Covariance estimation using high-frequency data: An analysis of Nord Pool electricity forward data

Gudbrand Lien

Faculty of Economics and Organization Science, Lillehammer University College, NO-2624 Lillehammer, Norway; email: gudbrand.lien@hil.no

Erik Haugom

Faculty of Economics and Organization Science, Lillehammer University College, NO-2624 Lillehammer, Norway; email: erik.haugom@hil.no

Sjur Westgaard

Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway; email: sjur.westgaard@iot.ntnu.no

Per Bjarte Solibakke

Molde University College, Postbox 2110,NO- 6402 Molde, Norway; email: per.b.solibakke@hiMolde.no

Abstract

The modeling of volatility and correlation is important in order to calculate hedge ratios, value at risk estimates, CAPM betas, derivate pricing and risk management in general. Recent access to intra-daily high-frequency data for two of the most liquid contracts at the Nord Pool exchange has made it possible to apply new and promising methods for analyzing volatility and correlation. We apply the concepts of realized volatility and realized correlation, and this first study statistically describes the distribution (both distributional properties and temporal dependencies) of electricity forward data from 2005 to 2009. The main findings show that the logarithmic realized volatility is approximately normally distributed, while realized correlation seems not to be. Further, realized volatility and realized correlation have a long-memory feature. There also seems to be a high correlation between realized correlation and volatilities and positive relations between trading volume and realized volatility and between trading volume and realized correlation. These results are to a large extent consistent with earlier studies of stylized facts of other financial and commodity markets.

Index words: Realized volatility and correlation, High-frequency data, Distribution properties, Temporal dependence, Nord Pool forward data

Acknowledgements: This paper is a part of the project "ELCARBONRISK - Modeling and forecasting risk in electricity, carbon and related energy markets (Oil, Gas, Coal)," financed by Research Council of Norway, Eidsiva Energi and Tafjord Kraftproduksjon AS. The authors wish to thanks the participants of the 2nd Trondheim Winter Energy Workshop 4-5 March 2010 for their comments.

1 INTRODUCTION

Balance risk and expected returns within a portfolio approach constitute some of the key concepts in modern finance. Accordingly, the estimation and forecasting of both volatility and correlation is arguably among the most important pursuits in empirical asset pricing, asset allocation and risk management. Examples of the crucial role that volatility and correlation estimates and applications play in finance include the calculation of hedge ratios, the calculation of portfolio value at risk estimates, the calculation of CAPM betas, option and derivate pricing, and volatility transmission between assets and markets.

Many classical models in financial economics assume constant volatilities and correlations, even if the dynamic properties of volatilities and correlations have been widely accepted. The research on dynamic properties of volatilities and correlation has typically been based on the estimation of parametric univariate and multivariate ARCH (and its generalizations) and stochastic volatility models. One drawback with these models (and parametric models in general) is that they depend on specific distributional assumptions, which reduce the robustness of the empirical findings (Andersen *et al* 2006).

The recognition of the limitations of the traditional volatility and correlation models has led to a different approach. The development of computer technology over the past decades and the increased availability of high-frequency financial data have opened up a new field of research within this framework (Dacorogna *et al* 2001). The idea has been to use historically relevant and reliable high-frequency data in order to improve the modeling and forecasting of outcome variability and correlation. The availability of high-frequency data has made it possible to construct ex post facto realized daily volatilities and correlations, through the summation of squares and the cross-products of intraday observations, respectively. This approach has the advantage of allowing for the characterization of the distributional features of the volatilities and correlation without an attempt to fit multivariate conditional and stochastic volatility models.

Andersen and Bollerslev's (1998) seminal paper shows that realized volatility computed from high-frequency intraday returns is effectively a model-free volatility measure. On the basis of mainly the same ideas and procedures as for univariate realized volatility, both Andersen *et al* (2001a) and Barndorff-Nielsen and Shephard (2004a) have spelled out the concept of realized covariances and correlations. This framework permits one to treat volatility and correlation simply as observables (and not latent or modeled), which has opened up several opportunities. We can, for example, rely on conventional statistical techniques for characterizing ex post facto realized volatilities and correlations measures' distributional properties. Further, we can use these observed measures in dynamic forecasting with simple standard "regression" techniques.

Some studies have also given some indication that nonparametric models based on high-frequency data provide superior out-of-sample forecasts of volatility compared to parametric volatility models based on daily data (e.g. Andersen *et al* 2003; Racicot *et al* 2008).

Until about the late 1980s only sparse data were available, and the research consisted of formulating adequate models and verifying that the models reproduced the data patterns and gave reasonable predictions for the future, with an emphasis on the methodology. In a situation with sparse data, it was important to garner as much as possible information from what little there was, and so it was natural for the researcher to take great pains to make sure that the methodology was correct. Of course, it is equally important for studies to employ the correct methodology and to implement certain analytical improvements. However, owing to the availability of high-frequency data, the research community has begun to focus on indepth, assumption-free (or with a minimum set of assumptions) analyses in order to discover the fundamental statistical properties of the data. These studies aim to document the statistical characteristics of the financial returns or the “stylized facts”, as it often is named in the econometric and finance literature, where the understanding of the market is of interest in itself and can give useful information for the specification and estimation of predictive models. Notable examples of studies within the financial field of stylized facts are those of individual stocks and stock indices (Andersen *et al* 2001b; Areal and Taylor, 2002; Thomakos and Wang 2003; Chen *et al* 2006), bonds (Thomakos and Wang 2003), currencies (Andersen *et al* 2001a; Dacorogna *et al* 2001; Thomakos and Wang 2003; Chen *et al* 2006) and agricultural commodities (Chen *et al* 2006). Some of these studies, in addition to variance and volatility analyses, also focus on the stylized facts of realized covariances and correlations between assets investigated (Andersen *et al* 2001a; Andersen *et al* 2001b; Thomakos and Wang 2003). To our knowledge, Wang *et al* (2008) have been the first to examine the distributions of realized energy-futures volatilities and correlation in their study of the NYMEX light crude oil and natural-gas futures contracts.

Ulrich (2009) and Chan *et al* (2008) have analyzed high-frequency US and Australian electricity spot-market data. Haugom *et al* (2009) have studied univariate realized volatility on the basis of high-frequency electricity-futures data from the Nord Pool market. However, to our knowledge, nobody has analyzed realized covariance based on high-frequency financial electricity data.

The main findings obtained in these studies mentioned above show that the logarithmic realized volatility and correlation are approximately normally distributed. Further, realized volatility and correlation have a long-memory feature, which can be modeled by fractionally integrated processes, and there seems to be a high correlation between realized correlation and volatilities.

Recently, high-frequency, tick-by-tick data have also become available for the two most liquid electricity forward contracts (yearly and quarterly contracts) at the central Nord Pool data source. This is a quickly growing derivative market, and more knowledge and tools for risk management and trading applications are needed. In this study, we build on the framework employed in Andersen *et al* (2001a) and Andersen *et al* (2001b) on stylized volatility and correlation facts, and we apply high-frequency data in the electricity market to analyze the stylized volatility and correlation facts of these electricity forward contracts. As far as we are aware, we are the first to undertake this endeavor. As electricity is distinct in several important ways from other commodities (e.g., non-storability, uncertainty in load and generation, inelastic demand, oligopolistic generation), it is interesting to compare whether the use of high-frequency data reveals that the behavior of the financial electricity market differs significantly from traditional financial markets. The data covers the period of June 2005 to May 2009.

The next section includes a brief review of the concepts of realized volatility and correlation, and a description of the data set follows. The next section presents the main results of the analysis, and the last section offers some concluding comments.

2 THE CONCEPTS OF REALIZED VOLATILITY AND CORRELATION

Barndorff-Nielsen and Shephard (2001) and Andersen *et al* (2002) argue for the importance of decomposing the volatility of returns over short horizons in both a continuous sample path component and an occasional discontinuous jump component. Since one of our main focuses is covariance and correlation between series, we consider a N-dimensional log-price process $p(s)$ over the period $[t-1,t]$. We assume that the N-dimensional or multivariate log-price process is governed by a jump-diffusion process, which can be formulated as follows:

$$dp(s) = \mu(s)ds + \Omega(s)dw(s) + K(s)dq(s) \quad (1)$$

where the drift, $\mu(s)$, is a N-dimensional vector process, the instantaneous volatility, $\Omega(s), is\ a\ N \times N$ matrix such that $\sum(\mathrm{s}) = \Omega(s)\Omega^{'}(s)$ is the covariance matrix process of the continuous sample path component and $W(t)$ is a vector of N independent Brownian motions. Further, K is the $N \times N$ process controlling for the magnitude and transmission of jumps, and $q(s)$ is the N-dimensional jump-counting process such that $K(s)dq(s)$ is the contribution of the jump process to the log-price process.

Here we slightly simplify the notation by denoting κ_j as the contribution of the *j*-th jump in $[t-1, t]$ to the continuous price diffusion. Thus, (1) can be written as:

$$dp(s) = \mu(s)ds + \Omega(s)dw(s) + \kappa(s) \quad (2)$$

We assume that the value of this process is observed in equally spaced intervals, Δ, in the period $[t-1, t]$. We observe the log price every Δ units of time, where Δ is small, and we set $M \equiv [1/\Delta]$, as equally spaced intraday returns. Then, the *i*-th intraday return of day *t* is:

$$r_{t,i} = p(t + i\Delta) - p(t + (i-1)\Delta), \qquad i = 1, \cdots, M. \quad (3)$$

The realized quadratic covariance, $(RCov)$, of the log-price process for day t is

$$RCov_t = \sum_{i=1}^{M} r_{t,i}\, r_{t,i}^{'} \quad (4)$$

and thus realized variance, $(RVar)$, on day *t* is

$$RVar_t = \sum_{i=1}^{M} r_{t,\mathrm{i}}^{2} \quad (5)$$

as a special case of (4). As shown by Andersen *et al* (2001a), when $M \to \infty$, a consistent estimator over $[t-1, t]$ of the quadratic covariance in (1), $RCov_t$, is:

$$RCov_t = \int_{t-1}^{t} \Sigma(s)ds + \sum_{j=1}^{j_t} \kappa_j \kappa_j^{'} \quad (6)$$

where $j_t = \int_{t-1}^{t} dq^*(s)$, and where $q^*(s)$ represents the univariate counting process derived from $q(s)$ such that $q^*(s)$ increases by 1 whenever $q(s)$ changes. The continuous or integrated covariance matrix ($ICov$) over $[t-1, t]$ is the matrix

$$ICov_t = \int_{t-1}^{t} \Sigma(s)ds. \quad (7)$$

This means that if we consider price processes without jumps or co-jumps, the last component in (1) and (2) drops out, and we only need a consistent estimator for the integrated covariance matrix, which is simply (4). The non-parametric realized covariance estimator in (4) will, following (6), include all jumps and co-jumps, as well as continuous price-diffusion processes. In other words, the realized covariance represents an ex post facto measure of the true total price variation, including discontinuous jump and co-jump parts. This occurs because the realized covariance is the sum of squared high-frequency returns, and so any abnormally large positive or negative returns will be squared and thus have a large impact on the realized covariance measure. If we have an additional estimator for the integrated covariance matrix ($ICov$), that are robust to jumps/co-jumps, we can disentangle the continuous and jump/co-jump components in $RCov$, by decomposing the estimated $ICov$ from $RCov$. Yet, there still is some way to go. While in the univariate case the difference between $RVar$ and realized bipower variance ($RBPVar$) is an estimate of jumps, the difference is not zero even in absence of jumps. But, there exist a formal test to remove the non-significant differences in the univariate case. However, based on our knowledge, there does not exist any formal test that can distinguish significant differences between $RCov$ and realized bipower covariance ($RBPCov$) for the multivariate case. In other words, for the multivariate case we cannot know if the difference is significant and we cannot know if the difference is caused by jumps in one of the assets or by co-jumps jointly among a number of assets.

Still, studies of univariate series have shown that the jump component has both different distributional properties and temporal dependency properties other than the diffusion-volatility component (e.g. Huang and Tauchen, 2005; Andersen *et al* 2007). These studies show that the volatility jump component is both highly important and distinctly less persistent than the continuous component, and that a separation of the rough jump moves from the smooth continuous moves results in significant out-of-sample volatility forecast improvements. Therefore, for forecasting purposes, even without focusing on the jump/co-jump element, there may be reason to estimate a smooth or continues as possible realized covariance and correlation measures, that is, to estimate an approximation for $ICov$.

One candidate for cases with jumps/co-jumps for approximation of the integrated covariance matrix is the realized bipower-covariation process, originally proposed by Barndorff-Nielsen and Shephard (2004b). It is a multivariate extension of the univariate realized bipower-variation process. Both the realized bipower-variation estimator and the realized bipower-covariation estimator have some drawbacks (e.g. Boudt *et al* 2008: Andersen *et al* 2009). Given that the high-frequency data is observed over extremely short intervals, there is small chance of jumps/co-jumps affecting two neighboring returns, and so the impacts of jumps/co-jumps on the realized bipower-covariation estimator can be negligible. However, for electricity modeling, as for other assets, the returns are more typically observed over longer intervals, such as 15 or 30 minutes. Then, the realized bipower-covariation estimator is typically highly affected when jumps/co-jumps affect two or more neighboring returns. Further, the bipower measures are sensitive to the presence of "zero" returns in the sample. Several alternatives to the realized bipower-covariation estimator have been proposed, for example, the estimator of nearest-neighbor truncation by Andersen *et al* (2009) and the range-based covariance estimator by Bannough *et al* (2009).

In this study we use the realized outlyingness weighted quadratic covariation, $(ROCov)$, introduced by Boudt *et al* (2008) and Laurent (2009), as an nonparametric estimator for the integrated covariance matrix $(ICov)$. While realized quadratic covariance, $(RCov)$, is estimated by the sum of outer products of high-frequency returns, $ROCov$ equals the *weighted* sum of outer products of high-frequency returns. This continuous covariance estimator downweights returns that are local outliers relative to neighboring returns. Outlyingness arises because of jumps, co-jumps, or other reasons. The more "jumpy" a local return window is, the lower weight it receives on the $ROCov$ estimator. The analogue outlyingness weighted quadratic variance measure is denoted $ROVar$. Although Boudt *et al* (2008) have given a rigorous description of the $ROCov$ and $ROVar$ estimators, here we only briefly sketch the procedure.

The first step involves an estimation of a local multivariate outlyingness measure of the return vector.[1] The outlyingness measure, $d_{t,i}$, for the return vector, $r_{t,i}$, equals:

$$d_{t,i} = r_{t,i}^{'} S^{-1}(r_{t,i}) r_{t,i} \tag{8}$$

where $S(r_{t,i})$ is a robust estimator of the multivariate scale ("standard deviation") of the returns belonging to the same local window as $r_{t,i}$.[2] Our application estimates the multivariate scale $S(r_{t,i})$ with window length *M*, that is, one day (seven observations in our case, see below). The multivariate outlying measure is asymptotically chi-squared distributed (Laurent 2009).

[1] The outlyingness measure $d_{t,i}$ can be computed on raw returns and filtered returns. Filtered returns account for the intraday periodicity, and are applied in this study.

[2] Boudt *et al* (2008) use the Minimum Covariance Determinant (MCD) estimator to construct the multivariate outlyingness.

In the second step a weight function, $w(d_{t,i})$, for all high-frequency returns in the interval $[t-1,t]$, is chosen. Jumps that influence the weight function, $w(d_{t,i})$, need to be statistically significant. If the outlyingness measure, $d_{t,i}$, exceeds a threshold defined by a quantile of chi-square distribution, then it is statistically significant. For returns without jumps, the outlyingness measure $d_{t,i}$ does not raise any suspicion about jumps, and the weight function equals one. If significant jumps occur, the outlyingness measure $d_{t,i}$ will make the weight function less than one or zero, depending on method used for downweighting. Two methods are possible for downweighting those returns that are outliers in relation to the majority of the returns in their local window. If the "hard" rejection is applied, the weight function assumes the value zero when significant outlyingness exists. With the "soft" rejection (which this study uses), the weight function approaches zero as more jumps occur in returns. In other words, the hard-rejection weight function removes the extreme returns in the estimation of $ROCov$, while the soft-rejection weight function downweights these observations in the estimation of $ROCov$.

Having calculated both the outlyingness $d_{t,i}$ and the weight function $w(d_{t,i})$, from the high-frequency data, we can compute the $ROCov$ estimator as follows:

$$ROCov_t = c_w \frac{\sum_{i=1}^{M} w(d_{t,i}) r_{t,i} r_{t,i}^{'}}{\frac{1}{M}\sum_{i=1}^{M} w(d_{t,i})} \qquad (9)$$

where the correction factor c_w ensures that the $ROCov$ is consistent for $ICov$, (see Boudt *et al* (2008, p. 14) for further details).

In this study we look at the distributional properties and temporal dependencies for the following:

- Realized variance, $RVar$, and realized outlyingness weighted quadratic variance, $ROVar$.

- Root of $RVar$ and $ROVar$, namely, realized volatilities measures, henceforth RV and ROV, respectively.
- Logarithm of the root of $RVar$ and $ROVar$, that is, realized logarithmic volatilities measure, denoted $lnRV$ and $\ln ROV$, respectively.
- Realized covariance measure, $RCov$.
- Realized outlyingness weighted quadratic covariance measure, $ROCov$.
- Realized correlation, $RCorr = RCov_{t,s}/(RV_t \times RV_s)$.
- Realized outlyingness weighted correlation, $ROCorr = ROCov_{t,s}/(ROV_t \times ROV_s)$.
- Realized jump/co-jump share between asset t and asset s, defined as $RJ_{ts} = max(RCov_{t,s} - ROCov_{t,s}, 0)/RCov_{t,s}$ (Andersen *et al* 2007).

All estimates in this study were calculated with the OxMetrics package called RE@LIZED (Laurent, 2009).

3 DATA

The central Nord Pool data source includes transactional prices and trading volumes (contracts) in megawatt (MW). The data encompass forward prices for two financial contracts: 1) one-quarter-ahead prices traded the last quarter before maturity; 2) one-year-ahead prices traded the last year before maturity. The financial trading at Nord Pool takes place between 08:00 to 15:30. The series consist of approximately 16000 observations from June 2005 up to June 2009. Both future contract series are organized into 30-minute raw-price data observations, using the before nearby contract price as the interpolated price for the specified half hour. In order to avoid problems with large jumps in returns between contracts, the returns at 08:00 for the first trading day of the new contract are defined as missing for both data-sets. Accordingly, the sample used in this study consists of 995 daily-return data and daily-realized variability measures, for the period 1 June 2005 to 29 May 2009. We have also removed a few observations from the yearly or quarterly contracts that did not match each other in date. Some earlier studies have argued for dropping overnight returns in the volatility and correlation measures because these non-trading overnight hours may differ from the volatility and correlation during trading hours and consequently introduce more noise than useful information (e.g. Martens 2002). The contract-price data used in this study show very small or, more typically, no changes during the night, viz, between 16:00 and 08:00, which further supports dropping these observations. Hence, in this study we apply a sampling frequency of 15 (M=15) 30-minute intraday returns.

Table 1 shows the descriptive statistics of these data. For quarterly contracts the returns varies between -11.83% (8:30 returns – relative price change between 8:00 and 08:30) and 10.61% (8:30 returns). Daily standard deviation of returns varies between 0.37% (12:30 returns) and 1.87% (8:30 returns). For yearly contracts the returns vary between -10.24% (8:30 returns) and 5.23% (8:30 returns). The daily standard deviation of returns varies between 0.30% (12:30 returns) and 1.28% (8:30 returns). These numbers indicate that the shorter contract time series have higher variability than the longer contract over the day. The average is, as expected, close to zero for quarterly and yearly contract returns. All series also have fat tails (high positive kurtosis).

TABLE 1 Descriptive statistics for returns for the quarterly and yearly contracts. The table shows the number of observations, the mean, median, min, max, standard deviation, skewness and kurtosis. Data from 1st June 2005 to 29th May 2009.

Quarterly	Hours	N.Obs	Mean	Median	Min.	Max.	Std.Dev.	Skew.	Kurt.
	8:30	995	0.01 %	0.00 %	-11.83 %	10.61 %	1.87 %	0.06	4.86
	9:00	995	-0.10 %	0.00 %	-5.95 %	2.92 %	0.81 %	-1.11	6.42
	9:30	995	0.04 %	0.00 %	-2.21 %	4.81 %	0.66 %	1.49	9.56
	10:00	995	-0.01 %	0.00 %	-2.36 %	2.17 %	0.48 %	-0.17	2.40
	10:30	995	-0.04 %	0.00 %	-3.02 %	1.94 %	0.47 %	-0.19	3.79
	11:00	995	-0.01 %	0.00 %	-2.82 %	2.85 %	0.46 %	-0.38	6.48
	11:30	995	0.01 %	0.00 %	-2.55 %	1.73 %	0.38 %	-0.52	6.02
	12:00	995	0.01 %	0.00 %	-2.67 %	3.69 %	0.38 %	0.52	13.35
	12:30	995	-0.02 %	0.00 %	-2.42 %	1.72 %	0.37 %	-0.86	7.14
	13:00	995	0.01 %	0.00 %	-3.49 %	3.97 %	0.67 %	-0.06	4.39
	13:30	995	-0.01 %	0.00 %	-4.95 %	2.94 %	0.62 %	-0.49	6.25
	14:00	995	0.00 %	0.00 %	-2.71 %	2.86 %	0.51 %	-0.13	5.34
	14:30	995	-0.02 %	0.00 %	-3.06 %	4.14 %	0.51 %	0.22	8.67
	15:00	995	-0.01 %	0.00 %	-2.61 %	2.14 %	0.48 %	-0.25	3.06
	15:30	995	0.05 %	0.00 %	-2.58 %	3.09 %	0.57 %	0.24	3.17
Yearly	Hours	N.Obs	Mean	Median	Min.	Max.	Std.Dev.	Skew.	Kurt.
	8:30	995	0.03 %	0.00 %	-10.24 %	5.23 %	1.28 %	-0.85	7.26
	9:00	995	-0.05 %	0.00 %	-5.61 %	4.11 %	0.72 %	-0.94	9.67
	9:30	995	0.01 %	0.00 %	-2.51 %	5.81 %	0.53 %	1.37	17.25
	10:00	995	0.00 %	0.00 %	-2.64 %	2.31 %	0.43 %	-0.06	5.88
	10:30	995	-0.01 %	0.00 %	-2.41 %	2.45 %	0.42 %	0.13	5.76
	11:00	995	0.02 %	0.00 %	-3.05 %	2.36 %	0.37 %	-0.09	8.92
	11:30	995	0.02 %	0.00 %	-2.18 %	2.35 %	0.34 %	0.33	7.13
	12:00	995	-0.01 %	0.00 %	-2.67 %	1.95 %	0.301 %	-0.73	13.10
	12:30	995	0.00 %	0.00 %	-2.13 %	1.76 %	0.295 %	-1.05	8.49
	13:00	995	0.00 %	0.00 %	-2.84 %	2.29 %	0.37 %	-0.27	8.26
	13:30	995	0.00 %	0.00 %	-2.67 %	2.94 %	0.41 %	0.36	9.45
	14:00	995	0.00 %	0.00 %	-2.87 %	1.77 %	0.38 %	-0.51	6.34
	14:30	995	0.00 %	0.00 %	-2.34 %	3.49 %	0.42 %	0.56	10.15
	15:00	995	-0.01 %	0.00 %	-2.94 %	2.51 %	0.40 %	-0.24	6.91
	15:30	995	0.02 %	0.00 %	-2.86 %	3.94 %	0.49 %	0.46	7.17

4 RESULTS

4.1. The distribution of daily volatility, correlation and jumps/co-jumps

We start with some stylized facts of distributional properties. The summary statistics in Table 2 show that all realized variances (weighted for outlyingness or not), realized volatilities and realized covariance distributions for both quarterly and yearly contracts are significantly right-skewed and have significantly excessive kurtosis that exceeds the normal value of zero; the result is that the normal distribution is a poor approximation (based on a number of normality tests). The normality tests used are the Jarque-Bera test, the Anderson-Darling test and the Shapiro-Wilk test.[3]

The realized logarithmic volatilities ($lnRV$ and $lnROV$) seem to be approximately normal distributed, especially the variable for realized logarithmic volatilities for quarterly contracts and the realized logarithmic outlyingness weighted volatilities for yearly contracts (see the upper panel the last two columns in Table 2 and the two uppermost rows of panels in Figure 1).

Compared to the corresponding variance measures, the realized correlation ($RCorr$) and realized outlyingness weighted correlation ($ROCorr$) are more symmetric, with less skewness and kurtosis, but non-normally distributed (see the lower panel, columns three and four, in Table 2 and the lowermost rows of panels in Figure 1). We may also note, as expected, the strong positive realized correlation and realized outlyingness weighted correlation, on average, between the yearly and quarterly contracts. However, these correlations also display high variation, ranging from -0.67 to 0.99 for *RCorr* and from -0.99 to 1.00 for $ROCorr$.

[3] For data with long memory, the Jarque-Bera test over-rejects normality and is thus not recommended (Thomakos and Wang, 2003). Owing to this, we use a number of normality tests.

TABLE 2 Summary statistics of realized variance, volatility, covariance, correlation and jump/co-jump distributions.

Variance and volatility	RVar	ROVar	RV	ROV	lnRV	lnROV
Quaterly						
Mean	0.076 %	0.056 %	2.359 %	1.941 %	-3.90	-4.16
Median	0.039 %	0.026 %	1.980 %	1.621 %	-3.92	-4.12
Min.	0.001 %	0.000 %	0.348 %	0.100 %	-5.66	-6.91
Max.	1.987 %	1.788 %	14.096 %	13.373 %	-1.96	-2.01
Std.Dev.	0.119 %	0.103 %	1.429 %	1.349 %	0.56	0.68
Skew.	6.84 **	7.85 **	2.01 **	2.18 **	0.06	-0.34 **
Kurt.	81.72 **	99.80 **	7.66 **	9.24 **	-0.10	0.45 **
JB	284580 **	423121 **	3105 **	4331 **	1.06	27.49 **
AD	+Inf. **	+Inf. **	32.11 **	+Inf. **	0.30	1.31 **
SW	0.513 **	0.451 **	0.851 **	0.842 **	0.999	0.993 **
Yearly						
Mean	0.043 %	0.029 %	1.724 %	1.299 %	-4.24	-4.66
Median	0.018 %	0.009 %	1.359 %	0.947 %	-4.30	-4.66
Min.	0.000 %	0.000 %	0.209 %	0.050 %	-6.17	-7.61
Max.	1.852 %	1.782 %	13.608 %	13.351 %	-1.99	-2.01
Std.Dev.	0.082 %	0.074 %	1.140 %	1.117 %	0.58	0.82
Skew.	12.12 **	14.41 **	2.47 **	2.66 **	0.18 *	-0.20 **
Kurt.	240.95 **	315.58 **	13.95 **	15.82 **	-0.16	0.03
JB	2431220 **	4163367 **	9073 **	11558 **	6.39 *	6.79 *
AD	+Inf. **	+Inf. **	+Inf. **	+Inf. **	2.60 **	0.73
SW	0.390 **	0.323 **	0.816 **	0.791 **	0.994 **	0.995 *
Covariance, correlation and relative jumps/co-jumps						
	RCov	ROCov	RCorr	ROCorr	RJ	
Mean	0.037 %	0.026 %	0.598	0.528	39.03 %	
Median	0.016 %	0.007 %	0.669	0.603	36.34 %	
Min.	-0.034 %	-0.019 %	-0.677	-0.986	-2744.20 %	
Max.	1.819 %	1.719 %	0.986	1.000	555.14 %	
Std.Dev.	0.079 %	0.071 %	0.292	0.350	102.79 %	
Skew.	12.53 **	14.44 **	-0.955 **	-0.708 **	-20.053 **	
Kurt.	257 **	319 **	0.715 **	0.026	542.312 **	
JB	2765775 **	4246694 **	172 **	83 **	12259697 **	
AD	+Inf. **	+Inf. **	20.52 **	19.89 **	+Inf. **	
SW	0.384 **	0.321 **	0.923 **	0.931 **	0.278 **	

RVar = realized variance, ROVar = realized outlyingness quadratic weighted variance, RV = realized volatility, ROV = realized outlyingness quadratic weighted volatility, lnRV = logarithmic realized volatility, lnROV = logarithmic realized outlyingness quadratic weighted volatility, RCov = Realized covariance, ROCov = Realized outlyingness weighted quadratic covariance, RCorr = Realized correlation, ROCorr = Realized outlyingness weighted quadratic correlation, RJ = Realized relative jumps and co-jumps. 5% significant level is marked by *, 1 % by **. JB = Jarque-Bera test, AD = Anderson-Darling test, SW = Shapiro-Wilk test.

Our electricity forward-contract distributional property evidence on realized variance, covariance, volatility, logarithmic volatility and correlation is to a large extent consistent with earlier studies of individual stocks and stock indexes (e.g. Andersen *et al* 2001b; Thomakos and Wang 2003), bonds (e.g. Thomakos and Wang 2003), currencies (Andersen *et al* 2001a; Dacorogna *et al* 2001; Thomakos and Wang 2003) and oil and gas (Wang *et al* 2008).

The realized relative jumps/co-jumps component (RJ) affects 39% of the observations. This component is significantly left-skewed distributed, and implies a larger probability for smaller (than median) jumps/co-jumps than larger (than median) jumps/co-jumps.

As expected, the outlyingness weighted variance and volatilities measure show less fluctuation through time than the non-smoothed distributions, and we can infer that some jumps and co-jumps (or other noise) have been removed with the outlyingness weighted estimator.

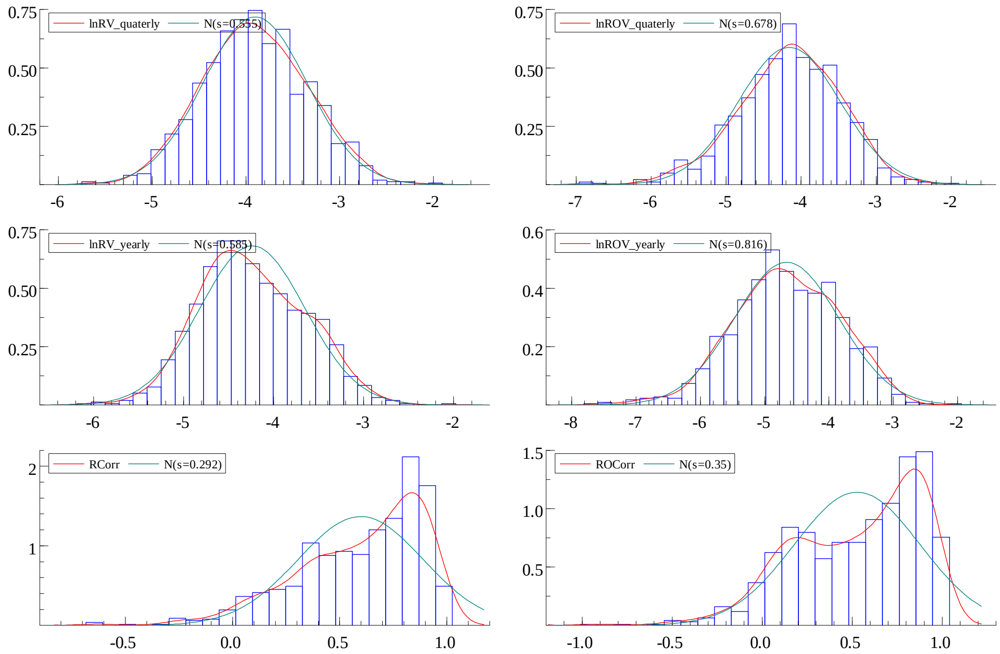

FIGURE 1 Density distribution, histogram and normal reference distribution of quarterly contract, lnRV_quarterly, for yearly contract, lnRV_yearly and for realized correlation between quarterly and yearly contracts, RCorr to the left. Corresponding graphs for the outlyingness weighted measure to the right.

4.2. Temporal dependence

In this section we explore the time-series properties of the realized variance, volatility, covariance, correlation, and relative jumps/co-jumps variable of the quarterly and yearly forward contact data. Table 3 reports a number of test statistics about the temporal-dependency properties for the aforementioned variables. The lines denoted Q20 summarizes the value of the standard Box-Pierce test for joint significance of the first 20 autocorrelations. The null hypothesis of no autocorrelation is clearly rejected for all realized variance, volatility, covariance and correlation measures analyzed, except for the realized relative jumps and co-jumps variable (RJ). The right panels of Figure 2 depict the autocorrelations for logarithmic realized volatility ($lnRV$) and logarithmic realized outlyingness weighted quadratic volatility ($lnROV$) for quarterly and yearly contacts. From the figure, we see that autocorrelations are systematically above the conventional Bartlett 95% confidence error bounds (at least for more than 100 lags), which confirm the autocorrelation tests from the Box-Pierce tests. Moreover, the realized correlation ($RCorr$) and realized outlyingness weighted correlation ($ROCorr$) measure in the right panels of Figure 3 show strong autocorrelation, but weaker than for the individual variance and correlation variables. The realized relative jumps and co-jumps (RJ) show insignificant autocorrelation patterns, which means that it is difficult to model this variable with standard univariate time-series techniques.

From Figure 2 we observe that the autocorrelations for the volatility variable start around 0.5 and decay very slowly, which suggests a long-memory pattern. The correlation measures are less autocorrelated and have fewer significant lags than the volatility measures (cf. Figure 3). In any case, the autocorrelation functions seem to decay at a slow hyperbolic rate, in contrast to the geometric decay rate associated with the conventional stationary $I(0)$ process, or alternatively to an infinite persistence pattern resulting from a non-stationary unit-root $I(1)$ process. The hyperbolic-decay process is a fractionally integrated process with a fractional order ranging from 0 to 1. When the fractional order is between 0 and 0.5, the process is mean-reverting stationary (Chen *et al* 2006). To test the decay rate we used both the non-stationary unit-root ADF test (Dickey and Fuller 1979) and the stationary KPSS test (Kwiatkowski *et al* 1992). Further, we calculated the fractional integrated estimate, d_{GPH}, with the method proposed by Geweke and Porter-Hudak (1983). For all measures analyzed (in Table 3), the null hypothesis of non-stationarity, or unit root (via ADF test), is rejected. Conventional stationarity, examined with the KPSS method, is also rejected for all contracts and measures investigated. These two tests provide initial evidence for long memory and hyperbolic decay of the autocorrelation function. The estimated d_{GPH} values for the variance/volatility, covariance/correlation and jump/co-jump measures all have $0 < d_{GPH} < 0.5$, showing that these measures are fractionally integrated and mean-reverting stationary. We may also note that the realized volatility measures are in general more persistent (i.e., higher d_{GPH} values) than the realized covariance/correlation measures, whereas the jump/co-jump measures are not statistically significant fractionally integrated. The temporal characteristics of realized variance/volatilities and covariance/correlations for the electricity forward data are consistent with the findings from Andersen *et al* (2001a) for exchange rates, Andersen *et al* (2001b) for stocks and Thomakos and Wang (2003) for various futures contracts. In contrast to our results, Wang *et al* (2008) did not find that realized correlation between futures contract prices for crude oil and natural gas exhibit any long-memory patterns.

TABLE 3 Temporal dependence of realized variance, volatility, co-variance, correlation and jumps/co-jumps.

Variance and volatility												
	RVar		ROVar		RV		ROV		lnRV		lnROV	
Quaterly												
Q20	509.24	**	326.16	**	1 911.18	**	1 455.73	**	3 115.86	**	2 612.49	**
ADF	-6.18	**	-9.36	**	-6.27	**	-7.12	**	-4.84	**	-5.67	**
KPSS	1.373	**	1.893	**	4.142	**	4.649	**	3.055	**	5.362	**
AR	9		4		4		4		10		6	
d_{GPH}	0.230	**	0.209	**	0.302	**	0.291	**	0.358	**	0.344	**
Yearly												
Q20	669.03	**	460.28	**	3 072.44	**	2 969.85	**	3 883.26	**	3 062.85	**
ADF	-9.34	**	-9.87	**	-5.35	**	-6.49	**	-4.52	**	-5.16	**
KPSS	2.407	**	2.880	**	3.617	**	6.689	**	3.297	**	5.962	**
AR	4		4		6		4		9		6	
d_{GPH}	0.335	**	0.297	**	0.408	**	0.395	**	0.353	**	0.290	**
Covariance, correlation and relative jumps/co-jumps												
	RCov		ROCov		RCorr		ROCorr		RJ			
Q20	344.85	**	241.97	**	347.00	**	385.27	**	5.27			
ADF	-9.81	**	-10.53	**	-5.65	**	-7.79	**	-21.37	**		
KPSS	2.173	**	2.097	**	1.402	**	2.900	**	0.421	*		
AR	4		4		10		6		1			
d_{GPH}	0.251	**	0.258	**	0.190	**	0.199	**	0.010			

RVar = realized variance, ROVar = realized outlyingness quadratic weighted variance, RV = realized volatility, ROV = realized outlyingness quadratic weighted volatility, lnRV = logarithmic realized volatility, lnROV = logarithmic realized outlyingness quadratic weighted volatility, RCov = Realized covariance, ROCov = Realized outlyingness weighted quadratic covariance, RCorr = Realized correlation, ROCorr = Realized outlyingness weighted quadratic correlation, RJ = Realized relative jumps and co-jumps. The Q20 is the Box-Pierce statistics, where the null hypothesis rejects zero autocorrelation from lag 1 up to 20. ADF is the augmented Dickey-Fuller test (where the null-hypothesis is non-stationarity) with an intercept and no time trend. KPSS test of Kwiatkowski *et al* (1992), where the null-hypothesis is conventional stationarity, also has no time trend. Number of lags in the ADF and KPSS tests are determined by the Akaike's Information Criteria (AIC), and is given in the rows denoted AR. d_{GPH} is the long-memory test by Geweke and Porter-Hudak (1983). A level of significance of 5% is marked by *, 1% by **.

From Figure 3 we also observe that the realized correlation ($RCorr$) is very volatile. In many periods the realized correlation is close to one, but in some periods close to zero and sometimes clearly negative. The same applies to a large extent for realized outlyingness weighted correlation ($ROCorr$), shown in the middle-left panel in Figure 3.

In Table 4 we document to what extent the volatility measures linearly move together, and how volatilities measures and correlation measures are correlated, which is frequently termed the volatility-in-correlation effect. Table 4 indicates a positive association at 0.27 and 0.40 between realized correlation ($RCorr$) and logarithmic realized volatility ($lnRV$) for quarterly and yearly contracts, respectively. Between realized outlyingness weighted correlation and logarithmic realized outlyingness weighted volatility, the numbers are 0.40 and 0.57, respectively. These results, consistent with the findings by, for example, Andersen *et al* (2001a) on currency or Andersen *et al* (2001b) on stocks, imply that correlations between yearly and quarterly forward contracts on electricity are on average higher when the volatility in the yearly and quarterly contracts is high, compared to when the volatility in these contracts are in a low-volatility state. This again suggests that models based on constant correlation, such as mean-variance efficiency analysis, are misguided.

The forward contract prices not only tend to move together, as indicated by the positive means for covariance and correlation reported in Table 3, but their volatility measure do so as well, reported in Table 4. Correlation between realized volatility for quarterly and yearly contracts are 0.67 for the non-outlyingness weighted volatility measure, and 0.57 for the outlyingness weighted volatility measures. The tendency of realized volatility to vary in tandem was also found, say, for currencies (Andersen *et al* 2001a) or for stocks (Andersen *et al* 2001b). There is also a quite highly positive correlation at over 0.70 between non-outlyingness and outlyingness weighted measures.

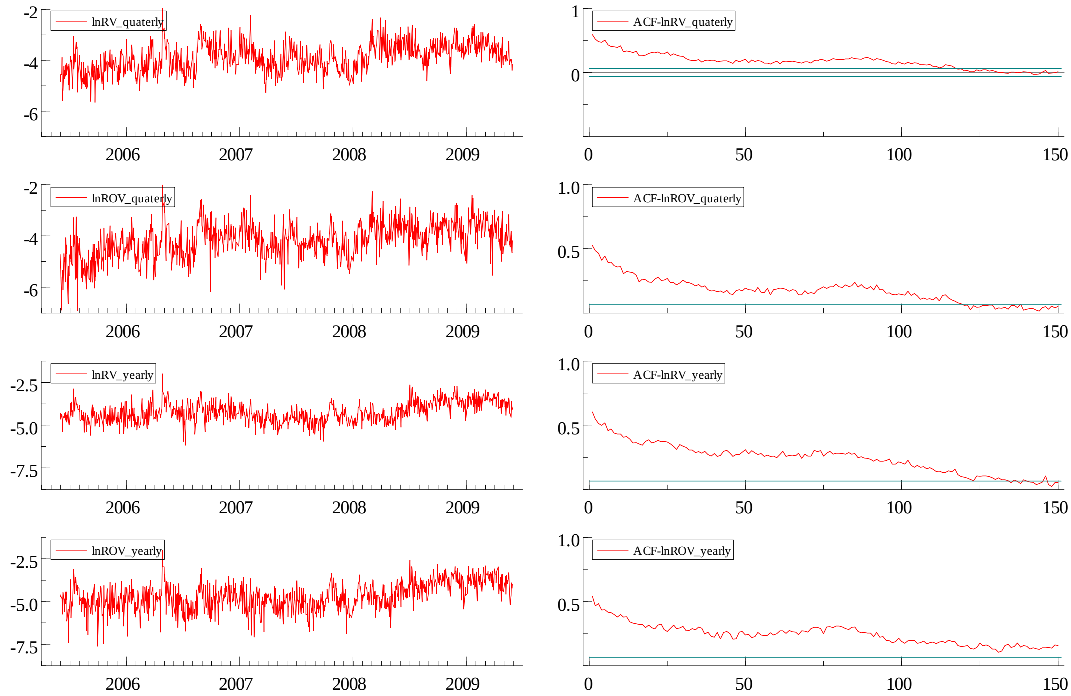

FIGURE 2 Plots of logarithm of realized variance and autocorrelation function (ACF) up to 150 lags for the yearly and monthly electricity forward contracts.

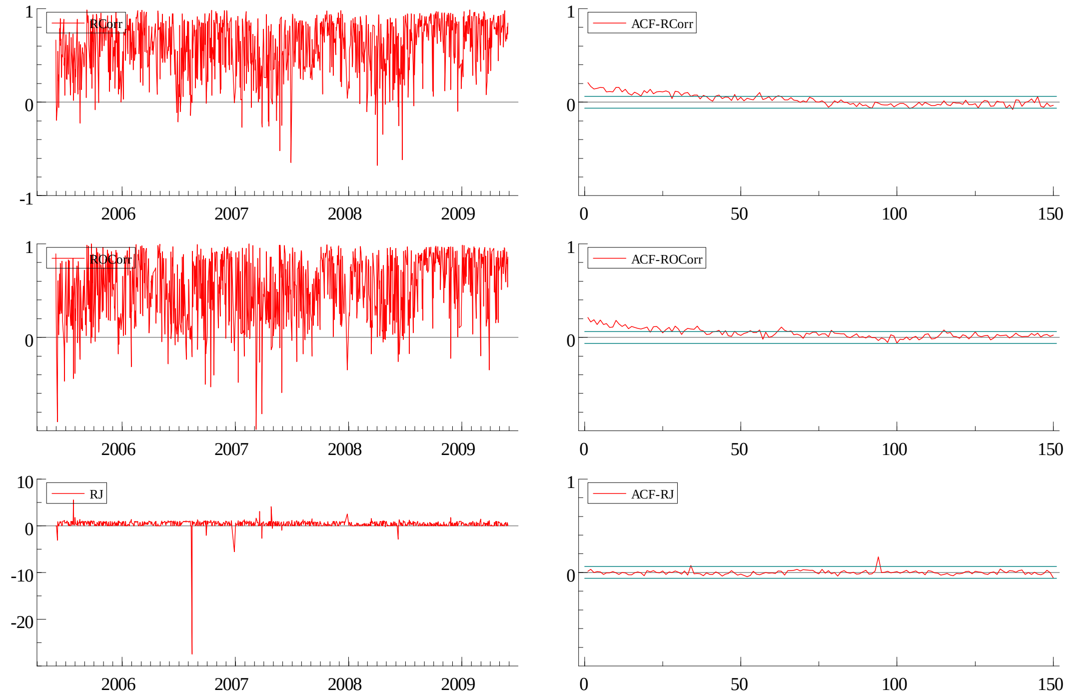

FIGURE 3 Plots of realized correlation and realized relative jumps and co-jumps between the yearly and monthly electricity forward contracts, in addition to the autocorrelation function (ACF) up to 150 lags.

TABLE 4 Pearson's correlation coefficients between realized correlation (both outlyingness and non-outlyingness weighted), the corresponding logarithmic realized volatility (both outlyingness and non-outlyingness weighted), time-to-delivery (in days) and change in trading volume (from day to day).

Variables	lnRV yearly	lnROV yearly	lnRV quarterly	lnROV quarterly	RCorr	ROCorr
lnRV yearly			**0.670**		**0.397**	
lnROV yearly	**0.768**			**0.568**		**0.572**
Time-to-delivery yearly	-0.027	0.016			**-0.105**	-0.053
Volume change yearly	**0.139**	**0.126**			**0.122**	**0.134**
lnRV quarterly					**0.270**	
lnROV quarterly			**0.823**			**0.400**
Time-to-delivery quarterly			**-0.235**	**-0.300**	-0.062	**-0.069**
Volume change quarterly			**0.131**	**0.109**	**0.109**	**0.098**
ROCorr					**0.720**	

Values in bold are different from 0 with 5% significance level. Non-relevant correlation coefficients are not reported.

As also shown in Table 4, the total trading volume change (from day to day) are positive correlated with both realized volatility and realized correlation. This result is consistent with the findings of Giot *et al* (2009), who state that much of the empirical literature document a positive contemporaneous relation between volume and volatility in financial markets. Moreover, the multivariate analysis for stock returns by Gallant *et al* (1992) also reports evidence of contemporaneous relationship in volatility.

According to Samuelson (1965), the volatility of futures price returns should increase as time-to-maturity decrease. For the yearly contracts, the realized volatility measure shows no statistically clear evidence of the Samuelson effect. However, the quarterly-contract series show evidence of increasing realized volatility as time-to-delivery decreases (significant at the 5% level). Table 4 also shows that decreasing time-to-delivery is positively related to increasing realized correlation (but this association does not apply for all measures found to be significant at the 5% level).

For modeling and forecasting realized volatilities and correlation, future studies would do well to take the stylized facts described in this section into account.

5 CONCLUDING COMMENTS

The main findings obtained in this study of Nord Pool electricity forward data show that the logarithmic realized volatility are approximately normally distributed, while realized correlation seems not to be. Further, realized volatility has a long-memory feature, and there seem to be a high correlation between realized correlation and volatilities. These results are to a large extent consistent with earlier, similar studies of stylized facts of other financial and commodity markets. In contrast to the study of crude oil and natural gas by Wang *et al* (2008), we also found a long-memory pattern in realized correlation.

The results from this study have implications for future research and analyses. First, electricity volatility modeling and forecasting based on high-frequency data can draw from similar studies of other financial and commodity markets. Secondly, realized volatilities and probably also realized correlation should be modeled by fractionally integrated processes (ARFIMA model). Thirdly, models based on constant correlation seem inappropriate. Fourthly, realized correlation is, as expected, positive, but very volatile and so the hedge position with these quarterly and yearly electricity forward contracts points to high risks.

More research is needed to investigate both the usefulness of smoothed outlyingness measures as compared to unsmoothed realized measures and how to disentangle the co-jump component from the continuous diffusion process. Future research could combine realized correlation with similar measures from multivariate conditional volatility models, multivariate stochastic volatility models and implied correlation from option-market data. Finally, realized correlation analyses between various energy markets and between energy and other markets would be fruitful areas for future research.

REFERENCES

Andersen, T.B, Benzoni, L., and Lund, J. (2002). An empirical investigation of continuous-time equity return models. *Journal of Finance* 57, 1239-1284.

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 and Statistics* 89, 701-720.

Andersen, T.G, Bollerslev, T., Diebold, F.X., and Labys, P. (2001a). Distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96, 42-55.

Andersen, T.G, Bollerslev, T., Diebold, F.X., and Ebens, H. (2001b). Distribution of realized stock rate volatility. *Journal of Financial Economics* 61, 43-76.

Andersen, T.G, Bollerslev, T., Diebold, F.X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica* 71, 579-625.

Andersen, T.G., and Bollerslev, T. (1998). Answering the sceptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39, 885-905.

Andersen, T.G., Bollerslev, T., Christoffersen, P.F., and Diebold, F.X. (2006). Volatility and correlation forecasting. In: *Handbook of Economic Forecasting, Volume 1*. Ed by Elliot, G., Granger, C.W.J., Timmermann, A., pp. 777-878. North-Holland, Amsterdam.

Andersen, T.G., Dobrev, D., and Schaumburg, E. (2009). Jump-robust volatility estimation using nearest neighbor truncation. CREATES Research Paper 2009-52. Access internet: http://ideas.repec.org/p/aah/create/2009-52.html.

Areal, N.M.P.C., and Taylor, S.J. (2002). The realized volatility of FTSE-100 futures prices. *Journal of Futures Markets* 22, 627-648.

Bannough, K., van Duk, D., and Martens, M. (2009). Range-based covariance estimation using high-frequency data: the realized co-range. *Journal of Financial Econometrics* 7, 341-372.

Barndorff-Nielsen, O.E., and Shepard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economies. *Journal of the Royal Statistical Society B* 63, 167-241.

Barndorff-Nielsen, O.E., and Shephard, N. (2004a). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica* 72, 885-925.

Barndorff-Nielsen, O.E., and Shephard, N. (2004b). Measuring the impact of jumps in multivariate price processes using bipower covariation. Discussion paper, Nuffield College, Oxford University.

Boudt, K., Croux, C., and Laurent, S. (2008). Outlyingness weighted quadratic covariation. Access internet: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1149728.

Chan, K.F., Gray, P., and van Campen, B. (2008). A new approach to characterizing and forecasting electricity price volatility. *International Journal of Forecasting* 24, 728-743.

Chen, Z., Daigler, R.T., and Parhizgari, A.M. (2006). Persistence of volatility in futures markets. *Journal of Futures Markets* 26, 571-594.

Dacorogna, M.M, Gencay, R., Muller, U., Olsen, R.B., and Pictet, O.V. (2001). *An Introduction to High-Frequency Finance*, Academic Press, San Diego, CA.

Dickey, D.A., and Fuller, W.A. (1979). Distributions of the estimators for autoregressive time series with unit root. *Journal of the American Statistical Association* 74, 427-431.

Gallant, A.R., Rossi, P.E. and Tauchen, G. (1992). Stock prices and volume, *Review of Financial Studies* 5, 199-242.

Geweke, J., and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis* 4, 221-238.

Giot, P., Laurent, S., and Petitjean, M. (2009). Trading activity, realized volatility and jumps, *Journal of Empirical Finance* (forthcoming).

Haugom, E., Westgaard, S., Solibakke, P.B., and Lien, G. (2009). Modelling and forecasting realized volatility for electricity ticker forward prices. Submitted to *Energy Economics*.

Huang, X., and Tauchen, G. (2005). The relative contribution of jumps in total price variance. *Journal of Financial Econometrics* 3, 456-499.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of unit root. *Journal of Econometrics* 54, 159-178.

Laurent, S. (2009). *Estimating and Forecasting ARCH Models Using G@RCH™ 6*. Timberlake Consultants, UK.

Martens, M. (2002). Measuring and forecasting S&P 500 index-futures volatility using high-frequency data. *Journal of Futures Markets* 22, 497-518.

Racicot, F-É., Théoret, R., and Coën, A. (2008). Forecasting irregularly spaced UHF financial data: Realized volatility vs UHF-GARCH models. *International Atlantic Economic Society*. DOI 10.1007/s11294-008-9134-2

Samuelson, P.A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* 6, 41-49.

Thomakos, D.D., and Wang, T. (2003). Realized volatility in the futures markets. *Journal of Empirical Finance* 10, 321-353.

Ulrich, C.J. (2009). Realized volatility and price spikes in electricity markets: The importance of observation frequency, Available at SSRN: http://ssrn.com/abstract=1342586.

Wang, T., Wu, J., and Yang, J. (2008). Realized volatility and correlation in energy futures markets. *Journal of Futures Markets* 28, 993-1011.