AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

Modeling and Forecasting the Volatility of Energy Forward Returns - Evidence from the Nordic Power Exchange

Asger Lunde and Kasper V. Olesen

CREATES Research Paper 2013-19

Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Modeling and Forecasting the Volatility of Energy Forward Returns*

Evidence from the Nordic Power Exchange

Asger Lunde   Kasper V. Olesen

Aarhus University, Department of Economics and Business & CREATES

May 24, 2013

Abstract

We explore the structure of transaction records from NASDAQ OMX Commodities Europe back to 2006 and analyze base load forwards with the Nordic system price on electric power as reference. Following a discussion of the appropriate rollover scheme we incorporate selected realized measures of volatility in a Realized EGARCH framework for the joint modeling of returns and realized measures of volatility. Conditional variances are shown to vary over time, which stresses the importance of portfolio reallocation for risk management and other purposes. We document gains from utilizing data at higher frequencies by comparing to ordinary EGARCH models that are nested in the Realized EGARCH. We obtain improved fit, in-sample as well as out-of-sample. In-sample in terms of improved log-likelihood and out-of-sample in terms of 1-, 5-, and 20-step-ahead regular and bootstrapped rolling-window forecasts. The Realized EGARCH forecasts are statistically superior to ordinary EGARCH forecasts.

Keywords: Financial Volatility, Realized GARCH, High Frequency Data, Electricity, Power, Forecasting, Realized Variance, Realized Kernel, Model Confidence Set

JEL Classification: C10, C22, C53, C58, C80, G40

*Acknowledgements. Both authors acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation, and thank NASDAQ OMX Commodities Europe for data access. We also thank the AU Ideas Pilot Center, Stochastic and Econometric Analysis of Commodity Markets, for financial support. We are grateful to conference participants at the International Risk Management Conference 2012 in Rome, the CEQURA Conference 2012 in Munich, the Energy Finance Conference 2012 in Trondheim, and seminar participants at CREATES for helpful comments and suggestions. Send correspondence to Asger Lunde, Dept. of Economics and Business, Aarhus University, Fuglesangs Allé 4, 8000 Aarhus C, DK. E-mail address: alunde@creates.au.dk. All errors are ours.

1 Introduction

Over the past decades, the availability of high-frequency financial data has opened a new field of research and paved the way for improved measurement, modeling and forecasting of volatilities and covolatilities (see Barndorff-Nielsen and Shephard (2007), Andersen et al. (2009), and Hansen and Lunde (2011) for recent surveys). Recently, complete records of transaction prices have become available for a range of derivative contracts that have the system price on power in the Nordic countries as underlying reference.[1] These financial products, and similar in other regions, have emerged in the aftermath of the wave of liberalizations that has flooded numerous countries and regions, spread over the world, in the last couple of decades. The Nordic countries were the pioneers back in 1991 with Norway as spearhead, and as of today, the price setting in the Nordic physical spot markets, with actual delivery of power, is handled by Nord Pool Spot (NPS). Spot markets have in general been active fields of study in the academic literature (see Higgs and Worthington (2008) for a recent survey), whereas the financial markets have caught less attention. NPS serves as underlying reference for financial futures, forwards, and options traded on NASDAQ OMX Commodities Europe (NOMXC). The financial contracts were introduced in 1995 motivated by a need for hedging possibilities for market participants exposed to price changes.

In this paper, we show that the inclusion of intraday information and the use of sophisticated econometric methodologies improves the modeling and forecasting of volatilities in an important branch of the energy markets. Specifically, we investigate the liquidity in the NOMXC and conduct an analysis of the most liquid derivative products traded; the base load forwards (referred to as NOMXC forwards in the following).[2] We explore the transaction records and use realized measures of volatility to enrich the information set of the more conventional GARCH models. That is, we utilize the information contained in realized measures of volatility to estimate the conditional variance of daily returns within the Realized GARCH framework of Hansen et al. (2012*a*). In particular, we contribute by utilizing the intraday data at the highest resolution level. Because we have tick-by-tick data we can combine information from multiple realized mesures based on various sampling scales. This has not been investigated in previous studies. Most previous research conducted on this data use a daily, or lower, resolution (e.g. Benth and Koekebakker (2008) or Bauwens et al. (2012)), and Haugom et al. (2011*a*,*b*) has considered the use of 30 minute price information in reduced form time series models for volatility. Our main contribution is to show that frequency matters for the modeling and forecasting of volatility in the financial markets for power, and that novel inventions in the measurement of volatilites can easily be incorporated in applied work using the Realized

[1] We refer to "electric power" as "power" throughout.

[2] Strictly speaking, the contracts are swaps, as they specify an exchange of a fixed cash flow (the forward price) for a floating cash flow (the underlying reference) over a settlement period. We stick to the NOMXC terminology throughout.

GARCH framework.[3]

This is an important message as the power markets are distinct and requires analysis on their own. The importance is demonstrated in for example Geman and Roncoroni (2006), Escribano et al. (2011), and Veraart and Veraart (2013) for spot prices. Constrasting these studies our data set is from the derivatives market. We expect these prices to behave more as traditional financial assets than the spot price that has a very special structure in power markets (see e.g.Veraart and Veraart (2013)). Derivatives on flow commodities, such as power, are often settled against an average of the spot price over a specified delivery period. That is easily contracted for futures and forwards contracts and in the option space classified as Asian options. Consequently, futures and forwards stop trading before the delivery period, which creates the need for rollover between consecutive contracts in order to analyze continual time series. A topic we discuss in Section 3.2. Intraday, the derivative contracts trade as in other markets. This creates potential also for speculation and results in transactions to analyze for the researcher. For the specific contracts analyzed in this paper the trading activity counts more than $700,000$ transactions over our sample period from January 2006 to May 2011.

We find that Realized EGARCH models that utilize the information in realized ex-post measures of volatility massively outperform the benchmark EGARCH in-sample. This is shown through a comparison of the likelihood of the benchmark EGARCH with a partial likelihood contribution in the richer Realized EGARCH. We also explore the additional benefit from including multiple realized measures of volatility and find that this in some cases leads to even better in-sample performance. The superior performance in-sample manifests itself out-of-sample through statistically superior predictions of the conditional variance. Thus, our work shows that by means of the Realized GARCH framework one can in a simple way obtain more precise volatility estimates and short term predictions for the NOMXC forwards.

Our results are of major importance for producers, utility companies, and other participants in the power sector. These agents are exposed to the physical spot market for power, and the financial products traded on NOMXC provide important tools for hedging the risk inherent. The financial products provide protection to rapidly changing prices and to others they are vehicles for portfolio diversification. In active risk and portfolio management, volatility estimates are needed for traditional Markowitz portfolio theory, calculation of hedge ratios, value at risk (VaR) estimates, and so forth. Further, volatility estimates are needed in the pricing of options written on these products, which trade both on NOMXC and over-the-counter.

The paper is organized as follows. Section 2 briefly introduces the history of, and the products traded on, NOMXC and related markets. Section 3 outlines the sorting and filtering of

[3] It is important to note that the HAR models (see e.g. Andersen et al. (2007), and Corsi (2009)) can easily be embedded in the Realized GARCH framework, by allowing for a more flexible dynamic specification of the conditional variance. In contrast, the HAR is a reduced form model that does not allow the researcher to address questions about the return distribution. For this reason, extensions of the model specification to FIGARCH- or HAR-like volatility dynamics are of second order interest, and we defer this for future reseach.

the data, discusses the issue of rollover, and analyzes resulting continual series at the daily frequency. Section 4 presents the econometric methodology used, the Realized GARCH class of models, and Section 5 presents estimation results and diagnostics. In Section 6 we perform regular and bootstrapped rolling-window forecasting and evaluate the models using mean-squared-error (MSE) and predictive likelihood. To differentiate model performance we apply the "model confidence set" (MCS) approach of Hansen et al. (2011). Section 7 concludes.

2 The Financial Market for Nordic Power

Energy markets are commonly classified into three groups; "Fuels", "Power", and "Weather and Emissions (W&E)", which roughly corresponds to the historical pace at which they were opened (see e.g. Eydeland and Wolyniec (2003)). We focus on "Power" for which the wholesale markets were liberalized in numerous countries beginning in the early 1990s. Power is technologically considered non-storable, which creates a need for real-time balancing of locational supply and demand. In most countries, an independent transmission system operator (ITSO) maintains the system and manages provision, contracting, and infrastructure for all needed activities.[4] The ITSOs are non-commercial monopolies that operate the high-voltage grids and secure the supply of power. For example, actual consumption may exceed production, which makes the frequency of the alternating current fall below the target (50 Hz in the Nordic countries), and the ITSO procures "up regulating" (as opposed to "down regulating" in the opposite case). In regulating markets, prices are volatile and most market participants seek to forecast production or demand in order to take positions beforehand. Optimally, only discrepancies between the expectations and actual needs are settled in the regulating market. Positions can be taken bilaterally or through pools. NPS is, as the name suggests, a pool and operates a day-ahead double auction market, Elspot, where market participants submit supply and demand no later than noon the day before the energy is delivered to the grid. A market system clearing price for all hours in the following day is calculated and announced.[5] To further reduce the exposure to the balancing market, a cross-border intraday market, Elbas, opens two hours after the closure of NPS and closes one hour prior to the operation hour. The organization of physical exchanges varies between regions, and we limit ourselves to this brief exposition of the Nordic market.

The system clearing price from the NPS is the underlying reference price for a range of derivative contracts traded on NOMXC. Originally introduced as Eltermin by Nord Pool in late 1995, which among other entities was acquired by NASDAQ OMX in 2008 and merged into NOMXC, the exchange is among the leading and most liquid exchanges in the world for

[4] In the Nordic countries the ITSOs are Energinet.dk (Denmark), Statnett (Norway), Svenska Kraftnät (Sweden), and Fingrid (Finland).

[5] The exposition is simplified. See www.nordpoolspot.com for details. Furthermore, due to grid bottlenecks, bidding areas develop. Thus, different regions in the Nordic countries are exposed to different prices in some hours.

financial derivatives on power. In the financial markets the parties contract for the delivery of power in the future. This can be only days ahead or up to years ahead. NOMXC is exchange-based, but also bilateral or broker-based OTC trades are registered, and for some products one or more market makers post two-sided quotes in the order book.

In this paper we consider the set of base load forward contracts traded on NOMXC with the NPS as the underlying reference price. This is by far the most liquid subset of the traded products. Market participants enter into contracts with a specified delivery period; monthly, quarterly or yearly, and contracts kept until expiry are settled financially on every clearing day during this period. No physical supply or receipt of power occurs. Today, prices are quoted in EUR/MWh (contracts with a delivery period prior to 2006 were quoted in NOK/MWh), each contract specifies delivery of a continuous flow of power during the delivery period of 1MW, and contracts are cash settled to the average spot price.[6] Monthly contracts trade until the last trading day before the specified delivery period, and settlement is spot reference cash on every clearing day during the delivery period. Quarterly contracts are cascaded from yearly contracts and cascaded into monthly contracts on the last trading day before the specified delivery period. Yearly contracts are cascaded into quarterly contracts three trading days prior to the delivery period. All settling of accounts takes place through NOMXC, and the two parties involved do not know each other's identity. The settling of accounts is guaranteed by NOMXC. Swedish Vattenfall acts as market maker on base load forward contracts with commitments to continuously quote buy and sell prices in the order book within a maximum spread determined by volatility and price levels.[7]

3 Data

We have available transaction data for all contracts traded on NOMXC in the period 2 January 2006 to 31 May 2011 delivering 1359 distinct trading days. The raw data files have a total of $2,107,919$ transactions with information as outlined in the Appendix. We remove products which do not have the Nordic system price as spot reference (products on EEX Phelix, APX, and UK as well as trades in different European Union Allowances (EUAs) and Certified Emission Reductions (CERs) are also registered, see NASDAQ OMX Commodities (2011*a*)). A deal in CLICK Trade (the trading application provided by NOMXC) may consist of two or more transactions, i.e. there is not a one-to-one relation between buy and sell as one offer may hit several bids if price and volume match. Therefore, we remove duplicated transactions according to *DealNumber*, as they contain the same information. This leaves $954,248$ unique transactions.

[6]The structure of derivatives traded on Nord Pool/NOMXC has gone through a number of adjustments since the market opening but has remained unchanged since January 2006.

[7]See NASDAQ OMX Commodities (2011*a*) for details on traded products, and NASDAQ OMX Commodities (2011*b*) and NASDAQ OMX Commodities (2011*c*) for details on trading and clearing on NOMXC. Also, see nordpoolspot.com, nasdaqomxcommodities.com, and eex.com for further details on this Section.

The left part of Table 1 lists the division when sorted according to *DealSource* and *MainCategory*. The most traded contracts are among the base load futures and forwards. A little surprising, the trade in peak loads is limited, and also options are rarely traded. The time stamps on OTC ticks are imprecise and make them unfit for an intraday analysis. Another reason for exclusion is the often doubtful transaction prices. As an example, the ENOMMAR-08 contract was traded to 1 EUR/MWh and 528.6 EUR/MWh at the same time on 11 February 2008 implying that the trade may have been a part of a bigger deal. Focusing on exchange-traded base load futures and forwards alone we are left with 705,548 ticks.

Table 1: Overview of transactions in derivative contracts traded on NOMXC.

Product Specification	OTC Transactions	Exchange Transactions	Length of Delivery Period	Transactions
Base	212,365	699,928	Day (future)	5,620
BaseDay	155	5,620	Week (future)	31,278
CfD	19,975	11,038	Month (forward)	99,953
Option	5,064	23	Quarter (forward)	424,272
Peak	36	44	Year (forward)	144,425
	237,595	716,653		705,548

The sampling period is 02 January 2006 to 31 May 2011. Left part: Number of transactions sorted according to *DealSource* (OTC: Over The Counter, Exchange) and *MainCategory* (Base: forward and future base load contracts with a weekly, monthly, quarterly, or yearly delivery period, BaseDay: future base load contracts with a daily delivery period, CfD: Contract for Differences, Option: European call and put options (written on quarterly forwards), Peak: forward and future peak load contracts with a weekly, monthly, quarterly, or yearly delivery period). Right part: Number of exchange transactions in future and forward base load contracts sorted according to length of delivery period.

From the right part of Table 1 quarterly contracts are seen to be the most traded and the futures are the least traded. With liquidity deemed important, we limit ourselves to consider the most traded contracts, the forwards. In the period 2 January 2006 to 31 May 2011, the specifications (delivery period, contract size, currency quote, etc.) of the forwards have remained unchanged. Further, the opening hours of the exchange have remained the same and no half trading days are present as in some other markets.

3.1 Realized Measures of Volatility

We consider a collection of different realized measures of volatility below, which we briefly introduce in this section. We follow the implementation of the authors of the original paper as closely as possible and omit detailed descriptions of each estimator in the interest of space. Instead, the interested reader is referred to the original papers.

The fundamental theory of asset prices says that the log-price at time t, $p_t^\star$, must, in a frictionless arbitrage-free market, obey a semi-martingale process (see e.g. Back (1991) for an introduction). The object of interest, that realized measues are aimed at estimating, is usually the quadratic variation (QV) of $p^\star$ defined by

$$[p^\star]_t = \text{plim}_{n\to\infty} \sum_{j=1}^{t_j \le t} \left(p^\star_{t_j} - p^\star_{t_{j-1}} \right)^2 ,$$

for any sequence of deterministic partitions $0 = t_0 < t_1 < \ldots < t_n = T$ with $\sup_j \left\{ t_{j+1} - t_j \right\} \to 0$ for $n \to \infty$ (see e.g. Protter (2005)). We use subscript index j to refer to the (possibly) irregular spaced sequence of the raw data series at day t. That is, we may partition the interval $[0, T]$ into m sub-intervals, such that for a fixed m, the jth sub-interval is given by $\left[t_{j-1,m}, t_{j,m} \right]$, where $0 = t_{0,m} < t_{1,m} < \ldots < t_{m,m} = T$. The time gap, i.e. the length of the jth sub-interval, is given by $\delta_{j,m} := t_{j,m} - t_{j-1,m}$, such that the length of each sub-interval shrinks to 0 as m increases. The intraday log-returns are defined by

$$r^\star_{t_{j,m}} := p^\star_{t_{j,m}} - p^\star_{t_{j-1,m}}, \quad j = 1, \ldots m,$$

and calculated from intraday log-prices spanning the period 8:00 AM to 15:30 PM (the hours the exchange is open). We often omit subscripts on t for convenience. The standard realized variance of $p^\star$ is defined as

$$RV^{(m)}_\star := \sum_{j=1}^{m} \left(r^\star_{j,m} \right)^2 ,$$

and it is consistent for the QV as $m \to \infty$ (see e.g. Protter (2005)), i.e. the realized variance yields a precise estimate of volatility when prices are observed continuously and without measurement error. A feasible asymptotic distribution theory was established by Barndorff-Nielsen and Shephard (2002). $RV^{(m)}_\star$ is ideal, but infeasible as $p^\star$ is latent. For the observable price process, p, we observe on day t the RV based on m intraday returns, which is generally inconsistent for the QV (see e.g. Zhang et al. (2005)).

In our analysis, we consider various versions of $RV^{(m)} := \sum_{j=1}^{m} \left(r_{j,m} \right)^2$, where $r_{j,m}$ is the observed intraday return, i.e. a noisy version of $r^\star_{j,m}$. The estimates are based on different sampling schemes. When $t_{j,m}$, $j = 0, \ldots, m$, are equidistant in calender time it is referred to as calender time sampling (CTS), which we implement using previous-tick interpolation (see e.g. Hansen and Lunde (2006)). Under CTS, we write $RV^{(x\,sec)}$, when $\delta_{j,m} = x$ seconds. We select two sampling frequencies: 5 min, $RV^{(300\,sec)}$, and 30 min, $RV^{(1800\,sec)}$, and in some instances we make use of sub-sampling, which is indicated by an asterisk in the superscript, e.g. $RV^{(300\,sec)*}$. When $t_{j,m}$ denotes the actual time of a transaction we have tick time sampling (TTS), and one may choose $t_{j,m}$, $j = 0, \ldots, m$ to be the time of every second transaction, say. Under TTS, we write $RV^{(y\,tick)}$ when each intraday return spans y ticks.

As noted above, the realized variance is sensitive to market frictions when applied to returns recorded over shorter time intervals. To our knowledge, this has not been studied for NOMXC forwards. Realized kernels are a broad set of estimators designed to be robust to most types of frictions. The realized kernel of Barndorff-Nielsen et al. (2011) is in the univariate case with time gap $\delta_m > 0$ given by

$$K\left(p_{\delta_m}\right) = \sum_{h=-m}^{m} k\left(\frac{h}{H+1}\right)\Gamma_h, \quad \Gamma_h = \sum_{j=|h|+1}^{m} r_{t_{j,m}} r_{t_{(j-h),m}},$$

where $k(x)$, for $x \in \mathbb{R}$, is a non-stochastic weight function and Γ_h is the hth realized autocovariance. In the implementation, we follow the suggestions in Barndorff-Nielsen et al. (2011) with respect to the jittering of the initial and final time points each day, the selection of the optimal bandwith, H, for the recommended "non-flat-top Parzen" kernel as the choice of $k(x)$, and for the estimation of the noise-to-signal-ratio ξ^2 by $\xi^2 = \frac{\left(RV^{(1\,tick)}/(2 \cdot N)\right)}{RV^{(1800\,sec)}}$.[8] Below, we refer to this simply as the realized kernel (RK).

3.2 Rollover Schemes

The limited lifespan of the individual forwards creates a need for rollover schemes between the multiple time series, covering different trading days, in order to create continual time series. At any point in time, there will be contracts trading for several expirations. A rollover scheme defines the linking procedure between them, i.e. the point in time at which one contract is switched for the subsequent one (e.g. ENOQ1-10 to ENOQ2-10). We refer to this point in time as the rollover date with the last possible rollover date being the last trading day of the expiring contract. To the best of our knowledge, no rigorous theoretical justification has yet suggested that one linking method is better than the other, and there does not appear to be a consensus in the literature regarding the exact procedure for creating a rollover scheme (see e.g. Ma et al. (1992), Holton (2003), and references herein). It depends on the use to which the data will be put. In our case liquidity is the primary concern, as we work with realized measures of volatility. This affects our choice below. Alternatives to rollover schemes are "artificial" series constructed from constant-maturity prices, the most liquid contract, the mean of all traded contracts, etc. However, we want the continual series to reflect a prespecified trading strategy, e.g. holding a contract and switching to the next contract on a prespecified day before the expiry of the current contract, that does not rely on ex post measures.[9]

We refer to a first nearby as the time series comprising the price of the nearest-to-expiration

[8] We set $H = c^{\star}\xi^{4/5}m^{3/5}$, where $c^{\star} = \left\{k''(0)/k_{\bullet}^{0,0}\right\}^{1/5}$ and $k_{\bullet}^{0,0} := \int_0^1 k(x)^2\,dx$.

[9] The example given is obviously not a self-financing trading strategy. It will lead to negative (positive) "roll yield" in contango (backwardation) and may be affected by seaonality. However, such strategies are used by many traders in different disguises in markets where rollover is necessary. This includes exchange-traded funds which are responsible for a substantial percentage of the traded volume in e.g. some oil futures markets.

contract. The second nearby comprises the price, at each point in time, of the second nearest-to-expiration contract, etc. These are possible outcomes of rollover schemes that are easy to implement. In the final days prior to a contract's last trading date, distortions can be pronounced, as positions are closed out and liquidity migrates to the subsequent contract. Thus, the number of transactions may provide indications on the behavior of traders that close positions. Table 2 summarizes the liquidity on average in the final days prior to maturity (recall the earlier cascading of the yearly contracts).

Table 2: Liquidity in base load forwards traded on NOMXC.

Monthly Forwards

	First Nearby			Second Nearby		
$T-t$	#	$\overline{RK}$	$\overline{RV}^{(1tick)}$	#	$\overline{RK}$	$\overline{RV}^{(1tick)}$
0	44	4.429	3.819	25	2.616	2.715
1	36	3.213	3.180	23	3.106	3.194
2	33	3.399	2.946	22	2.913	2.511
3	39	3.119	3.380	21	3.961	3.676
4	39	3.783	3.412	19	3.253	3.201
5	34	2.922	2.681	16	2.183	2.150
6	37	3.528	3.035	20	2.929	2.690
7	37	3.672	3.389	20	3.082	2.995
8	34	3.568	3.498	21	3.222	2.993
9	36	3.120	3.053	18	2.709	2.847
10	38	3.272	2.841	17	2.556	2.765

Quarterly Forwards

	First Nearby			Second Nearby		
$T-t$	#	$\overline{RK}$	$\overline{RV}^{(1tick)}$	#	$\overline{RK}$	$\overline{RV}^{(1tick)}$
0	73	3.230	2.279	137	3.252	2.332
1	80	1.984	1.997	109	2.590	2.337
2	86	2.201	2.129	87	2.122	1.928
3	124	2.633	2.355	90	2.623	2.178
4	137	3.535	2.577	77	2.411	2.411
5	173	3.627	2.944	65	2.411	2.133
6	181	3.910	3.332	72	2.856	2.416
7	187	3.658	3.361	71	2.556	2.646
8	206	3.665	3.279	64	2.585	2.590
9	225	3.670	3.052	61	2.300	2.503
10	237	4.073	3.638	56	2.568	2.665

Yearly Forwards

	First Nearby			Second Nearby		
$T-t$	#	$\overline{RK}$	$\overline{RV}^{(1tick)}$	#	$\overline{RK}$	$\overline{RV}^{(1tick)}$
-	0	–	–	29	0.316	0.270
-	0	–	–	23	0.562	0.876
0	26	1.030	0.603	26	0.406	0.359
1	38	1.363	1.393	24	0.519	0.509
2	53	2.119	1.659	33	1.500	0.942
3	52	1.412	1.642	33	0.843	0.862
4	48	1.261	1.409	23	1.074	1.235
5	63	2.007	1.904	30	1.145	1.343
6	66	1.595	1.527	33	1.086	1.306
7	66	1.798	1.675	49	1.117	1.231
8	67	3.396	2.834	31	1.807	1.327

The sampling period is 02 January 2006 to 31 May 2011. The first nearby refers to the time series comprising the price, at each point in time, of the nearest-to-expiration contract. The second nearby comprises the price, at each point in time, of the second nearest-to-expiration contract. $T-t$ is time-to-maturity in days for the first nearby. The number of transactions (#), the realized kernel (RK) of Barndorff-Nielsen et al. (2011), and the realized variance (RV) are daily averages over the sampling period for the given time-to-maturity.

For the monthly and yearly series liquidity does not appear to migrate. That is, the first nearby remains the most liquid until it expires. For the quarterly series the second nearby becomes the most traded two days prior to the maturity day of the first nearby on average. This may indicate that speculators, with no interest in daily cash settlements during the delivery period, are more active in the quarterly contracts. Further, we notice that the market for the second nearby quickly becomes thin for the quarterly contract as time-to-maturity is increased. Thus, to avoid thin market concerns we choose the rollover date to be the maturity day for months and years, and two days prior to maturity for quarters. The heterogeneity of consecutive contracts introduces characteristics in the data which are pure artifacts of the rollover scheme. The literature is rich on suggestions on how to remedy this, e.g. scaling the old or the new contract. We stick to the trading strategy and propose in Section 4 different ways to cope with the inherent seasonality, and thus avoiding that artificially large positive and negative returns muddle the results.

3.3 Stylized Facts for the Price Process at the Daily Frequency

The existing literature on modeling and forecasting volatility in power futures and forwards has mainly utilized observations at a daily frequency (see e.g. references in the introduction and also Malo and Kanto (2006), Benth et al. (2008*a*), and Pen and Sévi (2010)). The classical GARCH framework often utilizes daily returns (or lower frequencies) to extract information about the current and future level of volatility. To motivate the application of GARCH models we document in this section the common stylized facts of financial series at the daily frequency: unpredictability of returns, volatility clustering, leptokurtosis, asymmetries, and so forth. In the discrete Realized GARCH framework we use r_t^{oc} and r_t^{cc} to denote open-to-close and close-to-close daily returns at day t, respectively, with realized measures of volatility denoted by x_t. Open-to-close returns are the daily changes in the logarithm of the first and last transaction price each day, and close-to-close returns at time t are the daily changes in the logarithm of the last transaction price at time $t-1$ and t. The information set is then given by $\mathcal{F}_t = \{r_t^{oc}, r_t^{cc}, x_t, r_{t-1}^{oc}, r_{t-1}^{cc}, x_{t-1}, \ldots\}$, which of course is richer than in the conventional GARCH framework. Thus, the Realized GARCH should be more responsive than conventional GARCH models.

In Figure 1 we present the continual monthly, quarterly, and yearly series in levels, and in Figure 2 the daily changes in the logarithm of these, close-to-close returns, all at the daily frequency with rollover schemes as defined in Section 3.2. Returns that straddle a rollover date are indicated as a black dot.[10] We see that the level series, covering different delivery horizons, have similar patterns over time, but with changes more pronounced in the contracts with shorter delivery periods. This is especially clear in the second half of the period. A possible reason being that contracts with delivery close by is more affected by news and changes in fundamentals or the arrival of such is more related to these contracts. In the later years, the monthly and quarterly contracts have not traded with a discount in spring as compared to the yearly contracts, which was the case in 2008 and 2009. Weather fundamentals ("dry winters") being a likely explanation.

From Figure 2 we encounter periods with large changes in absolute sense (often characterized as volatility clustering), and the daily changes appear more erratic for the monthly series and less for the yearly. Returns that straddle rollover dates are "extracted" from the series for clarity. Some are large in absolute value (some are even outside the range of the plot), but a seasonal pattern is only partly evident. For example, the monthly rollover returns are mainly positive from late spring and then turn negative in the first months each year. The quarterly

[10] Some observations are outside the chosen range: rollover from ENOMAUG-07 to ENOSEP-07 resulted in a return of around 21 pct., ENOSEP-07 to ENOOCT-07 17 pct., ENOOCT-07 to ENONOV-07 25 pct., ENOMAY-08 to ENOJUN-08 17 pct., ENOJUN-08 to ENOJUL-08 25 pct., ENOJUL-08 to ENOAUG-08 31 pct., ENOMAR-10 to ENOAPR-10 -23 pct., ENOJAN-11 to ENOFEB-11 -16 pct., ENOQ3-07 to ENOQ4-07 32 pct., ENOQ4-07 to ENOQ1-08 30 pct., ENOQ2-08 to ENOQ3-08 16 pct., ENOQ3-08 to ENOQ4-08 34 pct., ENOQ1-11 to ENOQ2-11 -23 pct., and ENOY-11 to ENOY-12 -20 pct.

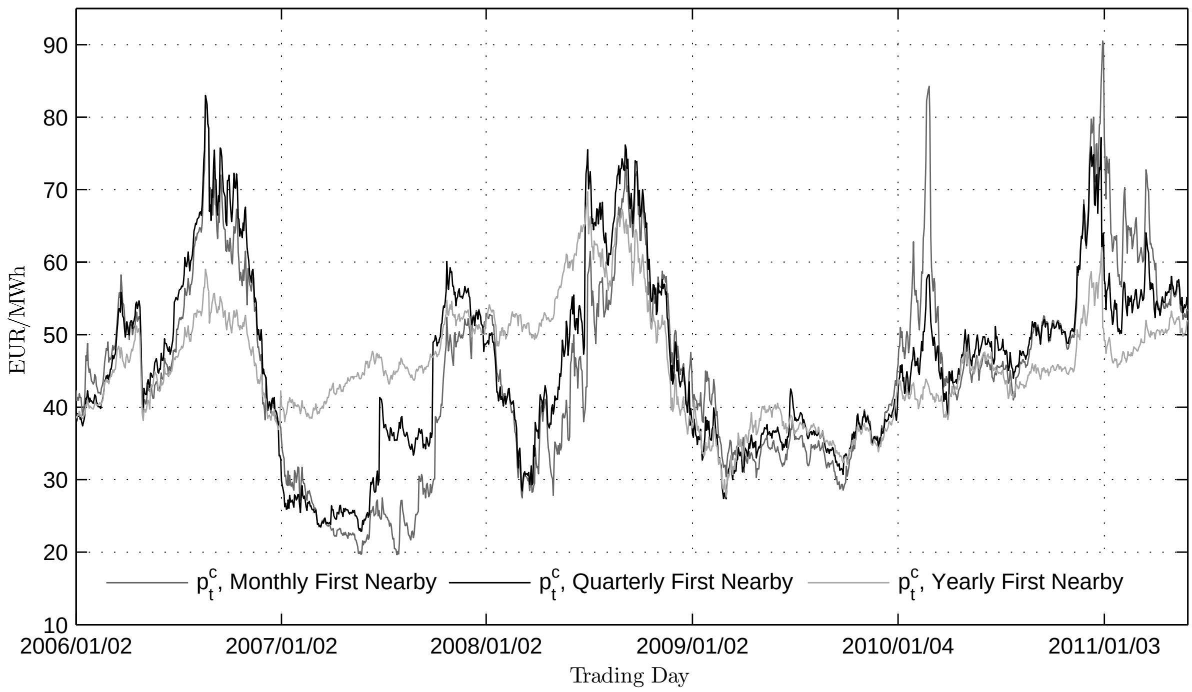

Figure 1: Time series in levels of base load forwards traded on NOMXC. The sampling period is 02 January 2006 to 31 May 2011 and results are reported at a daily frequency. All series are constructed using rollover schemes as specified in Section 3.2.

rollover returns have a similar pattern. However, the changes are quite different from year to year.

In Table 3 we present summary statistics along with a range of diagnostic tests. Rollover returns are omitted from the close-to-close returns as test statistics are sensitive to observations large in value. Excluding rollover returns in r_t^{cc} removes the largest observations (in absolute sense) such that the minimum, the maximum and the kurtosis are roughly similar for r_t^{oc} and r_t^{cc}. This hints that the information flow outside the market opening hours does not generate large (absolute) price changes, which would show up as outliers in r_t^{cc} (if present they are damped over the trading day). On the other hand, r_t^{cc} has a slightly higher standard deviation and is as such more volatile. Normality is clearly rejected in all cases by the Jarque-Bera test. This was expected for the skew level series and the slightly skew and leptokurtic returns. We are unable to reject the presence of auto-correlation in some of the series, which motivates an ARMA structure for the conditional mean. Overall, the small empirical autocorrelations, in most cases insignificant from zero, render the returns unpredictable. For the squared series the null of no autocorrelation is readily rejected in all r_t^{oc} and r_t^{cc} series, which points towards the presence of (G)ARCH effects and the possibility of volatility prediction. For the price levels we are unable to reject the presence of a unit root for all specifications of the DGP and the model in the ADF-regression using augmented Dickey-Fuller tests with automatic lag length selection. The presence of a unit root is readily rejected for all return series.

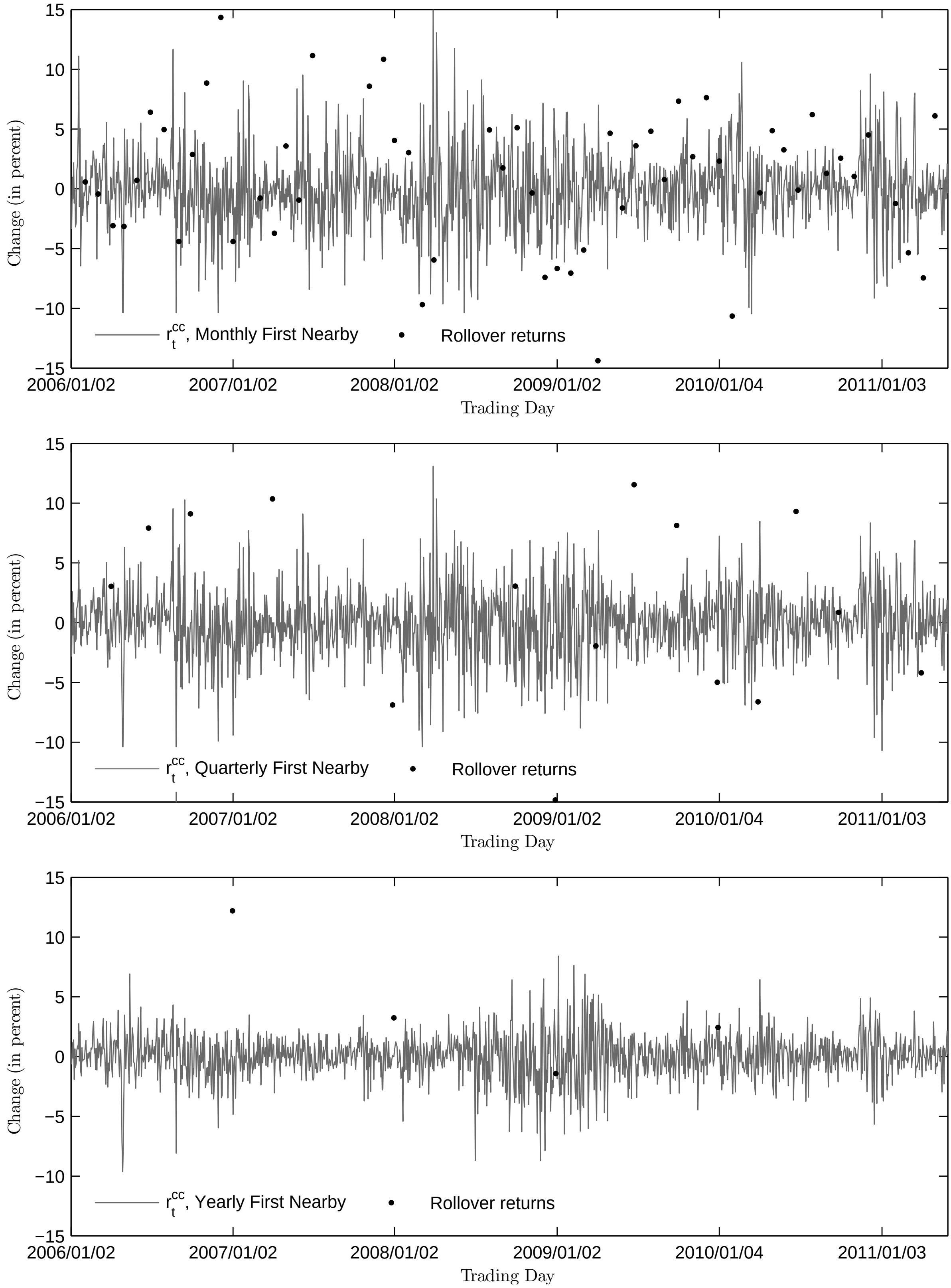

Figure 2: Log-returns of base load forwards traded on NOMXC. The sampling period is 02 January 2006 to 31 May 2011, and results are reported at a daily frequency. The top row presents the monthly first nearby, the middle row the quarterly first nearby, and the bottom row the yearly first nearby. All with rollover schemes as specified in Section 3.2. Returns that straddle a *rollover date* are indicated as a black dot with a few observations outside the chosen range (see Footnote 10).

Table 3: Descriptives and diagnostic tests for base load forwards traded on NOMXC.

	Levels			Open-to-Close Returns			Close-to-Close Returns		
	Month	Quarter	Year	Month	Quarter	Year	Month	Quarter	Year
Mean	45.010	46.102	45.786	−0.066	−0.021	−0.016	−0.084	−0.050	0.024
Median	44.800	45.700	45.200	0.000	0.000	0.000	0.000	0.000	0.099
Min	19.500	22.860	27.600	−8.931	−13.911	−9.118	−12.104	−15.649	−9.639
Max	90.500	83.000	69.700	10.407	11.902	8.829	15.536	13.090	8.413
Std.Dev.	13.496	12.518	7.540	2.177	2.268	1.574	3.073	2.858	1.888
Skewness	0.338	0.361	0.458	0.087	−0.118	−0.016	0.128	−0.189	−0.387
Kurtosis	2.764	2.585	2.996	5.104	5.895	6.924	5.157	5.061	5.840
Jarque − Bera	29.094**	39.237**	47.593**	252.454**	477.819**	872.147**	267.103**	248.520**	490.264**
$ACF(1)$	0.991**	0.991**	0.992**	0.105**	−0.022	−0.079**	0.116**	0.054**	0.050
$ACF(2)$	0.979**	0.981**	0.983**	0.014	−0.016	−0.102**	0.030	−0.012	−0.051
$ACF(4)$	0.955**	0.961**	0.968**	0.009	0.005	0.079**	0.003	0.001	0.043
$ACF(5)$	0.943**	0.952**	0.960**	−0.058**	−0.047	−0.058**	−0.006	−0.017	−0.017
$Q(4)$				17.208**	1.035	31.433**	21.728**	4.966	9.929*
$Q(12)$				31.965**	9.975	52.378**	25.444*	15.318	14.319
$Q^2(4)$				100.929**	52.936**	263.093**	41.879**	40.315**	139.715**
$Q^2(12)$				296.237**	203.358**	478.923**	164.109**	169.655**	363.200**
$Aug.DF^*$	−0.533	−0.486	0.005	−33.100**	−37.621**	−12.872**	−32.751**	−34.846**	−26.710**
$Aug.DF^\star$	−2.472	−2.784	−2.303	−33.115**	−37.611**	−12.887**	−32.759**	−34.843**	−26.706**
$Aug.DF^\bullet$	−2.593	−2.861	−2.315	−33.142**	−37.659**	−12.996**	−32.805**	−34.885**	−26.745**
#Obs	1359	1359	1356	1359	1359	1356	1358	1358	1355

The sampling period is 02 January 2006 to 31 May 2011. The Jarque-Bera (JB) tests the null of normality and ACF(L) the empirical auto-correlation function at lag L (only lag lengths with estimates significantly different from zero for one or more of the returns series are shown). $Q(L)$ is the Ljung-Box test of no auto-correlation in up to L lags and $Q^2(L)$ is the Ljung-Box test on squared log-returns to test for homoscedasticity. The augmented Dickey-Fuller (DF) tests use automatic lag selection (by the Akaike Information Criterion) and have a DGP and an estimated model in the ADF-regression with no deterministic trends ($Aug.DF^*$), no deterministic trend in the DGP but a constant and a time-trend in the estimated model ($Aug.DF^\star$), and a constant or time trend in the DGP and both in the estimated model ($Aug.DF^\bullet$). Two asterisks indicate rejection at 1 percent significance level (for ACF that the estimate is outside the 95 pct. confidence band), and one asterisk rejection at the 5 percent level.

4 Econometric Methodology

We propose to model returns and realized measures of volatility jointly within the Realized GARCH framework of Hansen et al. (2012*a*). The key variable of interest is the conditional variance $h_t = \text{Var}\left[r_t \,|\mathcal{F}_{t-1}\right]$. In the Realized GARCH framework, h_t depends on its own (truncated) past as usual, but as opposed to lagged squared returns included in the traditional framework, the Realized GARCH incorporates a realized measure of volatility (or a vector of these), x_t.[11] As such, the model defines a class of GARCH-X models as in Engle (2002) and Barndorff-Nielsen and Shephard (2007), where x_t is exogenous.[12] However, an additional equation that ties the realized volatility measure to the latent volatility completes the model, and the dynamic properties of both returns and the realized measure are specified. We use a variant of the Realized

[11] Hansen et al. (2012*a*) finds that the ARCH-term is insignificant in the Realized GARCH model and we omit it in the model formulation and in the empirical analysis.

[12] See also Hansen et al. (2012*a*) for a comparison to the Multiplicative Error Model (MEM) by Engle and Gallo (2006) and the HEAVY model by Shephard and Sheppard (2010).

GARCH model suggested and analysed in Hansen and Huang (2012) and used in Hansen et al. (2012*b*). This model is called the Realized EGARCH as it shares certain features with the EGARCH model of Nelson (1991). As outlined below, the Realized GARCH framework can be compared to nested and more standard GARCH models, which provides an elegant way to certify possible benefits from utilizing high-frequency data in a particular market. In applications Hansen et al. (2012*a*) showed with DJIA stocks that the Realized GARCH class led to substantial improvements in in-sample and out-of-sample empirical fit. Similarly, the chosen subclass here can verify the potential in the unexplored transaction data considered in this paper. We stress this as the primary interest of the paper and a horserace between competing models is left for future research.

4.1 The Realized EGARCH Model

We follow the exposition in Hansen and Huang (2012) for convenience. Let $\tilde{h}_t = \log h_t$ and $\tilde{x}_{k,t} = \log x_{k,t}$. Then the Realized EGARCH model for returns and K realized measures of volatility is given by the following mean, GARCH and measurement equation(s), respectively

$$\begin{aligned} r_t &= \mu_t + e^{-\tilde{h}_t/2} z_t, \\ \tilde{h}_t &= \omega + \beta \tilde{h}_{t-1} + \tau\left(z_{t-1}\right) + \gamma' \tilde{\mathbf{x}}_{t-1}, \\ \tilde{x}_{k,t} &= \xi_k + \varphi_k \tilde{h}_t + \delta_k\left(z_t\right) + u_{k,t}, \qquad k = 1, \ldots K. \end{aligned} \tag{1}$$

The autocorrelation documented in Table 3 for some of the series motivates the inclusion of autoregressive terms in μ_t, and rollover dummies or other exogenous terms can similarly be included. We assume market efficiency, i.e. that such terms are already incorporated in the price, and a refinement of the mean equation is left for future research. Similarly, we include results for multiple lags in the GARCH equation. The so-called Samuelson effect is often claimed to be present in energy markets (see e.g. Benth et al. (2008*b*)), which made us include different time-to-maturity dependencies in the GARCH equation. However, none of these were found to be significantly different from zero. The absence of the Samuelson effect is further stressed in Table 2, where the realized measures of volatility do not show signs of time-dependence. In general, the model allows for the inclusion of exogenous variables in the GARCH equation. In the measurement equations the inclusion of multiple realized measures of volatility is a neat way to show the superiority of some measures over others. The functions $\tau\left(z_t\right)$ and $\delta_k\left(z_t\right)$ are called leverage functions, because they model aspects related to the leverage effect, which refer to the dependence between returns and volatility.[13] Hansen et al. (2012*a*) found that a simple

[13] We refer to the *leverage effect* as the asymmetry in volatility following big price increases and decreases, respectively. In the classic Black 76' leverage story an increase in financial leverage level leads to an increase in equity volatility level with business risk held fixed. A financial leverage increase can come from stock price decline while the debt level is fixed. In energy markets there is evidence of a so-called *inverse leverage effect*. The volatility tends to increase with the level of prices, since there is a negative relationship between inventory and prices (see for instance

second-order polynomial form provides a good empirical fit and that $\log z_t^2$ was inferior to z_t^2. We will adopt this structure and set $\tau(z_t) = \boldsymbol{\tau}' a(z_t)$ and $\delta_{(k)}(z_t) = \boldsymbol{\delta}_k' b(z_t)$, $k = 1, \ldots, K$, where $a_t = b_t = \begin{bmatrix} z_t & z_t^2 - 1 \end{bmatrix}'$ such that we can express the GARCH and measurement equation(s) as

$$\begin{aligned} \tilde{h}_t &= \boldsymbol{\lambda}' \mathbf{g}_{t-1}, \\ \tilde{x}_{k,t} &= \boldsymbol{\psi}_k' \mathbf{m}_t + u_{k,t}, \qquad k = 1, \ldots K, \end{aligned} \tag{2}$$

where

$$\boldsymbol{\lambda} = \begin{bmatrix} \omega & \beta & \boldsymbol{\tau} & \gamma \end{bmatrix}', \mathbf{g}_t = \begin{bmatrix} 1 & \tilde{h}_t & a_t & \tilde{\mathbf{x}}_t \end{bmatrix}', \boldsymbol{\psi}_k = \begin{bmatrix} \xi_k & \varphi_k & \boldsymbol{\delta}_k \end{bmatrix}', \mathbf{m}_t = \begin{bmatrix} 1 & \tilde{h}_t & b_t \end{bmatrix}'.$$

$\mathrm{E}[\tau(z_t)] = 0$ and $\mathrm{E}\left[\delta_{(k)}(z_t)\right] = 0$, $k = 1, \ldots, K$, for any distribution of z_t as long as $\mathrm{E}[z_t] = 0$ and $\mathrm{Var}[z_t] = 1$. We adopt a Gaussian specification by assuming $z_t \sim N(0, 1)$ and $\boldsymbol{u}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ with z_t and $\boldsymbol{u}_t$ mutually independent. This is motivated by the empirical fact that realized volatility is approximately log-normal and that returns standardized by realized volatility are close to normal, respectively. We standardize returns by the conditional variance, which incorporates the realized measure. The Gaussian specification may be too restrictive and for that reason we compare regular and bootstrapped forecasts in Section 6. That is, the Gaussian specification is not crucial to the estimation, and as such we follow Hansen and Huang (2012) and apply a QMLE approach to assess the precisions of the parameter estimates. The conjectured limiting theory follows with minor modifications from Theorem 2 in Hansen and Huang (2012).

The "intercept" ξ and "slope" φ add flexibility to the measurement equation, which is required when we link realized measures of volatility that span a shorter period than the return. As long as x_t and h_t are roughly proportional we should expect $\varphi \approx 1$ and $\xi < 0$. The presence of z_t in the measurement equation provides a simple way to model the joint dependence between r_t and x_t. Tying the realized measure to the conditional variance is nicely motivated by the fact that the return equation implies that $\log r_t^2 = \log h_t + \log z_t^2$, and because x_t is similar to r_t^2 in many ways, albeit a more accurate measure of h_t, one may expect that $\log x_t \approx \log h_t + f(z_t) + error_t$. This motivates a logarithmic measurement equation, which further makes a logarithmic GARCH equation convenient. A logarithmic specification automatically ensures a positive variance, and as $\log r_{t-1}^2$ does not appear in the GARCH equation (it is replaced by $\log x_{t-1}$) zero returns do not cause havoc for the specification.

Equation (1) gives a GARCH equation similar to an EGARCH-type model motivating the

Deaton and Laroque (1992)). Little available inventory means higher and more volatile prices. A similar economic interpretation does not apply in the forward market, where a leverage effect, if present, must be contributed to different investor behavior for prices moving up and down, respectively.

benchmark

$$\tilde{h}_t = \tilde{\boldsymbol{\lambda}}' \tilde{\mathbf{g}}_{t-1}, \qquad \text{where} \quad \tilde{\boldsymbol{\lambda}} = \begin{bmatrix} \omega & \beta & \boldsymbol{\tau} \end{bmatrix}' \quad \text{and} \quad \tilde{\mathbf{g}}_t = \begin{bmatrix} 1 & \tilde{h}_t & a_t \end{bmatrix}' \tag{3}$$

The log-likelihood function of (1) can be expressed such that it can be directly compared to the log-likelihood function of (3). Therefore, we can easily detect whether realized measures of volatility lead to improved empirical fit. In-sample in terms of the log-likelihoods in optimum, which is the goal of Section 5, and out-of-sample in terms of the forecasting performance, which is the goal of Section 6.

4.2 Estimation

To estimate the parameters in (1) we consider the quasi log-likelihood function given by

$$\ell\left(r, \boldsymbol{x}; \boldsymbol{\theta}, \boldsymbol{\Sigma}_u\right) = -\frac{1}{2} \cdot \sum_{t=1}^{T} \left(\log 2\pi + \tilde{h}_t + z_t^2 + K \cdot \log 2\pi + \log |\boldsymbol{\Sigma}| + \boldsymbol{u}_t' \boldsymbol{\Sigma}^{-1} \boldsymbol{u}_t \right).$$

where $|\mathbf{A}|$ is the determinant of the matrix $\mathbf{A}$ and $\boldsymbol{\theta} = \begin{bmatrix} \tilde{h}_t & \mu & \boldsymbol{\lambda}' & \boldsymbol{\psi}_1' & \ldots & \boldsymbol{\psi}_K' \end{bmatrix}'$ is the model parameters. Thus, we treat the initial value of the conditional variance as unknown. As usual, the value of $\boldsymbol{\Sigma}_u$ that maximizes the likelihood among the class of all symmetric positive definite matrices is $\hat{\boldsymbol{\Sigma}}_u(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{u}_t(\boldsymbol{\theta})\, \boldsymbol{u}_t'(\boldsymbol{\theta})$, and as $\sum_{t=1}^{T} \boldsymbol{u}_t(\boldsymbol{\theta})\, \hat{\boldsymbol{\Sigma}}(\boldsymbol{\theta})^{-1} \boldsymbol{u}_t'(\boldsymbol{\theta}) = \operatorname{tr}\left[\sum_{t=1}^{T} \hat{\boldsymbol{\Sigma}}(\boldsymbol{\theta})^{-1} \boldsymbol{u}_t(\boldsymbol{\theta})\, \boldsymbol{u}_t'(\boldsymbol{\theta})\right] = TK$, which does not depend on $\boldsymbol{\theta}$, we can express the objective function as

$$\ell\left(r, \boldsymbol{x}; \boldsymbol{\theta}, \hat{\boldsymbol{\Sigma}}(\boldsymbol{\theta})\right) \propto \underbrace{-\frac{1}{2} \sum_{t=1}^{T} \left[\tilde{h}_t + z_t^2\right]}_{=\ell(r)} - \underbrace{\frac{T}{2} \log \left|\hat{\boldsymbol{\Sigma}}(\boldsymbol{\theta})\right|}_{=\ell(\boldsymbol{x}|r)}. \tag{4}$$

This reduces the parameter set that the optimizer has to search over and, using `C` for the recursive filter, the optimization is completed in a few seconds.[14] If one is only interested in one-step ahead modeling, specifying the measurement equation becomes redundant and the parameters in the model are further reduced. Standard GARCH models do not model $\boldsymbol{x}_t$, so the log-likelihood we obtain for these models (here the EGARCH model) cannot be compared to those of the Realized GARCH model. However, the expression for the log-likelihood above proves useful in this respect as $\ell(r)$ is a partial log-likelihood, which is directly comparable to the log-likelihood of standard models.

[14]For the results presented we experienced with different initial values and made sure results coincided. Overall, the optimization is fairly robust.

5 Estimation Results

In Table 4 we present the estimation results for the Realized EGARCH model with a range of specifications. The discussion is confined to the quarterly contract and focus on close-to-close returns.[15] We use open-to-close returns only for one specification of the Realized EGARCH and its benchmark to highlight characteristics of the parameters in the measurement equation. The realized kernel is primarily used as the realized measure of volatility (columns $2-9$), but results for five realized variances for one model specification is presented for comparison.[16] We denote by EG the conventional EGARCH in (3) and REG denotes the Realized EGARCH in (1) with the choice of p and q in parenthesis.

Focusing on parameter estimates, the realized measure loadings are large and the typical GARCH effects are of a smaller magnitude compared to conventional GARCH models. However, the lagged conditional volatility is still the dominant term. These findings are in line with results for individual stocks in Hansen et al. (2012*b*). The estimates of the parameters in the leverage function δ are similar to the ones reported in Hansen et al. (2012*a*) and Hansen et al. (2012*b*) and describe an asymmetric volatility response to positive versus negative shocks. However, the τ_1 estimate is in most cases found to be insignificant from zero, where Hansen et al. (2012*a*) and Hansen et al. (2012*b*) found negative parameter estimates. This is in line with the comments above on the interpretation of leverage effects in energy markets. The presence of a leverage effect in NOMXC forwards is questionable, and omitting τ leads only to a negligible drop in the log-likelihood function. On the other hand, omitting δ leads to a rather large drop in the log-likelihood.[17] The rollover effect was included in different ways and results are reported for the simplest one; one dummy in the mean equation. The estimate is in most cases not significantly different from zero and other parameter estimates are similar to those of the other specifications. The realized measures are based on data spanning the trading session only, and as expected ξ is negative and φ close to one in all cases. Notice also that the estimates are smaller (in absolute sense) as expected for the open-to-close series.

Comparing models, $\text{REG}(1,1)^{\star}$ has the best empirical fit in-sample measured by the log-likelihood, $\ell(r,x)$, and the realized kernel appear inferior. However, for the partial log-likelihood the results are more unclear, and the performance of models based on the realized kernel appear similar to that of models utilizing realized variances.[18] The partial log-likelihood is directly comparable to the log-likelihood of traditional EGARCH models, and the improvements are unquestionable for all series.[19] That is, utilizing transactions data improves the empirical fit

[15] The quarterly contracts are of particular interest as they are the more liquid subclass of forwards and because they serve as the underlying asset for exchange-traded options.

[16] $RV^{(1\,tick)}$, $RV^{(1800\,sec)}$, $RV^{(1800\,sec)*}$, $RV^{(300\,sec)}$ $RV^{(300\,sec)*}$ in columns $10-14$, respectively.

[17] However, omitting either of the leverage functions increases the search time for the optimal solution significantly.

[18] One should compare column 4 to columns $10-14$.

[19] One should compare column 1 to 2, and column 3 to 4.

Table 4: Results for the Realized EGARCH model for the quarterly first nearby for base load forwards traded on NOMXC.

	Open-to-Close Returns		Close-to-Close Returns											
Model:	EG(1,1)	REG(1,1)	EG(1,1)	REG(1,1)	REG(1,1)	REG(1,1)	REG(1,1)	REG(2,2)	REG(1,1)^{roll}	REG(1,1)*	REG(1,1)*	REG(1,1)$^{\bullet}$	REG(1,1)†	REG(1,1)‡
Panel A: Point Estimates														
$\log h_1$	1.730 (0.459)	1.472 (0.633)	1.108 (0.445)	1.135 (0.331)	1.135 (0.381)	1.066 (0.350)	1.410 (0.312)	0.524; 1.577 (0.484) (0.250)	1.141 (0.594)	1.334 (0.354)	0.898 (0.193)	0.878 (0.289)	1.164 (0.259)	1.148 (0.288)
μ	−0.002 (0.001)	0.034 (0.071)	0.063 (0.071)	0.090 (0.101)	0.091 (0.141)	0.093 (0.100)	0.082 (0.084)	0.086 (0.081)	0.025 (0.565)	0.103 (0.135)	0.097 (0.073)	0.089 (0.099)	0.097 (0.110)	0.099 (0.078)
$\phi_2; \phi_1$							0.047 (0.026)	−0.017; 0.051 (0.031) (0.028)						
η									4.274 (3.214)					
ω	0.059 (0.019)	0.240 (0.027)	0.265 (0.084)	0.779 (0.095)	0.773 (0.097)	0.769 (0.104)	0.774 (0.095)	0.230 (0.068)	0.729 (0.096)	0.880 (0.096)	0.543 (0.080)	0.733 (0.101)	0.786 (0.104)	0.750 (0.116)
$\beta_1; \beta_2$	0.960 (0.012)	0.633 (0.017)	0.891 (0.032)	0.538 (0.029)	0.542 (0.029)	0.542 (0.033)	0.539 (0.028)	−0.267; 1.129 (0.053) (0.079)	0.558 (0.031)	0.483 (0.020)	0.643 (0.029)	0.546 (0.034)	0.505 (0.025)	0.489 (0.022)
$\gamma_1; \gamma_2$		0.316 (0.017)		0.323 (0.072)	0.318 (0.070)	0.324 (0.072)	0.323 (0.072)	−0.241; 0.340 (0.066) (0.077)	0.309 (0.059)	0.411 (0.074)	0.291 (0.039)	0.338 (0.062)	0.363 (0.063)	0.363 (0.070)
τ_1	−0.010 (0.020)	0.000 (0.000)	−0.047 (0.061)	0.006 (0.005)		0.002 (0.017)	0.005 (0.009)	0.001 (0.002)	0.003 (0.010)	−0.003 (0.011)	0.004 (0.004)	0.009 (0.006)	0.005 (0.004)	0.006 (0.003)
τ_2	0.109 (0.017)	0.068 (0.006)	0.037 (0.020)	−0.002 (0.001)		−0.002 (0.004)	−0.002 (0.002)	0.001 (0.001)	0.001 (0.002)	0.000 (0.000)	−0.000 (0.000)	−0.002 (0.001)	−0.001 (0.003)	−0.001 (0.001)
ξ		−0.493 (0.060)		−1.991 (0.678)	−2.010 (0.687)	−1.970 (0.678)	−1.979 (0.677)	−1.999 (0.708)	−1.962 (0.614)	−1.771 (0.483)	−1.459 (0.380)	−1.739 (0.541)	−1.767 (0.524)	−1.662 (0.596)
φ		0.976 (0.036)		1.250 (0.274)	1.258 (0.277)	1.241 (0.272)	1.246 (0.274)	1.256 (0.284)	1.258 (0.254)	1.100 (0.203)	1.052 (0.166)	1.158 (0.222)	1.190 (0.221)	1.234 (0.245)
δ_1		−0.029 (0.012)		−0.062 (0.004)	−0.061 (0.004)		−0.061 (0.004)	−0.060 (0.006)	−0.040 (0.018)	−0.056 (0.003)	−0.064 (0.004)	−0.068 (0.004)	−0.063 (0.004)	−0.061 (0.004)
δ_2		0.159 (0.007)		0.014 (0.002)	0.014 (0.002)		0.013 (0.002)	0.013 (0.002)	0.012 (0.007)	0.013 (0.002)	0.024 (0.003)	0.017 (0.002)	0.016 (0.002)	0.014 (0.002)
Panel B: Log-Likelihood and Auxiliary Statistics														
$\ell(r,x)$		−3798.7		−4601.2	−4601.6	−4612.5	−4596.7	−4589.6	−4580.1	−4375.0	−4823.0	−4685.4	−4531.8	−4499.3
$\ell(r)$	−2926.0	−2906.0	−3569.0	−3501.7	−3501.6	−3501.6	−3497.1	−3494.4	−3475.8	−3500.4	−3486.7	−3492.7	−3500.9	−3503.8
$\hat{\sigma}_u^2$		0.218		0.296	0.297	0.301	0.297	0.295	0.298	0.213	0.420	0.340	0.268	0.254

The sampling period is 02 January 2006 to 31 May 2011. EG denotes the conventional EGARCH in (3) and REG denotes the Realized EGARCH in (1) with the choice of lags of $\tilde{h}_t$ and $\tilde{x}_t$ in parenthesis. Panel A contains parameter estimates. Panel B contains the full (when applicable) and the restricted log-likelihood, and $\hat{\sigma}_u^2$ is the estimated second moment of $\boldsymbol{u}_t$ as outlined in 4.2, i.e. for $K = 1$ we write $\hat{\sigma}_u^2$ in place of $\hat{\boldsymbol{\Sigma}}_u$. Columns one and two use open-to-close returns and the remaining columns uses close-to-close returns. The realized kernel is used as our realized measure of volatility in columns $2 - 9$ and REG(1,1)* uses $RV^{(1\,tick)}$, REG(1,1)* uses $RV^{(1800\,sec)}$, REG(1,1)$^{\bullet}$ uses $RV^{(1800\,sec)*}$, REG(1,1)† uses $RV^{(300\,sec)}$, and REG(1,1)‡ uses $RV^{(300\,sec)*}$. The underlined model specification is used in following sections. Robust standard errors in parenthesis are calculated from the numerical scores and Hessian matrix of the log-likelihood function following Theorem 2 in Hansen and Huang (2012).

of the model. To visualize the time-dependent level and the pattern of the conditional volatility, which stress the importance of active risk and portfolio management, we present in Figure 3 the resulting conditional variances over time for the underlined model specification in Table 4.

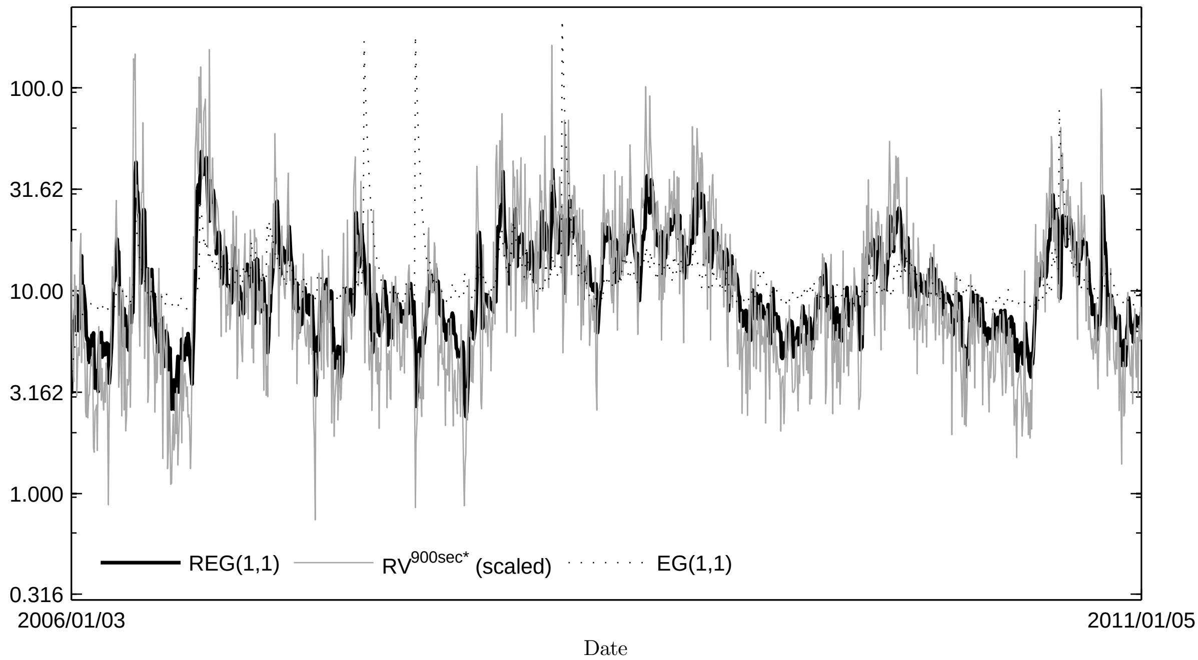

Figure 3: Conditional variances for base load forwards traded on NOMXC. The sampling period is 02 January 2006 to 31 May 2011. EG denotes the conventional EGARCH in (3) and REG denotes the Realized EGARCH in (1). Results are for the quarterly first nearby utilizing the realized kernel as the realized measure of volatility. The model specification is the one underlined in Table 4. The benchmark is the $RV^{(900\,sec)*}$(scaled) plotted in thick dim gray.

5.1 Multiple measurement equations

The Realized GARCH framework provides an elegant way to incorporate multiple realized measures of volatility and verifies possible superiority of some over others. We are especially interested in the comparison of the realized kernel with the different realized variances computed. In Table 5 we present results for pairwise comparisons with $RV^{(1\,tick)}$, $RV^{(1800\,sec)}$, $RV^{(1800\,sec)*}$, $RV^{(300\,sec)}$, and $RV^{(300\,sec)*}$. Parameter estimates are overall in line with those of Table 4. The γ_1 estimate to the left corresponds to the realized kernel. The estimates are of equal magnitude, and one is not superior to the other. The estimates of $\boldsymbol{\psi}_k = \begin{bmatrix} \xi_k & \varphi_k & \boldsymbol{\delta_k} \end{bmatrix}'$ are roughly equal for $k = 1$ and $k = 2$, respectively, where the left hand estimate belongs to the realized kernel. Finally, the large differences in log-likelihood values follow from det $\left(\hat{\boldsymbol{\Sigma}}_u\right)$ being well below 1.

Table 5: Results for the Realized EGARCH model with multiple measurement equations.

Model	REG(1,1)	REG(1,1)*	REG(1,1)$^{\bullet}$	REG(1,1)†	REG(1,1)‡
	Panel A: Point Estimates				
$\log h_1$	1.188 (0.292)	1.092 (0.292)	0.997 (0.530)	1.081 (0.387)	1.096 (0.356)
μ	0.095 (0.072)	0.090 (0.062)	0.088 (0.153)	0.092 (0.089)	0.093 (0.105)
ω	0.904 (0.106)	0.736 (0.088)	0.783 (0.100)	0.843 (0.103)	0.816 (0.114)
β_1	0.468 (0.023)	0.549 (0.032)	0.524 (0.031)	0.486 (0.024)	0.482 (0.023)
$\gamma_1; \gamma_2$	0.189; 0.205 (0.052) (0.039)	0.230; 0.104 (0.067) (0.034)	0.148; 0.198 (0.039) (0.042)	0.173; 0.195 (0.067) (0.042)	0.166; 0.197 (0.058) (0.036)
τ_1	−0.001 (0.001)	0.006 (0.007)	0.008 (0.009)	0.006 (0.004)	0.005 (0.007)
τ_2	−0.001 (0.001)	−0.003 (0.002)	−0.003 (0.002)	−0.002 (0.002)	−0.002 (0.003)
ξ	−1.978; −1.844 (0.614) (0.560)	−1.839; −1.726 (0.528) (0.513)	−1.869; −1.840 (0.595) (0.592)	−1.956; −1.831 (0.603) (0.594)	−2.026; −1.686 (0.650) (0.641)
φ	1.245; 1.130 (0.252) (0.229)	1.188; 1.160 (0.219) (0.212)	1.201; 1.198 (0.241) (0.240)	1.235; 1.217 (0.248) (0.245)	1.263; 1.245 (0.264) (0.260)
δ_1	−0.059; −0.058 (0.003) (0.003)	−0.062; −0.064 (0.005) (0.006)	−0.061; −0.068 (0.006) (0.007)	−0.059; −0.065 (0.004) (0.004)	−0.059; −0.064 (0.007) (0.007)
δ_2	0.014; 0.012 (0.002) (0.001)	0.014; 0.021 (0.002) (0.003)	0.013; 0.017 (0.002) (0.003)	0.014; 0.015 (0.002) (0.002)	0.014; 0.013 (0.002) (0.002)
	Panel B: Log-Likelihood and Auxiliary Statistics				
$\ell(r,x)$	−4710.0	−5102.6	−4742.9	−4612.9	−4497.4
$\ell(r)$	−3501.5	−3497.1	−3496.7	−3501.7	−3503.5
$\hat{\Sigma}_u$	$\begin{bmatrix} 0.29 & 0.20 \\ 0.20 & 0.21 \end{bmatrix}$	$\begin{bmatrix} 0.30 & 0.29 \\ 0.29 & 0.41 \end{bmatrix}$	$\begin{bmatrix} 0.29 & 0.28 \\ 0.28 & 0.34 \end{bmatrix}$	$\begin{bmatrix} 0.29 & 0.24 \\ 0.24 & 0.26 \end{bmatrix}$	$\begin{bmatrix} 0.29 & 0.24 \\ 0.24 & 0.25 \end{bmatrix}$

The sampling period is 02 January 2006 to 31 May 2011. Results are for the underlined model specification in Table 4. REG denotes the Realized EGARCH in (1) with the choice of p and q in parenthesis. REG(1,1) uses RK and $RV^{(1\,tick)}$ as the realized measures of volatility, REG(1,1)* uses RK and $RV^{(1800\,sec)}$, REG(1,1)$^{\bullet}$ uses RK and $RV^{(1800\,sec)*}$, REG(1,1)† uses RK and $RV^{(300\,sec)}$, and REG(1,1)‡ uses RK and $RV^{(300\,sec)*}$. For all models the first estimates are for RK. Robust standard errors in parenthesis below the estimates (see Table 4).

5.2 Diagnostics

To assess the in-sample fit of the model we computed some standard model diagnostics. We only report results for the REG$(1,1)$ model as these are representative for all the realized GARCH type models we estimated. First, to check the model assumption of normality we present quantile-quantile plots in Figure 4, which compare the empirical distribution of the standardized residuals, $\hat{z}_t$ and $\hat{u}_t$, to a normal distribution. The normality of $\hat{z}_t$ is clearly rejected with excess kurtosis far beyond 0. The results suggest that for forecasting one may consider a bootstrap approach (see Section 6).

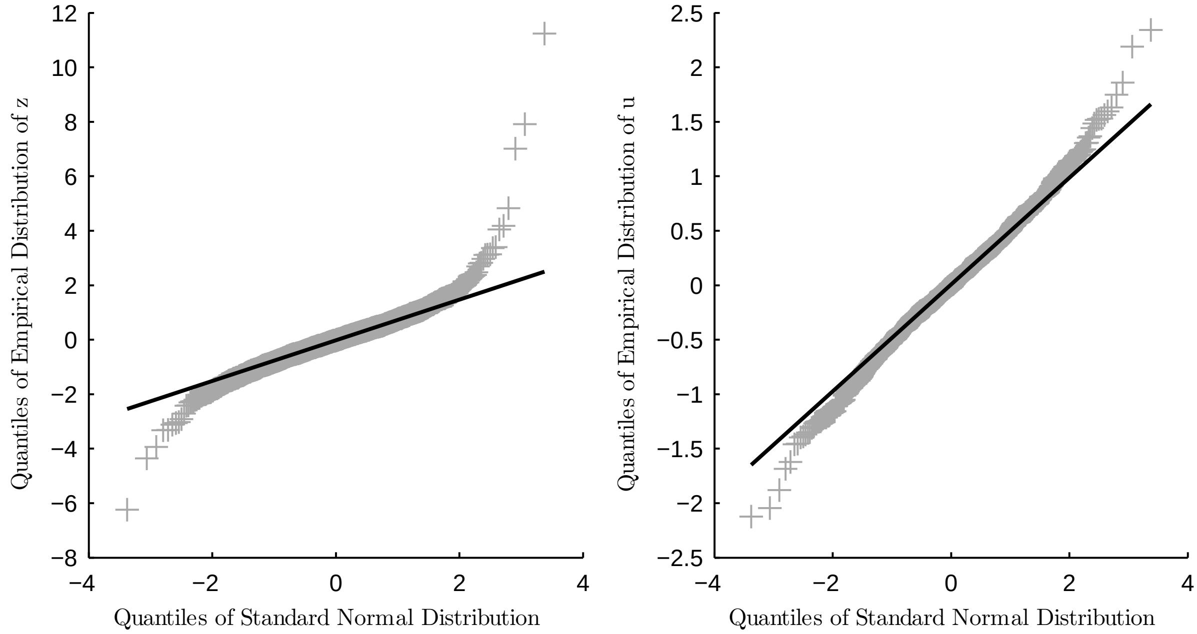

Figure 4: Residual QQ-plots for the REG$(1,1)$ fitted to base load forwards traded on NOMXC. The sampling period is 02 January 2006 to 31 May 2011. Results are for the standardized residuals from the REG$(1,1)$ for quarterly first nearby utilizing the realized kernel as the realized measure of volatility, and using the underlined model specification in Table 4. Jarque-Bera test probabilities are insignificant from zero in both cases.

Table 6 presents results on autocorrelation and heteroskedasticity in the standardized residuals. We do not detect signs of heteroscedasticity, and hence, the model is successful in this sense. With respect to autocorrelation, the results point towards the inclusion of an autoregressive term in the mean equation.

Table 6: Residual diagnostics for the REG$(1,1)$ fitted to base load forwards traded on NOMXC.

	ACF(1)	ACF(2)	ACF(3)	ACF(4)	ACF(5)	$Q(5)$	$Q(10)$	$Q(15)$
$\hat{z}_t$	0.06**	−0.01	0.04	−0.01	0.00	7.29	11.68	14.95
$\hat{z}_t^2$	−0.00	−0.00	0.02	0.01	0.01	0.89	1.61	3.56

The sampling period is 02 January 2006 to 31 May 2011. Results are for the standardized residuals from the REG$(1,1)$ for the quarterly first nearby utilizing the realized kernel as the realized measure of volatility, and using the underlined model specification in Table 4. ACF(L) is the empirical autocorrelation function at lag L, and $Q(L)$ is the Ljung-Box test of no autocorrelation in up to L lags. Two asterisks indicate rejection at 1 percent significance level (for ACF() that the estimate is outside the 95 pct. confidence band) and one asterisk rejection at the 5 percent level.

6 Forecasting

The Realized GARCH framework presented in Section 4 allows us to construct multi-step predictions of volatilities and return density forecasts. We discuss this for the Realized EGARCH(1,1) model with one measurement equation. Generalizations are straightforward. Point forecasts are easy to obtain owing to the fact that the vector $\tilde{h}_{t+1}$ can be represented as an ARMA(1,1) system. Substituting the measurement equation into the equation for the corresponding conditional moment one obtains

$$\tilde{h}_{t+1} = C + A\tilde{h}_t + \boldsymbol{B}\boldsymbol{\epsilon}_t,$$

where $\boldsymbol{\epsilon}_t = \left[\begin{array}{ccc} \delta(z_t) & \tau(z_t) & u_t \end{array}\right]'$, and C, A and $\boldsymbol{B}$ are given by

$$C = \ \omega + \gamma\xi, \qquad A = \beta + \gamma\varphi, \qquad \boldsymbol{B} = \left[\begin{array}{ccc} \gamma & 1 & \gamma \end{array}\right],$$

which follow from the estimation up until time t. The innovation process, $\boldsymbol{\epsilon}_t$, is a martingale difference sequence, and it follows that

$$\mathrm{E}(\tilde{h}_{t+k}|\tilde{h}_t) = A^k\tilde{h}_t + \sum_{j=0}^{k-1} A^j C,$$

which can be used to produce a k-step ahead forecast of $\tilde{h}_{t+k}$. We note that non-linearity implies $\mathrm{E}\left[\exp(\tilde{h}_t)\right] \neq \exp\mathrm{E}\left[\tilde{h}_t\right]$. Forecasts of the conditional distribution of $\tilde{h}_{t+k}|\mathcal{F}_t$, which can be used to deduce unbiased forecasts of the non-transformed variables, e.g., $h_t = \exp(\tilde{h}_t)$, can be obtained by simulation methods or the bootstrap. In the simulation approach, we first generate

$$\boldsymbol{\zeta}_t = \left(\begin{array}{c} z_t \\ u_t \end{array}\right) \sim N\left(\boldsymbol{0}, \left[\begin{array}{cc} 1 & 0 \\ 0 & \sigma_u^2 \end{array}\right]\right), \qquad t = 1, \ldots T,$$

and $\boldsymbol{\epsilon}_t$ can be computed. Based on S simulations we estimate h_{t+k} as $\frac{1}{S}\sum_{s=1}^{S}\exp(\tilde{h}_{t+k})$. Alternatively, a bootstrap approach can be preferable if the Gaussian assumption concerning the distribution of $\boldsymbol{\zeta}_t$ is questionable. From the estimated model we obtain residuals, $(\hat{\boldsymbol{\zeta}}_1, \ldots, \hat{\boldsymbol{\zeta}}_T)$, from which we draw re-samples instead of sampling from the Gaussian distribution. The time series for $\tilde{h}_t$ can now be generated from the bootstrapped residuals $\{\hat{\boldsymbol{\zeta}}_t^*\}$ in the same manner as with the simulated $\{\boldsymbol{\zeta}_t\}$.

In Figure 5 we present point forecasts and results for 1-step-ahead predictions, and 5-step and 20-step ahead predictions with bootstrapped innovations.[20] We set $S = 1000$ and use the REG$(1,1)$ with no deterministic components and with the realized kernel as the realized mea-

[20]Results based on simulated innovations can be found in the web appendix. However, differences to the ones presented here are not visible.

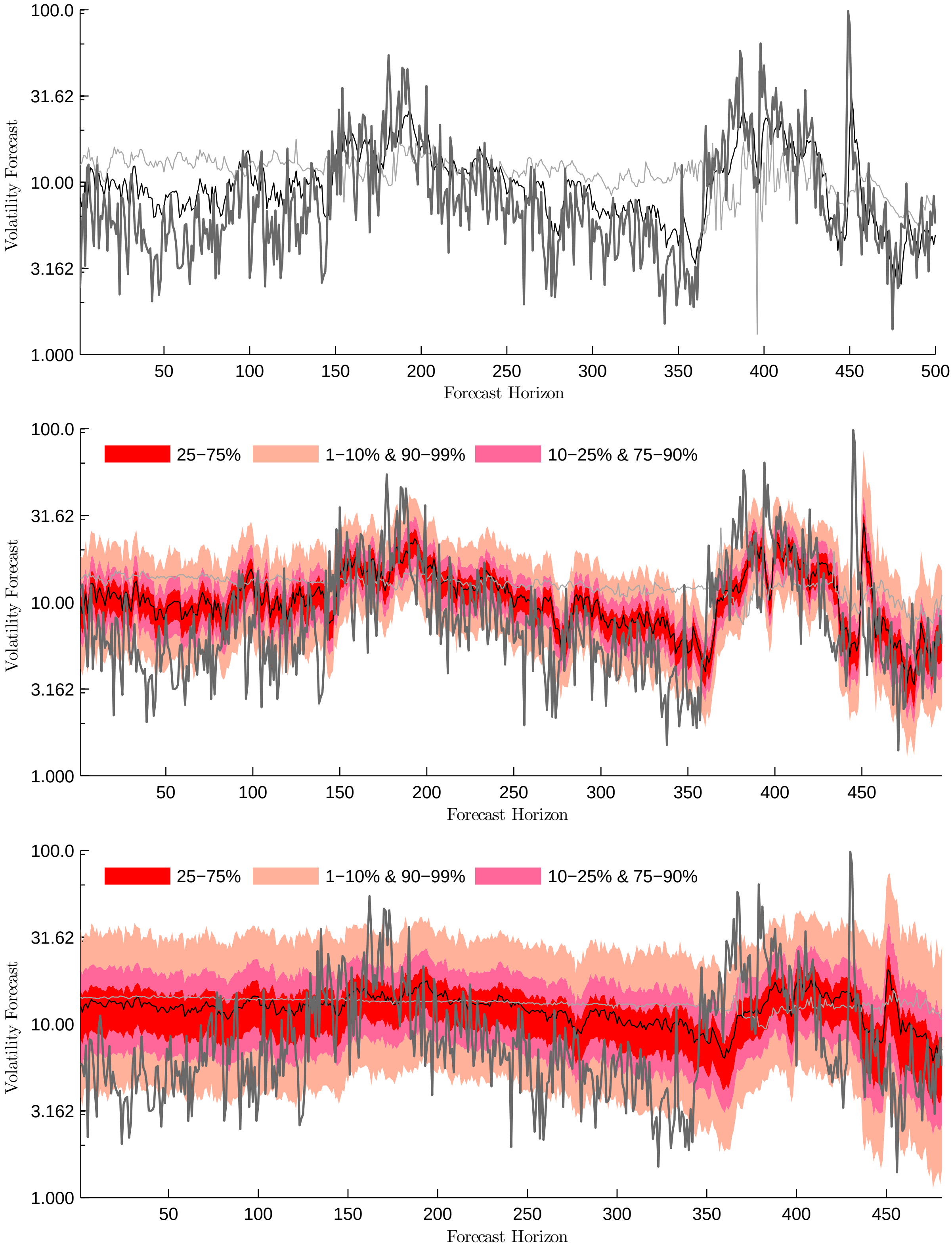

Figure 5: k-step ahead volatility forecasts (bootstrapped innovations) for base load forwards traded on NOMXC (top row has $k = 1$, the middle row has $k = 5$, and the bottom row has $k = 20$). The sampling period is 02 January 2006 to 31 May 2011. Results are for the REG(1,1) and EG(1,1) using the underlined model specification in Table 4, and for the first nearby constructed using rollover schemes as specified in Section 3.2. The dark gray curve is the scaled realized kernel used as benchmark, the black curve is the REG(1,1) k-step ahead volatility forecasts, and the dim gray curve is the EG(1,1). The shaded areas in the middle and bottom rows display percentiles of the bootstrapped REG(1,1) as stated.

sure of volatility. The horizon is set to 500 leaving 858 observations for each estimation in a rolling-window setup. The thick dim gray line is the 15 minute realized variance subsampled every minute, $RV^{(900\,sec)\star}$, scaled by a factor of 3.2 $(= 24/7.5)$ used as benchmark, the black line is constructed from forecasts from the REG$(1,1)$, and the gray line is constructed from forecasts from the EG$(1,1)$. For the 5-step and 20-step ahead predictions, shaded areas represent the percentiles of the S simulations.

For the 1-step-ahead predictions, in the top panel of Figure 5, improvements are visible even to the naked eye, and similar results appear in the middle and bottom panel of the figure. Improvements in the sense that the REG$(1,1)$ is, as expected, more responsive than the EG$(1,1)$, which is inadequately smooth. The REG$(1,1)$ does preserve the smoothing characteristic of GARCH models, which is more pronounced for longer forecast horizons, but it is more capable of picking up abrupt changes in the underlying volatility than the EG$(1,1)$. For 20-step ahead predictions the EG$(1,1)$ results are basically constant whereas the REG$(1,1)$ does conserve some dynamics. Finally, notice the widening of the shaded areas as the forecasting horizon is increased and, in particular, the relatively large percentiles. This shows that our choice of S is, as expected, too conservative in an application in practice.

Table 7: MCS results for base load forwards traded on NOMXC.

Model (innovations)	$s = 1$ (1-step) *Loss*	p_{MCS}	$s = 5$ (5-step) *Loss*	p_{MCS}	$s = 20$ (20-step) *Loss*	p_{MCS}
$Loss = \Sigma\left[MSE\left(RV^{(900sec)\star}, h_{t+s}\right)\right] \cdot 10^{-5}$						
REG (sim.)	0.661	1.0000∗	0.876	0.6037∗	1.041	1.0000∗
REG (boot.)	0.661	1.0000∗	0.875	1.0000∗	1.047	0.1036∗
EG (sim.)	1.121	0.0000	1.145	0.0000	1.160	0.0030
EG (boot.)	1.121	0.0000	1.134	0.0000	1.186	0.0029
$Loss = \Sigma\left[\log h_{t+s} + z_{t+s}^2\right] \cdot 10^{-3}$						
REG (sim.)	1.497	1.0000∗	1.510	1.0000∗	1.566	0.7115∗
REG (boot.)	1.497	1.0000∗	1.511	0.1032∗	1.565	1.0000∗
EG (sim.)	1.590	0.0001	1.576	0.0000	1.597	0.0121
EG (boot.)	1.590	0.0001	1.581	0.0000	1.600	0.0004

Loss and MCS p-values for the different forecasts using two different loss functions. The forecasts in $\hat{\mathcal{M}}^{*}_{90\%}$ are identified by an asterisk. Results are for the underlined model specification in Table 4. EG denotes the conventional EGARCH in (3) and REG denotes the Realized EGARCH in (1).

To more formally compare the forecast performance of the models we apply the MCS of Hansen et al. (2011) in Table 7. The objective of the MCS procedure is to determine the set of "best" models, $\mathcal{M}^{\star}$, from a collection of models $\mathcal{M}_0$, where "best" here is defined by a loss function. Thus, one can view the MCS as a set of models that includes the "best" models with a given level of confidence. With informative data the MCS will consist only of the best model, and less informative data may result in a MCS with several models. We refer to Hansen et al. (2011) for details. In Table 7, $\mathcal{M}_0$ consists of two models in the case of 1-step ahead predictions as no simulations are needed, and four models for 5- and 20-step ahead predictions as we want to compare results using simulated and bootstrapped innovations, respectively. We present results for two different loss functions. The mean squared error (MSE), where the "true value" is taken to be $RV^{(900\,sec)\star}$, and the predicted value of the (partial) likelihood, which the estimation seeks to minimize.[21] The total loss is the sum of the loss function over the forecasting horizon. One asterisk indicates that the model belongs to $\hat{\mathcal{M}}^{\star}_{90\%}$. Overall, the REG$(1,1)$ clearly beats the EG$(1,1)$. For all choices of s, the REG$(1,1)$ is consistently in $\hat{\mathcal{M}}^{\star}_{90\%}$ and EG$(1,1)$ is not. For 1-step ahead predictions $p_{MCS} \simeq 1$ for both loss functions. For 5- and 20-step ahead predictions the MCS does not favour either the REG$(1,1)$ with bootstrapped or simulated Gaussian innovations. They are both consistently in $\hat{\mathcal{M}}^{\star}_{90\%}$.

7 Conclusion

In this paper we have explored the transaction records from NOMXC back to January 2006 and discussed the issue of construction of continual nearby contracts. We have estimated a range of realized measures of volatility, and we used these to enrich the information set of GARCH models in the Realized GARCH framework of Hansen et al. (2012*a*), Hansen and Huang (2012) and Hansen et al. (2012*b*). Estimations clearly reveal a gain from utilizing data at higher frequencies. Compared to ordinary EGARCH models, which is nested in the Realized EGARCH considered, improved empirical fit is obtained, in-sample as well as out-of-sample. The out-of-sample assessment is based on 1-, 5-, and 20-step-ahead regular and bootstrapped rolling-window forecasts. The Realized EGARCH outperforms its nested benchmark visually and in terms of the MCS of Hansen et al. (2011) using two different loss functions. Throughout, the obtained results illustrate the importance of careful volatility estimation as the level is time-dependent but predictable. An appealing extension is in the multivariate setting, where covolatilities between the different NOMXC forwards, and between the forwards and other energy markets such as coal, gas, and oil, are important to many (e.g. energy and utility companies). This could for example extent the daily analysis in Bauwens et al. (2012) to incorporate intraday data. Another interesting venue would be to adapt the Hansen et al. (2012*b*) approach to the energy markets.

[21] We do not intend to forecast r_{t+s} to calculate z_{t+s} but use the true values known ex-post. That is, our focus is fixed on differences in the forecasts of the conditional variance, h_{t+s}. Furthermore, in calculating the predicted value of the partial likelihood for $s > 1$ we do not include the preceding steps. For each forecast we only include the value owing to the specified value of s. One may argue that this only reduce our ability to show that frequency matters as we approach the long-run variance for large s.

References

Andersen, T. G., Bollerslev, T. and Diebold, F. X. (2007), 'Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility', *Review of Economics & Statistics* **89**, 707–720.

Andersen, T. G., Bollerslev, T. and Diebold, F. X. (2009), Parametric and nonparametric volatility measurement, *in* L. P. Hansen and Y. Ait-Sahalia, eds, 'Handbook of Financial Econometrics', North-Holland, Amsterdam.

Back, K. (1991), 'Asset pricing for general processes', *Journal of Mathematical Economics* **20**(4), 371 – 395.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2011), 'Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading', *Journal of Econometrics* **162**, 149–169.

Barndorff-Nielsen, O. E. and Shephard, N. (2002), 'Econometric analysis of realised volatility and its use in estimating stochastic volatility models', *Journal of the Royal Statistical Society* **B 64**, 253–280.

Barndorff-Nielsen, O. E. and Shephard, N. (2007), Variation, jumps and high frequency data in financial econometrics, *in* R. Blundell, T. Persson and W. K. Newey, eds, 'Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress', Econometric Society Monographs, Cambridge University Press, pp. 328–372.

Bauwens, L., Hafner, C. M. and Pierret, D. (2012), 'Multivariate volatility modeling of electricity futures', *Journal of Applied Econometrics, forthcoming* .

Benth, F. E., Álvaro Cartea and Kiesel, R. (2008*a*), 'Pricing forward contracts in power markets by the certainty equivalence principle: Explaining the sign of the market risk premium', *Journal of Banking & Finance* **32**(10), 2006 – 2021.

Benth, F. E., Benth, J. S. and Koekebakker, S. (2008*b*), *Stochastic Modelling of Electricity and Related Markets*, World Scientific Publishing, Toh Tuck Link, Singapore.

Benth, F. E. and Koekebakker, S. (2008), 'Stochastic modeling of financial electricity contracts', *Energy Economics* **30**(3), 1116 – 1157.

Corsi, F. (2009), 'A simple approximate long-memory model of realized volatility', *Journal of Financial Econometrics* .

Deaton, A. and Laroque, G. (1992), 'On the behavior of commodity prices', *Rev. Financ. Econ. Stud.* **59**, 1–23.

Engle, R. F. (2002), 'New frontiers for ARCH models', *Journal of Applied Econometrics* **17**, 425–446.

Engle, R. F. and Gallo, G. (2006), 'A multiple indicators model for volatility using intra-daily data', *Journal of Econometrics* **131**, 3–27.

Escribano, A., Peña, J. I. and Villaplana, P. (2011), 'Modelling electricity prices: International evidence', *Oxford Bulletin of Economics and Statistics* **73**, 622–650.

Eydeland, A. and Wolyniec, K. (2003), *Energy and power risk management: new developments in modeling, pricing, and hedging*, John Wiley and Sons.

Geman, H. and Roncoroni, A. (2006), 'Understanding the fine structure of electricity prices', *Journal of Business* **79**, 1225–1261.

Hansen, P. R. and Huang, Z. (2012), 'Exponential garch modeling with realized measures of volatility', *Working Paper, CREATES* .

Hansen, P. R., Huang, Z. and Shek, H. H. (2012*a*), 'Realized garch: A joint model of returns and realized measures of volatility', *Journal of Applied Econometrics* **27**, 877–906.

Hansen, P. R. and Lunde, A. (2006), 'Realized variance and market microstructure noise', *Journal of Business and Economic Statistics* **24**, 127–218. The 2005 Invited Address with Comments and Rejoinder.

Hansen, P. R. and Lunde, A. (2011), Forecasting volatility using high-frequency data, *in* M. Clements and D. Hendry, eds, 'The Oxford Handbook of Economic Forecasting', Oxford: Blackwell, chapter 19, pp. 525–556.

Hansen, P. R., Lunde, A. and Nason, J. M. (2011), 'The model confidence set', *Econometrica* **79**, 456–497.

Hansen, P. R., Lunde, A. and Voev, V. (2012*b*), 'Realized beta garch: A multivariate garch model with realized measures of volatility and covolatility', *Working Paper, CREATES* .

Haugom, E., Westgaard, S., Solibakke, P. B. and Lien, G. (2011*a*), 'Modelling day-ahead nord pool forward-price volatility: Realized volatility versus garch models', *International Research Journal of Finance and Economics* .

Haugom, E., Westgaard, S., Solibakke, P. B. and Lien, G. (2011*b*), 'Realized volatility and the influence of market measures on predictability: Analysis of Nord Pool forward electricity data', *Energy Economics* **33**, 1206 – 1215.

Higgs, H. and Worthington, A. (2008), 'Stochastic price modeling of high volatility, mean-reverting, spike-prone commodities: The australian wholesale spot electricity market', *Energy Economics* **30**(6), 3172 – 3185. Technological Change and the Environment.

Holton, G. A. (2003), *Value-at-Risk, Theory and Practice*, Academic Press, London.

Ma, C. K., Mercer, J. M. and Walker, M. A. (1992), 'Rolling over futures contracts: A note', *Journal of Futures Markets* **12**(2), 203–217.

Malo, P. and Kanto, A. (2006), 'Evaluating multivariate garch models in the nordic electricity markets', *Communications in Statistics - Simulation and Computation* **35**(1), 117–148.

NASDAQ OMX Commodities (2011*a*), *Productcalendar*.

NASDAQ OMX Commodities (2011*b*), *Trade at NASDAQ OMX Commodities Europe's Financial Market*.

NASDAQ OMX Commodities (2011*c*), *Trading and Clearing Financial Power Contracts*.

Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: A new approach', *Econometrica* **59**(2), 347–370.

Pen, Y. L. and Sévi, B. (2010), 'Volatility transmission and volatility impulse response functions in european electricity forward markets', *Energy Economics* **32**(4), 758 – 770. Policymaking Benefits and Limitations from Using Financial Methods and Modelling in Electricity Markets.

Protter, P. (2005), *Stochastic Integration and Differential Equations*, New York: Springer-Verlag.

Shephard, N. and Sheppard, K. K. (2010), 'Realising the future: forecasting with high-frequency-based volatility (HEAVY) models', *Journal of Applied Econometrics* **25**, 197–231.

Veraart, A. E. D. and Veraart, L. A. M. (2013), Modelling electricity day-ahead prices by multivariate levy semistationary processes, *in* Benth, Kholodnyi and Laurence, eds, 'Forthcoming in Quantitative Energy Finance', Springer, Berlin.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y. (2005), 'A tale of two time scales: Determining integrated volatility with noisy high frequency data', *Journal of the American Statistical Association* **100**, 1394–1411.

Appendix A

Information Contained in Data Files

Date The date of registration of a trade on the form 'yyyy-mm-dd hh:mm:ss' (time indication is superfluous, and 00:00:00 in all lines, and thus removed)

ContractTime The time of trade on the form 'yyyy-mm-dd hh:mm:ss'. For trades on exchange *Date* and *ContractTime* are in agreement. For over-the-counter trades, dates in *ContractTimes* are often earlier than *Date,* and time indications are imprecise - most appear to be a rough estimate, e.g. '12:00:00'.

DealNumber A unique identification number for all deals entered. Both buy-side and sell-side ticks are included in the data set with the same *DealNumber*. That is, a deal in CLICK Trade may consist of 2 or more trades, i.e. there is not a one-to-one relation between buy and sell, as one offer may hit several bids if price and volume match. Two for low volumes and only one counterpart. Three for larger volumes and two or more counterparts.

TradeNumber A numbering of each tick. Not used.

DealSource The origin of the trade. On exchange ("Exchange") or over-the-counter ("OTC").

BuySell A buyer ("B") or seller ("S") indication. In the CLICK Trade system a "B" ("S") is used for the deal if the buyer (seller) is the so-called "Aggressor" accepting the lowest ask (bid) quote(s). Unfortunately, trades are not saved using this definition, but instead using "B" for the buyside(s) and "S" for the sellside(s).

Market Id A market identification. All ticks in the raw data set have the same prefix, "ENO", which is an abbrevation of the underlying commodity (Electricity) and market (NOrdic). Not used.

MainCategory A product categorization into base load futures/forwards with a delivery period longer than one day ("Base"), base load futures with a delivery period of one day ("BaseDay"), forward contracts-for-difference ("CfD"), ("EUA"), European options ("Options").

Category For *MainCategory* "EUA" either "EUA Forward" or "EUASPOT and for all other *MainCategories* a delivery period specification; "Day", "Week", "Month", "Quarter" or "Year".

ContractTicker The ticker of the traded contract.

DealPrice The transaction price in EUR/MWh.

NbrOfContracts The traded number of contracts.

ContractSize The number of delivery hours in the traded contract.

Volume A multiple of NbrOfContracts and ContractSize.

Earlier versions of the files may use different headers and contain slightly different information (e.g. the column InstrumentTypeID containing the type of the traded product, i.e. forward (ENFW), future week (ENFU), future day (EDAY), forward CfD (ENCD), and European option (ENOC/ENOP), is no longer included).

Research Papers 2013

CREATES
Center for Research in Econometric Analysis of Time Series