model may temporarily settle on suboptimal parameters. However, we find this to be a good time/result tradeoff.

5.2.1 Empirical results

We present the detailed performance of the out-of-sample forecast for the different models in table 6 and 7 at the end of this section. Table 4 and 5 summarize the number of test rejections, while figure 15 present the sum of test rejections at 5% and

[10] We follow Manganelli and Engle (2004) and set $n = [10^4, 10^5, 10^5, 10^6]$ and $m = [10, 15, 15, 20]$ for the I-GARCH, I-GJR-GARCH, I-GARCHX and I-GJR-GARCHX models respectively.

[11] Tolerance levels for function and parameter values are set to 10^{-10}.

1% significance level respectively. The model that performs best is the I-GARCHX CAViaR model with demand forecast as explanatory variable.

All the models outperform the GARCH-N model that was included for benchmarking purposes. This model is punished for assuming normally distributed returns, thereby missing the kurtosis and skewness in the data. It also assumes that the volatility responds similarly to an increase/decrease in price. The model fails the DQ test in 13 out of 16 cases at 5% level and has similarly poor results in the other less strict tests. The skew-t GARCH-GJR model corrects for the mentioned shortcomings, and we observe a clear improvement in performance. It is rejected 6 out of 16 times by the DQ test on the 5% level, and only 4 out of 16 times by the conditional coverage test. We also observe that the model passes all the tests for the daily average price. This makes the skew-t GARCH-GJR model a sufficient performer.

	DQ	CC	UC	P4	P9	P18	DA	Total
(Number of tests)	(16)	(16)	(16)	(12)	(12)	(12)	(12)	(48)
GARCH-N	13	12	13	7	12	7	12	38
Skew-t GARCH-GJR	6	4	2	8	1	3	0	12
Qreg DF	13	4	1	6	1	5	6	18
Qreg Vol	4	1	1	5	1	0	0	6
EWQR	14	5	1	7	3	4	6	20
I-GARCH CAViaR	2	1	1	1	0	3	0	4
I-GJR-GARCH CAViaR	2	2	1	4	0	1	0	5
I-GARCHX CAViaR	2	0	0	0	1	1	0	2
I-GJR-GARCHX CAViaR	4	3	3	5	1	3	1	10

DQ is the number of rejections of the DQ test, *CC* the conditional coverage test and *UC* the unconditional coverage test, aggregated for all periods. *P4*, *P9*, *P18* and *DA* are the sum of rejections for all three backtesting models in period 4, 9, 18 and for the daily average price. *Total* is the sum of rejections for the three backtesting models over all periods. The numbers in parenthesis is the number of tests performed on the model.

Table 4: Summary of the VaR results in table 6 and 7. Number of test rejections at 5% significance level. *Note: Smaller values are better.*

Qreg DF fails the DQ test 13 out of 16 times and the CC test 4 out of 16 times at the 5% level. However, it fails only one time in period 9. This indicates that the demand forecast is a better determinant of the risk during peak periods than during off-peak periods of the day. Qreg Vol performs best of our linear quantile regressions, failing the DQ test only 4 out of 16 times, and the conditional coverage test one time at the 5% level. This model passes all the tests both for period 18 and for the daily average price and is rejected only one time for period 9 at 5% level. It passes all periods except period 4 at 1% level. Hence, Qreg Vol is one of the top performers.

The EWQR model performs quite poorly, failing 14 out of 16 DQ tests, and 5 out of 16 CC tests at the 5% level. At the 1% level it fails the DQ test 9 out of 16 times. One obvious drawback with the model is the fact that its VaR forecast is bound by the past 250 days of returns. Consecutive record returns will increase the autocorrelation in the violation of VaR forecasts, making the DQ test and CC test fail.

	DQ	CC	UC	P4	P9	P18	DA	Total
(Number of tests)	(16)	(16)	(16)	(12)	(12)	(12)	(12)	(48)
GARCH-N	11	10	10	3	12	7	9	31
Skew-t GARCH-GJR	3	2	0	3	0	2	0	5
Qreg DF	9	3	1	5	1	4	3	13
Qreg Vol	2	1	1	4	0	0	0	4
EWQR	9	3	0	5	1	3	3	12
I-GARCH CAViaR	1	1	1	0	0	3	0	3
I-GJR-GARCH CAViaR	1	1	0	2	0	0	0	2
I-GARCHX CAViaR	0	0	0	0	0	0	0	0
I-GJR-GARCHX CAViaR	3	1	0	1	0	3	0	4

DQ is the number of rejections of the DQ test, *CC* the conditional coverage test and *UC* the unconditional coverage test, aggregated for all periods. *P4*, *P9*, *P18* and *DA* are the sum of rejections for all three backtesting models in period 4, 9, 18 and for the daily average price. *Total* is the sum of rejections for the three backtesting models over all periods. The numbers in parenthesis is the number of tests performed on the model.

Table 5: Summary of the VaR results in table 6 and 7. Number of test rejections at 1% significance level. *Note: Smaller values are better.*

The I-GARCH CAViaR model performs well, failing the DQ test only two times and the CC test one time at 5% level. Extending the model to account for asymmetrical response to returns does not improve the model performance. Both CAViaR specifications without explanatory variables pass all tests for period 9 and for the daily average price. These results indicate that all time series have autocorrelation in the VaR estimate, with stronger effects in period 9 and for the daily average price. This makes CAViaR models in general well suited for out-of-sample VaR forecasts.

The innovation in this paper of including explanatory variables in the CAViaR models improves the performance slightly. By using demand forecast as the explanatory variable in the I-GARCHX CAViaR model, we get a model that fails the DQ test 2 out of 16 times and that passes both coverage tests for all periods and all quantiles at the 5% level. At 1% level it passes all tests for all periods and quantiles. This proves that the strong relation between prices and demand forecasts found in section 5.1, combined with the superior performance of the CAViaR models, provides accurate out-of-sample VaR forecasts. Extending the model to account for asymmetrical response to returns does not improve the model performance. This might be due to the increase in parameters leading to overfitting.

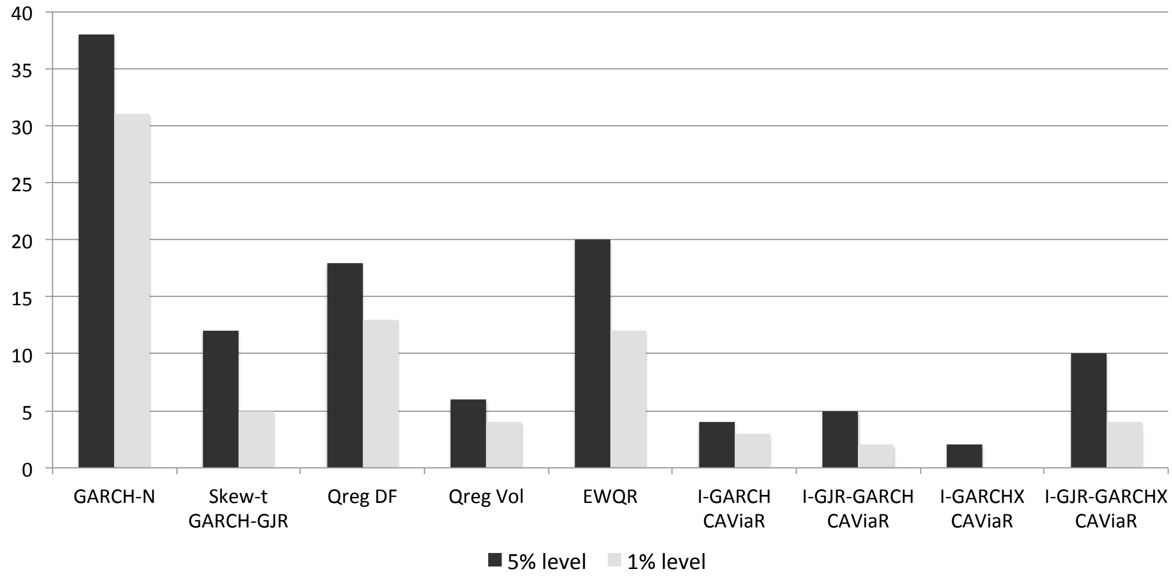

Figure 15: The total number of test rejections for the three backtesting models in table 4 and 5, at 5% and 1% significance level. The potential number of test rejections is 48 for all models and each significance level, the same as the total number of tests performed on each model. *Note: Smaller values are better.*

Data set	θ	GARCH-N			Skew-t GARCH-GJR			Qreg DF			Qreg Vol			EWQR		
		DQ	CC	UC	DQ	CC	UC	DQ	CC	UC	DQ	CC	UC	DQ	CC	UC
1-day horizon																
Period 4	1.0	0.00	0.00	0.00	8.25	4.70	2.04	2.20	13.83	6.71	41.72	45.81	28.70	0.00	21.87	11.40
	5.0	2.74	7.96	2.83	3.24	87.95	85.94	0.00	79.67	53.51	0.00	44.34	24.32	0.00	6.43	66.17
	95.0	12.02	1.03	1.44	0.01	0.27	1.06	0.00	18.97	72.52	2.47	0.95	0.23	0.00	0.61	66.17
	99.0	32.74	88.64	97.02	0.02	2.43	11.40	0.00	0.41	0.53	0.00	88.64	97.02	0.00	2.72	1.06
Period 9	1.0	0.00	0.00	0.00	86.61	45.81	28.70	95.55	59.74	34.39	99.43	75.27	52.62	0.37	30.11	80.40
	5.0	0.00	0.01	0.03	28.27	51.96	93.14	15.15	56.09	96.34	80.26	56.52	57.09	18.31	62.29	37.24
	95.0	0.00	0.00	0.00	2.13	8.09	41.03	82.83	57.46	30.35	28.11	49.95	24.32	4.62	82.50	72.52
	99.0	0.00	0.00	0.00	13.88	20.47	18.51	0.13	30.88	59.84	1.43	21.51	74.04	3.76	30.11	80.40
Period 18	1.0	0.00	0.00	0.00	33.41	30.11	80.40	0.00	25.31	28.70	99.23	84.23	80.40	33.19	73.92	59.84
	5.0	0.56	35.84	34.18	2.70	36.27	57.09	0.00	0.03	18.38	20.67	49.03	28.13	0.00	0.01	82.72
	95.0	18.54	41.20	18.38	0.34	0.38	22.87	1.80	18.14	28.13	34.91	41.20	18.38	1.22	33.71	75.83
	99.0	0.00	0.00	0.00	26.09	45.81	28.70	0.00	13.83	6.71	98.10	73.92	59.84	0.00	45.81	28.70
DailyAve	1.0	0.00	0.00	0.00	98.81	88.64	97.02	0.00	30.11	80.40	98.13	73.92	59.84	0.00	2.83	18.51
	5.0	0.00	0.67	0.24	26.81	28.69	96.34	0.00	0.01	41.03	11.39	10.25	93.14	0.02	0.97	93.16
	95.0	1.35	3.75	1.06	5.03	6.43	66.17	2.14	3.46	18.38	81.33	10.25	93.14	3.73	99.61	93.16
	99.0	0.00	0.00	0.00	12.50	32.64	18.51	1.13	45.81	28.70	28.43	84.23	80.40	1.96	45.81	28.70

Qreg DF and *Qreg Vol* are linear quantile regressions with demand forecast and volatility as its respective explanatory variable. θ is the confidence level at which VaR is calculated. *DQ* is the Dynamic Quantile test p-value of the null hypothesis that there is no autocorrelation, no measurement error and the correct fraction of exceedances in the *Hit* function defined in section 4.4. *CC* is the Conditional Coverage test with the null hypothesis that violations are spread evenly over time (as opposed to being clustered), and that the actual number of violations is equal to the expected number of violations. *UC* is the Unconditional Coverage test p-value with the null hypothesis that the actual number of violations is equal to the expected number of violations.

Note: Higher p-values are better.

Table 6: Backtests of the models used in the out-of-sample VaR analysis.

Data set	θ	I-GARCH CAViaR			I-GJR-GARCH CAViaR			I-GARCHX CAViaR			I-GJR-GARCHX CAViaR		
		DQ	CC	UC	DQ	CC	UC	DQ	CC	UC	DQ	CC	UC
1-day horizon													
Period 4	1.0	40.83	45.81	28.70	21.74	8.27	3.78	27.74	21.87	11.40	15.77	8.27	3.78
	5.0	3.81	28.34	19.17	4.01	32.01	24.32	20.54	38.16	37.24	22.56	4.93	24.32
	95.0	76.22	49.48	48.68	0.90	0.97	93.16	63.35	22.37	57.09	15.55	6.43	66.17
	99.0	96.14	84.23	80.40	34.08	20.47	18.51	85.70	45.81	28.70	0.11	1.44	3.78
Period 9	1.0	92.59	73.92	59.84	61.05	32.64	18.51	82.77	88.64	97.00	57.50	88.64	97.00
	5.0	17.43	81.96	93.16	20.77	86.44	96.33	21.03	49.95	24.32	13.04	10.80	4.42
	95.0	41.37	5.42	37.24	36.57	86.44	96.33	4.62	28.34	19.17	58.83	16.72	8.42
	99.0	99.07	84.23	80.40	41.65	30.88	59.84	12.67	73.92	59.84	10.62	30.11	80.40
Period 18	1.0	40.51	30.88	59.84	41.58	30.88	59.84	96.49	75.27	52.61	97.86	75.27	52.61
	5.0	20.28	81.90	57.09	6.62	4.92	85.93	1.34	38.57	93.16	54.10	87.08	62.73
	95.0	0.02	0.34	0.52	30.03	17.36	22.87	37.38	49.04	28.13	0.00	0.27	41.03
	99.0	99.23	84.23	80.40	15.62	75.27	52.61	19.52	45.81	28.70	0.02	60.24	42.45
DailyAve	1.0	40.89	25.31	28.70	74.78	32.64	18.51	98.84	88.64	97.00	62.23	45.81	28.70
	5.0	17.80	36.05	62.73	6.12	12.78	53.51	20.01	71.26	44.98	24.16	62.29	37.24
	95.0	31.35	5.42	37.24	71.66	28.64	14.84	64.56	41.78	72.52	12.54	33.64	30.35
	99.0	24.72	88.64	97.00	8.88	60.24	42.45	32.30	88.64	97.00	4.86	21.87	11.40

θ is the confidence level at which VaR is calculated. *DQ* is the Dynamic Quantile test p-value of the null hypothesis that there is no autocorrelation, no measurement error and the correct fraction of exceedances in the *Hit* function defined in section 4.4. *CC* is the Conditional Coverage test with the null hypothesis that violations are spread evenly over time (as opposed to being clustered), and that the actual number of violations is equal to the expected number of violations. *UC* is the Unconditional Coverage test p-value with the null hypothesis that the actual number of violations is equal to the expected number of violations.

Note: Higher p-values are better.

Table 7: Backtests of the models used in the out-of-sample VaR analysis *(Continued)*.

5.3 Practical implications

Electricity trading includes both the spot and the financial market. The derivatives traded on Nord Pool's Financial Market comprise base load and peak load futures, forwards and options. These are all based on the system price. It is therefore important for market participants to obtain knowledge on the system price formations to make good trading decisions. Insight in how the tail is formed is a part of this knowledge. Hence, an accurate forecast of price risk is important for participants in the financial market.

VaR is a risk measure that is useful for the clearinghouse in determining its margin levels. Nord Pool sets margin requirements and adjusts these after changes in the market. With accurate VaR forecasts these adjustments can be done at an earlier stage and thereby increase the predictability of the margin deposits required by market participants with positions in the financial market.

Risk forecasting, both in the short- and the long-term, could also be useful for producers and consumers to determine their respective bidding strategies. The amount of production or consumption reported by participants in the market depends on the price. A producer will place bids of high production when prices are high and low production when prices are low. Producers and consumers can take advantage of sudden changes in price by measuring the risk. An electricity producer that gets forecasts of high upside risk can offer a high volume to take advantage of the potential spike in price. For a forecast of high downside risk, a producer would restrict its production and offer a lower volume, as the risk of a price lower than the cost of production increases. The opposite will apply for an electricity consumer in the market.

In addition to accurate VaR forecasts, our linear quantile regression model makes it possible to gain insight on how fundamental factors affect the electricity price in various price ranges and for different periods of the day. Measuring these factors can provide useful information about the price risk.

Simple models reduce the computational complexity and increase the probability for models to be adopted by market participants. Linear quantile regression is easy to use and does not require time consuming analysis and deep knowledge in statistics. CAViaR models are more sophisticated and complex. This might cause linear quantile regression to be favored among market participants, in spite of CAViaR models' superior performance.

6 Conclusions

In this paper we discuss tail behavior of the Nord Pool system price and, in particular, the forecasts of the market risk measure Value at Risk (VaR). We compare linear- and nonlinear quantile regression models to different GARCH specifications.

First we run a linear quantile regression to characterize the effects of fundamental factors on the system price formations. We find that demand forecast and price volatility are the most dominant factors influencing the spot price. In general, the parameters show high significance in the tails with decreasing effect in the interior of the distribution. Based on these findings we develop a CAViaR specification taking in explanatory variables.

Secondly, we run out-of-sample testing of VaR forecasts. We find that the nonlinear quantile regression models, CAViaR, outperform the linear quantile regression and GARCH models. Of the CAViaR specifications, our I-GARCHX CAViaR innovation with demand forecast as explanatory variable is the best performer. A quantile regression with price volatility as explanatory variable also performs well, and outperforms the skew-t GARCH-GJR model.

The results indicate that quantile regression has a significant potential in giving accurate forecasts of VaR. A linear quantile regression model is easy to use and do not require deep knowledge in statistics; increasing the probability of market participants to adopt the method. The most accurate method, the I-GARCHX CAViaR, however, is a more sophisticated model, which may inhibit its use in spite of its superior performance.

Our findings have important implications for participants in both the auction and financial electricity market, as they provide insight on how fundamental factors affect the price and provide accurate forecast of the price risk. Financial market participants can adjust their positions in the derivatives market based on information from the models. Clearinghouses can provide more accurate margin requirements. In other words, our results can help market participants to tune their position in the market and reduce their exposure to risk.

Further research can develop the quantile regressions to get more sophisticated models. CAViaR models have shown the strength of nonlinear models in VaR forecasting. One can extend the CAViaR models further to get more accurate forecasts of the VaR. This will potentially be of greater value for the participants in the market, but the complexity of the model will increase. One can also apply the same analysis on Nord Pool area prices or enter new markets such as the European Energy Exchange (EEX) and the UK power market, to gain insight of market fundamentals in different markets.

Acknowledgement

We want to thank our supervisor, Sjur Westgaard, for feedback and supervision during the process. We would also like to thank TrønderEnergi and Siri Line Hove Ås for industry knowledge and detailed data sets on the Nord Pool market.

References

Alexander, C. (2008). *Market Risk Analysis: Value-at-Risk Models*. Wiley.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, (31).

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, (2).

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, (39).

Contreras, J., Espinola, R., Nogales, F. J., and Coneja, A. J. (2003). Arima models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, (3).

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, (50).

Engle, R. F. and Manganelli, S. (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, (4).

Escribano, A., Pena, J. I., and Villaplana, P. (2002). Modeling electricity prices: International evidence. *Economics Series*, (08).

Garcia, R. C., Contreras, J., van Akkeren, M., and Garcia, J. B. C. (2005). A garch forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems*, (2).

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, (48).

Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review*, (35).

Koenker, R. and Basset, G. (1978). Regression quantiles. *Econometrica*, (1).

Koenker, R. and D'Orey, V. (1987). Computing regression quantiles. *Applied Statistics*, (36).

Koenker, R., Ng, P., and Portnoy, S. (1997). Quantile smoothing splines. *Biometrica*, (81).

Koenker, R. and Park, B. J. (1996). An interior point algorithm for nonlinear quantile regression. *Journal of Econometrics*, (71).

Kuester, K., Mittnik, S., and Paolella, M. S. (2006). Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, (1).

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, (3).

Liu, K., Subbarayan, S., Shoults, R. R., Manry, M. T., Kwan, C., Lewis, F. L., and Naccarino, J. (1996). Comparison of very short-term load forecasting techniques. *IEEE Transactions on Power Systems*.

Lucia, J. J. and Schwartz, E. S. (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Reasearch*, (5).

Malki, H. A., Karayiannis, N. B., and Balasubramanian, M. (2004). Short-term electric power load forecasting using feedforward neural networks. *Exoert Systems*.

Manganelli, S. and Engle, R. F. (2004). A comparison of value at risk models in finance. *Risk Measures for the 21st Century*, (ed. G Szego).

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press.

Taylor, J. W. (2008). Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics*, (6).

Wei, W. W. S. (2005). *Time Series Analysis: Univariate and Multivariate Methods (2nd edition)*. Pearson.

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

7 Appendix

7.1 Variable description

Nord Pool system price

The spot price is the hourly Elspot day-ahead price from the Nord Pool Spot exchange. We analyze the natural logarithm of the daily average price/return and the intraday prices/returns for period 4, 9 and 18 with data covering the period January 1st 2005 to January 23rd 2011. Period 4 is the hour from 03:00-03:59, period 9 is the hour from 08:00-08:59 and period 18 is the hour from 17:00-17:59. The daily average price is the arithmetic average of the 24 intraday day-ahead prices.

Lagged prices

Lagged values of the spot price are included; price in the same trading period the previous day, and the average price of the same trading period of the last weeks (seven days) prices. This provides a link between the price today, yesterday and the price level the last week.

Price volatility

The price volatility is calculated separately for each of the periods and for the daily average price, using a skew-t GARCH-GJR(1,1) model on the price returns. Descriptive statistics for the price volatility is found in table 8 below. The results indicate high standard deviation, positive skew and high kurtosis for all series, especially for period 4 and 9. Period 4 and 9 also have high maximum values. The high kurtosis indicate fat tails and the positive skew indicate positive spikes in volatility.

	Descriptive Statistics							
	No. Obs.	Mean	Median	Minimum	Maximum	St. Dev.	Skewness	Kurtosis
Period 4	2213	0.06	0.00	0.00	9.65	0.36	14.84	295.33
Period 9	2213	0.06	0.02	0.02	8.31	0.25	21.68	627.66
Period 18	2213	0.02	0.01	0.00	0.40	0.04	4.68	25.35
Daily Ave	2213	0.01	0.00	0.00	0.33	0.02	6.97	62.36

Table 8: Descriptive statistics for the price volatility.

Gas price

The gas price used is the daily UK natural gas one-day forward price, from the main National Balancing Point (NBP) hub. The prices are converted from GBP to EUR using the exchange rate for the actual day. Weekend prices are calculated with linear interpolation.

Coal price

The coal price used is the daily API2 coal index for ARA (Amsterdam, Rotterdam, and Antwerp), taking into account the USD/EUR exchange rate. Weekend prices are calculated with linear interpolation.

Demand forecast

The demand forecasts is calculated using feedforward neural networks. The network is a two-layer feed-forward network with 20 sigmoid hidden neurons and linear output neurons trained with the Levenberg-Marquardt backpropagation algorithm. The following variables is used in the model; temperature forecast, one-day lagged demand, last days average demand, seven-day lagged demand, day of week [1,7], and weekend/holiday dummy. The parameters are found using 1000 in-sample data, and we do not re-estimate the model.

	Descriptive Statistics							
	No Obs	Mean	Median	Minimum	Maximum	St. Dev.	Skewness	Kurtosis
Period 4	2213	38.10	37.58	24.54	58.35	7.14	0.29	-3.99
Period 9	2213	47.39	46.53	26.42	73.61	9.25	0.14	-3.76
Period 18	2213	47.38	45.37	29.59	72.54	9.25	0.26	-4.14
Daily Ave	2213	44.50	43.43	27.96	67.77	8.08	0.23	-3.97

Table 9: Descriptive statistics for the demand forecast.

Descriptive statistics for the demand forecast is found in table 9. We observe that the demand forecasts vary across the periods and that period 9 and 18 have larger standard deviation with extreme maximum values. All series are positive skewed with negative kurtosis. This indicate that spikes occur rarely compared to the normal distribution, and that spikes happens more on the upside than on the downside.

Wind production forecast

To measure the amount of wind energy in the market, we use the forecasted production of wind energy for Denmark, provided by SKM Market Predictor[12].

Hydrological balance

Hydrological balance is "actual" deviation from normal magazine- and snow-reservoir levels calculated by SKM Market Predictor. It includes factors such as weather forecasts, reservoir level and inflow. This value can be both positive and negative, so we add a constant to get all values positive, as we take the natural logarithm of all variables.

Deviation temperature

The deviation between forecasted temperature and the normal (average over the last years) temperature for the next day. To avoid problems of multicollinearity with hydrological balance we use deviation from normal instead of actual forecasts. This value can be both positive and negative, so we add a constant to get all values positive, as we take the natural logarithm of all variables.

Deviation inflow

The deviation between forecasted inflow and the normal (average over the last years) inflow for the next day. To avoid problems of multicollinearity with hydrological balance we use deviation from normal instead of actual forecasts. This value can be both positive and negative, so we add a constant to get all values positive, as we take the natural logarithm of all variables.

Weekend dummy

The weekend dummy captures effects on the prices during weekends different than those during weekdays. It is constructed as a dummy variable taking the value one for saturdays and sundays and zero otherwise.

[12] SKM Market Predictor is an independent provider of power system analysis for the power market.

7.2 Box-Cox power transformation

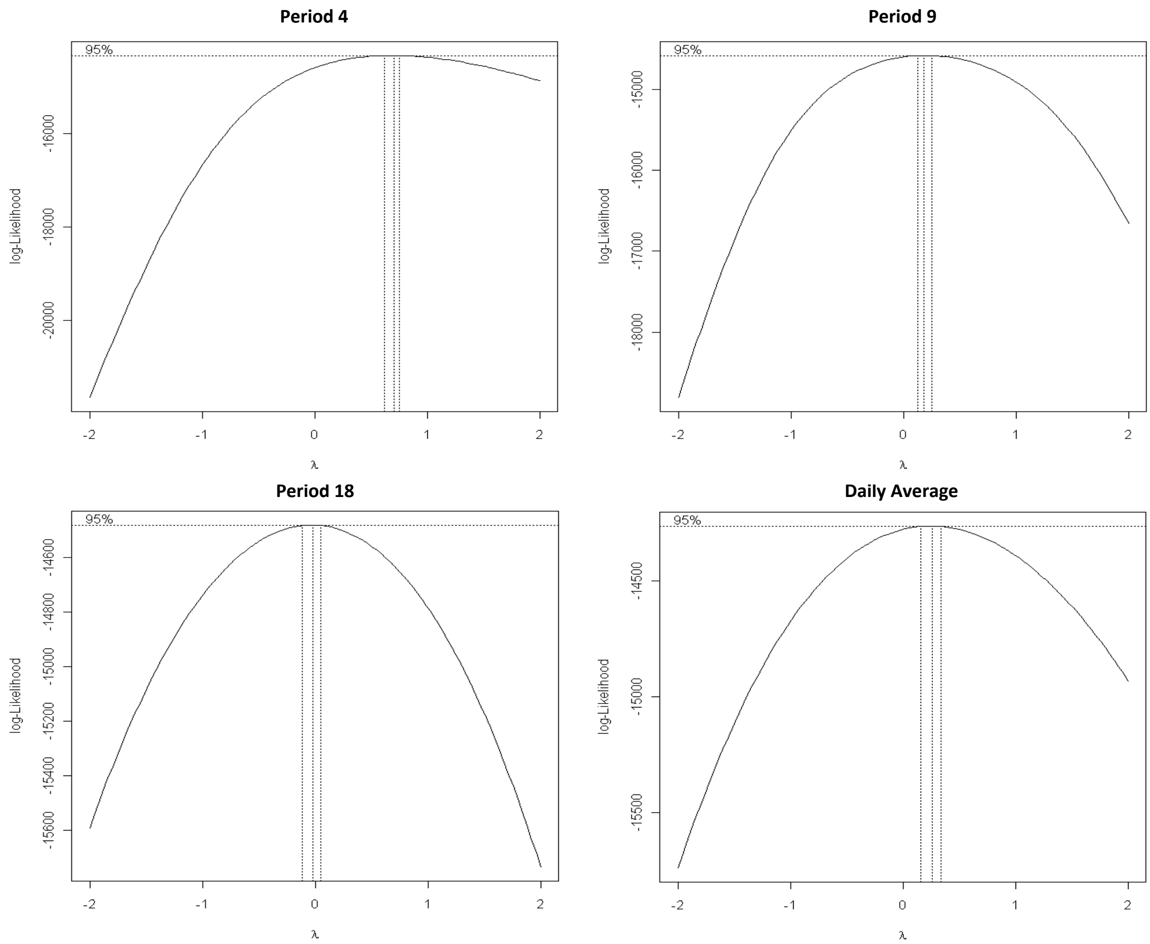

Figure 16: The results from the Box-Cox power transformation test.

Value of λ	Transfromation
-1.0	$\frac{1}{P_t}$
-0.5	$\frac{1}{\sqrt{P_t}}$
0.0	lnP_t
0.5	$\sqrt{P_t}$
1.0	P_t

Table 10: The Box-Cox transformation based on calculated λ.

We want to perform the same transformation on all time series. The average λ is close to zero, so we take the natural logarithm.

7.3 ARMA and GARCH models

An ARMA(p,q) model use autoregressive and moving average terms to approximate the given data, and is defined by Wei (2005) as[13]:

$$(1 - \sum_{i=1}^{p} \phi L^i) y_t = c + (1 + \sum_{i=1}^{q} \Theta_i L^i) \epsilon_t \tag{19}$$

Where L is the backshift operator, so that $L^k y_t = y_{t-k}$ This makes the model suited for autocorrelated time series. Traditional assumptions on the error term are $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \sigma_\epsilon^2$, and $E(\epsilon_t \epsilon_s) = 0$. To account for heteroskedasticity in the variance, ARCH specifications to model autocorrelation in squared residuals have been developed.

The generalized autoregressive conditional heteroskedasticity (GARCH) model is the most popular ARCH (Engle, 1982) specification in empirical research and was introduced by Bollerslev (1986). The model is specified in equation (20).

$$\begin{aligned} \epsilon_t | \psi_{t-1} &\sim N(0, h_t) \\ h_t &= \alpha + \sum_{i=1}^{s} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{r} \beta_i h_{t-i} \end{aligned} \tag{20}$$

Equation (21) presents the conditional volatility from one of the most used ARCH specifications, the GARCH-N(1,1).

$$h_t = (1 - \alpha - \beta)\sigma^2 + \alpha \epsilon_{t-1}^2 + \beta h_{t-1} \tag{21}$$

This model captures the volatility clustering of returns that is often experienced with financial asset returns. To achieve this, the model forecast the volatility to be a weighted combination of the unconditional variance, the previous squared residuals and the previous conditional variance. We include this model in the backtesting, and it will be referred to as GARCH-N. To find the model's VaR, we turn to the more general case. If X has distribution function $F(x)$ and density function $f(x)$ then:

$$VaR(\theta) = F^{-1}(\theta) \tag{22}$$

Under the assumption of normality, with mean μ and conditional variance σ_t this translates to the VaR for the GARCH-N(1,1) model (McNeil et al., 2005):

$$VaR_t(\theta) = \Phi^{-1}(\theta)\sqrt{h_t} + y_t \tag{23}$$

[13] We deviate from the regular notation θ, for the moving average weights, as it is used for quantiles throughout the paper

The GARCH-N(1,1), assumes that the returns are symmetrically distributed, an assumption that for most practical applications have been proved wrong. As skewed distributions with fat tails are common in financial asset returns (the same for the power market returns), a skew-t distribution modeling the returns could be of interest. However, a problem with the GARCH(1,1) model (irrespective of distribution) is that the conditional variance responds symmetrically to returns. Glosten et al. introduced the GJR-GARCH in 1993 and this ARCH specification models the asymmetry in the GARCH process. Equation (24) shows the GJR-GARCH(1,1) model.

$$\begin{gathered} h_t = (1-\alpha-\beta)\sigma^2 + \alpha(r_{t-1}-\mu)^2 + \beta h_{t-1} + \phi(r_{t-1}-\mu)^2 I_{t-1} \\ where \quad I_{t-1} = 0 \quad if \quad (r_{t-1}-\mu) \geq 0 \\ and \quad I_{t-1} = 1 \quad if \quad (r_{t-1}-\mu) < 0 \end{gathered} \tag{24}$$

This model has the same advantages as the GARCH introduced by Bollerslev, but also allow for the conditional volatility to respond asymmetrically to returns. Used with an assumption of skewed-t distributed returns, the model should be quite robust. We include this model with assumed Hansen skew-t distributed returns in the backtesting, and refer to it as the skew-t GJR-GARCH model. We proceed to find VaR and ETL under the model assumptions. From Hansen (1994) in equation (25) we have our density function:

$$g(z|\eta,\lambda) = \begin{cases} bc\left(1 + \frac{1}{\eta-2}(\frac{bz+a}{1-\lambda})^2\right)^{-(\eta+1)/2} & z < -a/b, \\ bc\left(1 + \frac{1}{\eta-2}(\frac{bz+a}{1+\lambda})^2\right)^{-(\eta+1)/2} & z \geq -a/b, \end{cases} \tag{25}$$

with $2 < \eta < \infty$, $-1 < \lambda < 1$, and constants;

$$a = 4\lambda c(\frac{\eta-2}{\eta-1}), \tag{26}$$

$$b^2 = 1 + 3\lambda^2 - a^2, \tag{27}$$

$$c = \frac{\Gamma(\frac{\eta+1}{2})}{\sqrt{\pi(\eta-2)}\Gamma\frac{\eta}{2}} \tag{28}$$

The distribution and specializes become the student's t distribution with $\lambda = 0$. For proof that this is a proper density function with zero mean and unit variance, see appendix in Hansen (1994). With the distributions PDF given by (25), and the CDF $G(x) = \int_{-\infty}^{x} g(z|\eta,\lambda)\,dx$, the VaR follows from equation (22):

$$VaR_t(\theta) = G^{-1}(\theta)\sqrt{h_t} + y_t \tag{29}$$

7.4 CAViaR models with explanatory variables modeling the return

An I-GARCHX CAViaR and I-GJR-GARCHX CAViaR model with explanatory variables modeling the return instead of the quantile as we do in this paper, will take the form of the equations below.

$$VaR_t(\beta) = (\beta_0 + \beta_1(VaR_{t-1} - \beta_3 z_{t-1})^2 + \beta_2(y_{t-1} - \beta_3 z_{t-1})^2)^{1/2} + \beta_3 z_t, \tag{30}$$

$$\begin{aligned} VaR_t(\beta) =& (\beta_0 + \beta_1(VaR_{t-1} - \beta_4 z_{t-1})^2 + \beta_2(y_{t-1} - \beta_4 z_{t-1})^2 \\ &+ \beta_3 I_{(y_{t-1}<0)}(y_{t-1} - \beta_4 z_{t-1})^2)^{1/2} + \beta_4 z_t, \end{aligned} \tag{31}$$

7.5 EWQR lambda estimates

Quantile, θ	Period 4	Period 9	Period 18	DailyAve
0.01	0.998	0.998	0.990	0.998
0.05	0.998	0.998	0.998	0.982
0.95	0.998	0.998	0.998	0.998
0.99	0.998	0.998	0.998	0.998

Table 11: The estimated λ weights used in the EWQR model.