EconomiX

http://economix.u-paris10.fr/

Document de Travail

Working Paper

2009-24

On the realized volatility of the ECX CO2 emissions 2008 futures contract: distribution, dynamics and forecasting

Julien Chevallier et Benoît Sévi

UMR 7166 CNRS

Université Université de Paris Ouest Nanterre La Défense
(bâtiments K et G)
200, Avenue de la République
92001 NANTERRE CEDEX

Tél et Fax : 33.(0)1.40.97.59.07
Email : secretariat-economix@u-paris10.fr

Université Paris X Nanterre

On the realized volatility of the ECX CO_2 emissions 2008 futures contract: distribution, dynamics and forecasting[1]

Julien Chevallier[2]

Imperial College London (Grantham Institute for Climate Change) and University of Paris 10 (EconomiX-CNRS)

Benoît Sévi[3]

University of Angers (GRANEM) and LEMNA

May 2009

Abstract: The recent implementation of the EU Emissions Trading Scheme (EU ETS) in January 2005 created new financial risks for emitting firms. To deal with these risks, options are traded since October 2006. Because the EU ETS is a new market, the relevant underlying model for option pricing is still a controversial issue. This article improves our understanding of this issue by characterizing the conditional and unconditional distributions of the realized volatility for the 2008 futures contract in the European Climate Exchange (ECX), which is valid during Phase II (2008-2012) of the EU ETS. The realized volatility measures from naive, kernel-based and subsampling estimators are used to obtain inferences about the distributional and dynamic properties of the ECX emissions futures volatility. The distribution of the daily realized volatility in logarithmic form is shown to be close to normal. The mixture-of-distributions hypothesis is strongly rejected, as the returns standardized using daily measures of volatility clearly departs from normality. A simplified HAR-RV model (Corsi, 2009) with only a weekly component, which reproduces long memory properties of the series, is then used to model the volatility dynamics. Finally, the predictive accuracy of the HAR-RV model is tested against GARCH specifications using one-step-ahead forecasts, which confirms the HAR-RV superior ability. Our conclusions indicate that *(i)* the standard Brownian motion is not an adequate tool for option pricing in the EU ETS, and *(ii)* a jump component should be included in the stochastic process to price options, thus providing more efficient tools for risk-management activities.

JEL Classification: C5, G1, Q4.

Keywords: CO_2 Price, Realized Volatility, HAR-RV, GARCH, Futures Trading, Emissions Markets, EU ETS, Intraday data, Forecasting.

---

[1] Acknowledgments to be added following the publication process.

[2] Address for correspondence: Imperial College of Science, Technology and Medicine, South Kensington Campus, London SW7 2AZ, UK. Email: **j.chevallier@imperial.ac.uk**

[3] Address for correspondence: Faculty of Law, Economics and Management, University of Angers, 13 allée François Mitterrand, BP 13633, 49036 Angers cedex 01, France. Email: **benoit.sevi@gmail.com**

1 Introduction

The distribution of financial asset returns is often modeled following mixtures of normal distributions that have different parameters (Dacorogna *et al.*, 2001). The distributional and dynamic properties of volatility appear especially important for risk-management purposes, since different specifications will yield to various pricing structures (Guillaume *et al.*, 1997). The investigation of such properties has been revivified by the recent literature on realized volatility, which relies on the use of intraday data. Since the seminal contributions of Andersen, Bollerslev, Diebold and Labys (henceforth ABDL, 2001), Andersen, Bollerslev, Diebold and Ebens (henceforth ABDE, 2001), and Barndorff-Nielsen and Shephard (henceforth BNS, 2002), among others, the literature on realized volatility measures has been very prolific[4].

This article uses tick-by-tick data of CO_2 emissions allowances, valid for compliance under the EU Emissions Trading Scheme (EU ETS), exchanged on the European Climate Exchange (ECX) based in London. More particularly, we use the futures contract of maturity December 2008 to examine the unconditional and conditional (dynamic) distributions of the ECX CO_2 emissions futures volatility. This new analysis appears important on such an emerging market, where the understanding of the volatility properties of CO_2 prices will allow a better characterization of the relevant stochastic process to price derivatives (Tucker (2001), Chevallier *et al.* (2009)). It appears also of primary importance to hedge against various kinds of institutional, economic or financial risks (see Busch and Hoffman, 2007). Hence, the research question developed in this article may be of precious use for risk-management purposes, which requires a careful understanding of the volatility of CO_2 prices.

The statistical properties of daily realized volatilities in futures markets has been investigated, among others, in Thomakos and Wang (2003). Their analysis of D-Mark, E-Dollar, S&P500 and T-bonds shows that standard deviations exhibit long memory, while standardized returns are serially uncorrelated. They also found that the unconditional distributions of daily returns' volatility are leptokurtic and highly skewed to the right, while the distributions of standardized returns and logarithmic standard deviations are close to a Gaussian distribution.

Luu and Martens (2003) test the *mixture-of-distributions-hypothesis* (MDH) (Clark (1973), Tauchen and Pitts (1983)) by comparing volatility models using daily and intraday data. Our approach consists in applying this research question to the study of ECX CO_2 emissions futures. The first use of intraday data for CO_2 emissions markets may be related to Benz and Klar (2008), who investigate the price discovery between various exchanges. To our best knowledge, our article constitutes the first attempt to derive the volatility properties of CO_2 emissions futures using realized measures.

Our data set contains one year of tick-by-tick data from ECX CO_2 emissions futures, corresponding to the 2008 futures contract. The choice to restrain our analysis to the 2008 contract is motivated by *(i)* the erratic behavior of *spot* prices during 2005-2007 due to banking restrictions (Alberola and Chevallier, 2009), which proved to be less robust than *futures* for price signalling in the medium-term; and *(ii)* the validity of the 2008 contract during Phase II (2008-2012), which offers the "bankability"

[4]Surveys may be found in Zivot (2005), McAleer and Medeiros (2008), Andersen and Benzoni (2009).

of CO_2 emissions allowances until the end of Phase III (2013-2020).

Since the end of 2007, both the liquidity of the EU ETS and the availability of high-frequency data have been increasing. ECX emissions futures are indeed the most heavily traded emissions contracts, followed by spot and option prices. The volume of intraday transactions recorded on the ECX CO_2 emissions futures market is approximately equal to one tenth of Foreign Exchange (FX) markets, which are opened 24 hours. With an average of 700 trades per day and 50 seconds between each transaction, the tick-by-tick data gathered for ECX CO_2 emissions futures is somewhat comparable to the values found on other financial markets, such as the level of daily transactions for the D-Mark as documented in Thomakos and Wang (2003).

This article provides the first empirical application of the methodology by ABDL (2001) and ABDE (2001) to ECX CO_2 emissions futures. We use one year of 15-minute returns[5] from the futures contract to estimate the daily realized volatility, and hence to describe the distribution and time-series properties of ECX CO_2 emissions futures. Compared to previous literature, the estimates of intraday volatility based on realized measures are more accurate than the estimates based on daily data which are used in Paolella and Taschini (2008), Benz and Truck (2009), Daskalakis *et al.* (2009) and Oberndorfer (2009), among others.[6]

Our methodology consists in dealing with the distributional, dynamic, and forecasting properties of realized volatility for ECX CO_2 emissions futures. We study the unconditional distributions of realized volatility measures, while testing for several transformations to approach normality. We also test whether the MDH holds for ECX CO_2 emissions futures. Then, we investigate the dynamics of realized volatility measures using an Heterogeneous Autoregressive Model of the Realized Volatility (HAR-RV) developed in Corsi (2009) versus GARCH specifications. We finally propose a forecasting exercise, by testing the predictive accuracy of the HAR-RV model versus other models of conditional volatility based on daily data.

Our main results may be summarized as follows. We first document the near normality of the logarithmic form of realized volatility measures for the ECX 2008 futures contract. This is standard in financial literature, as the "spot volatility" which governs the Brownian motion is generally assumed to be lognormally distributed. Nevertheless, the standardized returns (using realized volatility) are not normally distributed, which stands against the MDH. Standardized returns using GARCH volatilities are more normally distributed, which is not usual for financial series. Finally, the HAR-RV model with a daily and a weekly component outperforms unambiguously GARCH specifications in terms of dynamic modeling and forecasting accuracy. The latter result is due to the superiority of realized measures in estimation using intraday data over lower frequency variations.

Several directions may be pursued in extension of our work. The investigation of jump components in realized volatility measures appears of primary interest, by using standardized bi-power and tripower

[5] The optimal sampling frequency is chosen so as to limit the impact of market microstructure effects.

[6] Our analysis remains univariate. Using high-frequency data, a multivariate analysis such as Cartea *et al.* (2007) or Bunn and Fezzi (2007) does not seem appropriate, because of the complex relationships linking CO_2 emissions and energy markets. Thus, the study of realized covariance and realized correlations of ECX CO_2 emissions futures with other high-frequency energy futures price series is not considered here.

variation (Andersen, Bollerslev and Diebold, henceforth ABD, 2007). The formal determination of the optimal sampling frequency also appears as a promising area for future research using specific microstructure noise detection tests (see Awartani *et al.*, 2009).

The remainder of the article is organized as follows. Section 2 provides an overview of futures trading on the EU ETS. Section 3 reviews estimation methods for realized volatility, discusses optimal sampling frequency issues and maturity effects characteristic of futures contracts. Section 4 studies the unconditional distribution of ECX CO_2 emissions futures returns and realized volatility, as well as the distributional properties of returns and standardized returns, using several transformations for realized volatility measures. Section 5 investigates realized volatility dynamics, and especially long memory components using the HAR-RV model. Section 6 provides a forecasting exercise to test the accuracy of the HAR-RV model against the predictive power of daily GARCH forecasts. Section 7 concludes.

2 The European CO_2 emissions futures market

In this section, we present first the key design issues on the European CO_2 emissions futures market, second we discuss the main characteristics of futures trading on ECX, and third we proceed with a preliminary analysis of the intraday data used.

2.1 Design and transactions growth

Let us discuss first some key design issues, as well as the growth of transactions recorded on the European CO_2 emissions market since its creation on January 1, 2005.

2.1.1 Key design issues

The European Union Emissions Trading Scheme (EU ETS) has been created by the Directive 2003/87/CE. Across its 27 Member States, the EU ETS covers large plants from CO_2-intensive emitting industrial sectors with a rated thermal input exceeding 20 MWh. One allowance exchanged on the EU ETS corresponds to one ton of CO_2 released in the atmosphere, and is called an European Union Allowance (EUA). 2.2 billion allowances per year have been distributed during Phase I (2005-2007). 2.08 billion allowances per year will be distributed during Phase II (2008-2012). With a value of around €20 per allowance, the launch of the EU ETS thus corresponds to a net creation of wealth of around €40 billion. In January 2008, the European Commission extended the scope of the EU trading system to other sectors such as aviation and petro-chemicals by 2013, and confirmed its functioning for a third Phase until 2020.

2.1.2 Transactions growth

During Phase I of the EU ETS (2005-2007), the total volume of allowances exchanged has been steadily increasing. The number of transactions has been multiplied by a factor four between 2005 and 2006,

going from 262 to 809 million tons. This increasing liquidity of the market has been confirmed in 2007, where the volume of transactions recorded is equal to 1.5 billion tons. This peak of transactions may be explained by the growth of the number of contracts with delivery dates from December 2008 to December 2012, which represented 4% of total exchanges in 2005, and 85% in 2007. These transactions reached €5.97 billion in 2005, €15.2 billion in 2006, and €24.1 billion in 2007, thereby confirming the status of the EU ETS as the largest emissions trading scheme to date in terms of transactions.

In 2008, the carbon market was worth between €89 billion and €94 billion, up more than 80% year-on-year, according to analysts (Reuters). The launch of secondary certified emission reduction (CER)[7] contracts on ECX certainly fostered this growth rate of transactions.

2.2 Futures trading

As discussed below, due to the non-reliable behavior of *spot* prices during Phase I (2005-2007), we decide to use *futures* prices valid for Phase II (2008-2012). More specifically, we choose to investigate in this article the volatility dynamics of the December 2008 futures contract traded in €/ton of CO_2 on ECX.

2.2.1 Price development

ECX futures trading started on April 22, 2005 with varying delivery dates going from December 2005 to December 2012. Futures contracts with vintages December 2013 and 2014 were introduced on April 8, 2008. Daily closing prices trade at €13.32/ton of CO_2 as of January 15, 2009, and have reached a maximum price of €32.90/ton of CO_2 in 2008[8].

Insert Figure 1 about here

Figure 1 shows the futures price development for contracts of maturities December 2005 through 2014 from April 22, 2005 to January 16, 2009. We may observe that futures prices for delivery during Phase II (2008-2012) proved to be much more reliable than futures prices for delivery during Phase I (2005-2007), due to the banking restrictions enforced between the two Phases (Alberola and Chevallier (2009)). Market observers noticed a divergence between Phase I spot and futures prices - which decreased towards zero - and Phase II futures prices - which conveyed a medium-term price signal around €20/ton of CO_2 throughout the historical data available for the second phase of the scheme. The price development for Phase II futures comprises a lower bound around 15€/ton of CO_2 in April 2007, and an upper bound around 35€/ton of CO_2 in November 2008.

[7] According to the article 12 of the Kyoto Protocol, Credit Development Mechanism (CDM) projects consist in achieving GHG emissions reduction in non-Annex B countries. After validation, the UNFCCC delivers credits called CERs that may be used by Annex B countries for use towards their compliance position. CERs are fungible with EU ETS allowances with a maximum limit of around 13.4% on average.

[8] In the longer term, analysts forecast EUA prices of €20-25/ton of CO_2 over Phase II and €25-30/ton of CO_2 over Phase III, which will run from 2013-20 (Reuters).

2.2.2 Contract specifications

The ECX CO_2 emissions futures contract is a deliverable contract where each member with a position open at cessation of trading for a contract month is obliged to make or take delivery of emission allowances to or from national registries. The unit of trading is one lot of 1,000 emission allowances. Each emission allowance represents an entitlement to emit one tonne of carbon dioxide equivalent gas. Market participants may purchase consecutive contract months to March 2008, and then December contract months from 2008 to 2012[9]. Trading occurs from 07:00AM to 05:00PM GMT. Allowances delivery typically occurs by mid-month of the expiration contract date. The ECX December 2008 futures contracts expired on December 15, 2008. The first delivery of the underlying CO_2 allowance occurred on December 16, 2008, and the last delivery on December 18, 2008.

2.3 Preliminary analysis of the intraday data

Our sample contains one year of tick-by-tick transactions for the ECX futures contract of maturity December 2008, going from January 2 to December 15, 2008. This is equivalent to 240 days of trading after cleaning the data for outliers, and until the expiration of the contract. Intraday data with a one-year time horizon have been studied, for instance, by Taylor and Xu (1997) for the DM/$ exchange rate. The total amount of intraday observations in our sample is equal to 167,004. The ECX CO_2 emissions futures tick data thus corresponds to one tenth of the transactions recorded on FX markets - which are opened 24 hours and reach more than 15,000 trades per day. However, this level of transactions appears comparable to the values found for other markets. For instance, Thomakos and Wang (2003) note that the average number of price changes per day is 163 for the Eurodollar, 3,366 for the S&P500, and 1,710 for T-bonds. The average amount of transactions for the ECX CO_2 emissions futures tick-by-tick data is equal to 700 trades per day. This corresponds to an average of 50 seconds between each transaction.

In the next section, we detail how to compute realized volatility measures.

3 Estimation of realized volatility

In this section, first we review the theoretical background to derive realized volatility measures from intraday data, second we present different estimation methods, third we discuss the issue of optimal sampling frequency choice and the maturity effect in the futures contract.

3.1 Theory

Let $p(t)$ denote a logarithmic asset price at time t. Abstract from a jump process, the continuous-time diffusion process generally employed in asset and derivatives pricing may be expressed by a stochastic differential equation as:

[9]Note *spreads* between two futures contracts may also be traded.

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) \quad \text{with} \quad 0 \leq t \leq T \tag{1}$$

with $\mu(t)$ a continuous and locally bounded variation process, $\sigma(t)$ a strictly positive càdlàg (right continuous with left limits) stochastic volatility process, and $W(t)$ a standard Brownian motion. Note that the formulation in equation (1) is very general, includes most of the processes generally used in standard asset pricing theory (see ABDL (2001)) and may accommodate for long memory components.

Next, let us consider the quadratic variation (QV) for the cumulative return process $r(t) \equiv p(t) - p(0)$:

$$[r, r]_t = \int_0^t \sigma^2(s)ds \tag{2}$$

The QV simply equals the integrated volatility of the process described in equation (1). Now, assume that returns are sampled on a Δ-period yielding $r_{t,\Delta} \equiv p(t) - p(t - \Delta)$. The *realized variance*[10] (RV) is defined as the sum of the corresponding $1/\Delta$, which is assumed to be an integer for simplicity, high-frequency intraday squared returns, or:

$$RV_{t+1}(\Delta) \equiv \sum_{j-1}^{1/\Delta} r^2_{t+j.\Delta,\Delta} \tag{3}$$

Andersen and Bollerslev (1998) followed by ABDL (2001) and BNS (2002) among others demonstrated that, as the sampling frequency of the underlying returns increases, the RV converges uniformly in probability to the increment of the QV process, or:

$$RV_{t+1}(\Delta) \to \int_0^t \sigma^2(s)ds \tag{4}$$

when $\Delta \to 0$. Thus, RV is a consistent estimator for the integrated volatility used throughout asset pricing theory. In other words, as the sampling frequency increases, the estimation error of the RV diminishes.

3.2 Estimation methods

Theory suggests that optimal sampling corresponds to sampling at the highest possible frequency. In practice, this is far from being true as shown in a series of articles starting with Andersen and Bollerslev (1998). In fact, the logarithmic return process which is truly observed does not comply with the hypothesis of a semimartingale for the underlying process, which is a necessary hypothesis for deriving results discussed in the previous section. This issue is discussed in ABDL (2001) and Zhang *et al.* (2005) for instance. The latter authors describe this phenomenon as emerging from market microstructure problems, whose main examples are the existence of a bid-ask spread, non-synchronous trading, etc.

[10] Some authors refer to this as *realized volatility*, but we reserve this term for the square root of realized variance that is also considered in this article.

To mitigate the impact of microstructure noise, various methodologies have been employed in the empirical financial literature. These include the determination of the optimal Δ as described in Aït-Sahalia *et al.* (2005) after the noise has been modelled, the use of subsampling schemes as in Zhang *et al.* (2005), resorting to pre-filtering methods as in Andreou and Ghysels (2002) or kernel-based methodologies as in Zhou (1996) or Hansen and Lunde (2006). In order to investigate the relevance of different sampling methods for the analysis of the ECX emissions futures 2008 contract, we do not only follow ABDL (2001) as is the case in most of the existing empirical literature, but we also consider two other methodologies.

First, we consider the traditional un-weighted estimator used for instance in ABDL (2001) and BNS (2002). This estimator is the natural estimator in view of theoretical developments in quadratic variation and perfectly fits equation (3), as it is the sum of squared realized returns on a given sampling frequency. For each day d and sampling frequency $1/m$, we compute:

$$RV^{d,m} = \sum_{i=1}^{m} r_{i,m}^2 \tag{5}$$

Second, we estimate realized volatility following Zhang *et al.* (2005). Their sub-sampling method appears particularly relevant for use with the ECX emissions futures intraday data, because of the limited number of daily transactions compared to other more actively traded financial assets. The idea behind sub-sampling is that when a given sampling frequency, say $1/m$, is chosen in light of the microstructure noise limited impact, a large share of the data is ignored. To fully account for the available information, Zhang *et al.* (2005) propose to average the measure of realized volatility at $1/m$ frequency but for different starting times. Let:

$$RV^{d,m,p} = \sum_{i=1+p}^{m+p} r_{i,m}^2 \tag{6}$$

be the realized variance measure at sampling frequency $1/m$, but with the first observation chosen at $1+p$ with $p < \frac{1}{m}$. By evaluating $RV^{d,m,p}$ for starting times $1, 1+p, 1+2p, ..., 2$ and keeping the sampling frequency $1/m$, we move our estimation window, and thus exploit a larger part of the data set. Zhang *et al.* (2005) then propose to average the measure considering all starting values.

Third, we retain a kernel-based estimator as first proposed in Zhou (1996). After testing for various kernel estimates, such as the modified Tukey-Hanning kernel, our choice goes to the Bartlett kernel-based estimator, which shows better performance with respect to the variability of the estimators with respect to their inputs[11].

We then consider, as is now common in the literature[12], three different proxies for volatilities. First, we study the realized variance as defined in equation (3) with a sampling frequency of 15 minutes, in view of the volatility signature plots in Figure 3 (see more on this below). Second, following ABDL (2001)

[11] Hansen and Lunde (2006) discuss this issue, and provide more details on the practical application of kernel-based methods.

[12] See ABD (2007) and references therein.

we examine the square root of the realized variance, denoted realized volatility, or $RV_t^{1/2}$. Third, we consider the logarithm of the realized volatility, or $log(RV_t^{1/2})$, also known for its convenient properties in small samples[13]. As will be discussed below, the logarithmic transformation represents one among other power transformations. A better choice may emerge following Gonçalves and Meddahi (2008).

Insert Figure 2 about here

Figure 2 plots the three proxies of volatilities (left, middle and right panels) for the three estimation methods selected (top, middle and bottom panels). The time-series reveal the presence of jumps and structural breaks that may be taken into account using multipower variation measures.[14] Note also that the time-series on the left panel reflect the exclusion of the "once-in-a-generation" (Cai *et al.* (2001), ABDL (2003)) anomalous carbon price movement detected on October 13, 2008 which seems to coincide with the depressing effect of the "credit crunch" crisis on the prices of global commodity markets.

Insert Table 1 about here

Table 1 reports the descriptive statistics for the three proxies of volatility with the three estimation methods. We observe that the daily realized variance and the daily realized volatility in standard deviation form present nonzero skewness and excess kurtosis[15]. These descriptive statistics therefore reveal a "fat tailed" leptokurtic distribution for the ECX CO_2 emissions futures contract of maturity December 2008, except for the daily realized volatility in logarithmic form.

3.3 Optimal sampling frequency and maturity effect

As is usual, we need to estimate the highest frequency at which the microstructure noise can be neglected. To this purpose, we use volatility signature plots, where the realized volatility measure described in equation (5) is computed and plotted at different sampling frequencies.

Insert Figure 3 about here

Figure 3 shows the volatility signature plot for the full (top) and November-December (bottom) samples. As in ABDL (2001) and ABDE (2001), we use these volatility signature plots to estimate the range of sampling frequencies where the volatility is strongly increasing, indicating the increasing presence of microstructure noise.

For the full sample, it appears that the choice of 15-minute returns should allow to minimize the impact of the microstructure noise, while ensuring for each day a sufficient number of observations. The use of 15-minute returns for the ECX carbon tick data also appears as a conservative choice compared to

[13] Some articles (*e.g.* ABD (2007)) consider the series of the logarithm of the realized variance instead of the logarithm of the standard deviation of the realized variance. This is of course equivalent up to a scalar.
[14] This aspect is left for further research.
[15] Note for a normally distributed random variable skewness is zero, and kurtosis is three.

5-minute returns usually chosen for FX markets. Of course, the use of volatility signature plot as a simple graphical tool to determine the optimal frequency is questionable. To overcome this difficulty, Awartani *et al.* (2009) propose a statistical test allowing to assess the incremental impact of the microstructure noise between two possible frequencies. As such, a rolling version of their procedure can be viewed as a statistically robust implementation of the volatility signature plot method in ABDL (2001). Because our contribution remains more empirically-oriented, we choose to proceed with the graphical method.

Looking at Figure 3 reveals different patterns between the full sample and the end-of-year sub-sample. We observe that the level of volatility is slightly higher at the end of the year. This is a quite standard effect on commodity futures markets, also known as the "Samuelson effect". Samuelson (1965) advocated in his seminal article that volatility is increasing near the maturity of futures contract as a response to an increasing flow of information.[16] Thus, to verify the Samuelson hypothesis, we should observe that the futures price volatility increases as the futures contract approaches its expiration date. This characteristic of financial assets has been recently proven to be valid using intraday data for a wide range of futures market, including agricultural futures (Duong and Kalev, 2008).

The inspection of the volatility signature plots for the last months of 2008 tends to confirm this hypothesis. The effects of microstructure noise seem visually more important. More importantly, the dispersion of the estimator is larger due to the low level of observations used to compute the realized variance. For the November-December period, the realized volatility estimate can lie anywhere between 0.01 and 0.025 using a sampling frequency around 15 minutes. This variability is lower for the full sample, which goes from 0.015 to 0.020 for the same sampling frequency. Nevertheless, in view of the moderate effect that we observe at the end of the sample, we choose to keep a 15-minute interval between two observations as being representative of the optimal frequency for the entire sample.

In the next section, we explain the empirical results obtained.

4 Unconditional distribution of futures returns and realized volatility

In this section, we study the unconditional distribution of realized volatilities and returns for the ECX December 2008 futures contract. We first focus on the unconditional distribution of our three proxies for realized volatility. We then study the distributional properties of daily raw returns, RV-standardized and GARCH-standardized-returns.

4.1 Distribution of realized variance and volatility

Insert Figure 4 about here

[16] See also Illueca and Lafuente (2006) for an application of the realized volatility measure to the investigation of the expiration-day effect.

We first plot unconditional distribution of realized variances and realized volatilities in the left and middle panels of Figure 4. The distribution of these volatility measures appears strongly right-skewed.

Insert Table 2 about here

This is confirmed by normality test statistics in Table 2. The kurtosis of the series indicates fat tails compared to a Gaussian distribution.

Insert Figure 5 about here

Quantile-Quantile (QQ) plots against normality in Figure 5 unambiguously reject normality for realized variance and volatility. Next, we turn to the logarithmic transformation, which is common practice since ABDL (2001), to near normality.

4.2 Distribution of the logarithmic transformation of volatility

We begin our analysis by using the logarithmic transformation as in most of the existing literature. The kernel-based distributions plotted in the right panel of Figure 4 indicate a less skewed density than for realized variance or its square root. Indeed, in view of the plotted distributions and quantile-quantile plots in the right panel of Figure 5, it appears that the logarithmic transformation of the realized volatility, while remaining left-skewed, does a better job in nearing normality. It should be noted that our kernel-based distributions are only based on 240 trading days. This limited data availability may explain the departure from normality, which is expected in small sample experiments.

To sum up, our analysis shows that the logarithmic transformation of the daily realized volatility is closer to normality than other forms of volatility. This result is in line with previous literature on the modeling of stochastic volatility (see ABDL (2001, 2003) among others), which has practical applications in option pricing.

4.3 Alternative transformations

The logarithmic transformation is only one transformation among others. Alternative transformations have been proposed to improve the normal approximation in small samples. Chen and Deo (2004)'s transformation is based on a power transformation, from which the exponent is then estimated. Unfortunately, the exponent has to be estimated knowing the asymptotic variance of realized volatility, which is not the case in practice. Gonçalves and Meddahi (2008) thus coin this statistic as "infeasible", and rely on Edgeworth expansions to determine the optimal parameter β of the Box-Cox transformation to retain in order to eliminate the skewness. We tested various values of β to better take into account the residual skewness in our series. We did not find better transformations compared to the initial logarithmic transformation[17].

[17] The results of these tests are no reported here due to space constraints, but are available from the authors upon request.

4.4 Distributional properties of returns and standardized returns

Let R_t be the daily open-to-close continuously compounded return of the futures contract for day t.

Insert Figure 6 about here

Daily raw returns are plotted in Figure 6. As is common for financial time-series, returns exhibit volatility clustering.

Insert Table 3 about here

Descriptive statistics of daily returns are provided in Table 3. We observe that the unconditional distribution of returns is close to normality with a sample skewness of -0.047 and a sample kurtosis of 3.24, thus resulting in a Jarque-Bera statistic value of 0.69 corresponding to a p-value of 0.70.

Next, we compute the series of daily standardized returns. Following Clark's (1973) seminal contribution for cotton futures returns, the standardized returns should follow a normal distribution if the process governing the realized volatility is log-normal and the process governing returns is normal. According to Clark's vocable, the volatility process is the "*directing process*", and the distribution of standardized returns is said to be "*subordinated*" to the distribution of returns. The resulting process is thus a lognormal-normal mixture, so-called the "*mixture-of-distribution hypothesis*" (MDH) in the literature[18].

Insert Figure 7 about here

For the ECX CO_2 emissions 2008 futures data, it is obvious that standardized returns are not normally distributed (see Figure 7). Table 3 indicates a sample skewness of 0.89 and a sample kurtosis of 8.84. Gaussianity is clearly rejected at all confidence levels, and does not need further investigation. As in Areal and Taylor (2002), the rejection of the MDH may be due to *(i)* the imperfect estimation of the logarithmic volatility through the realized estimator[19], and *(ii)* the extreme outlier occurring on October 13, 2008, which strongly deforms our distribution. Another explanation for non-normality may be found in Fleming and Paye (2005), who argue that microstructure noise biases kurtosis estimates for standardized returns. The intuition behind this result is that microstructure noise is less likely to occur for large absolute returns, because large absolute returns are often associated with larger volumes. As such, the realized volatility is underestimated for large absolute return days, thus inflating the tails of the standardized returns distribution. Because of the limited number of observations in the present work, it appears difficult to verify this assumption. This would necessitate many large absolute return days and a thorough analysis of the microstructure bias conditionally on the presence of a large absolute return.

[18] A very clear presentation of the MDH is given in Jondeau *et al.* (2007), sections 3.3 and 3.4. This hypothesis is investigated for futures returns in Areal and Taylor (2002) and Martens and Luu (2003), among others.

[19] Note we did not introduce the possibility of jumps in our analysis through more robust estimators as bipower or tripower estimators (see ABD (2007)). Indeed, the presence of jumps may distort the distribution of standardized returns. This area is left for further research.

The rejection of the MDH for the ECX CO_2 emissions 2008 futures contract has strong implications for derivatives pricing in these markets[20]. The jump-free diffusion process which is commonly assumed for option pricing does not seem suitable for the CO_2 emissions allowance market. There may be two different explanations for that. First, the process may include jumps. Options would then be better priced using jump-diffusion models. Second, the independence assumption between the Brownian motion and the volatility process may be violated. This also has some consequences for the pricing of derivatives, as more complex models need to be considered.

We also investigate graphically the presence of leverage, *i.e.* an increase in volatility following negative returns. Such an asymmetry may have consequences in terms of volatility modeling, because a good working knowledge of returns would help to model volatility.

Insert Figure 8 about here

By contrast, the absence of asymmetric effect seems apparent in Figure 8, which provides a scatterplot of realized volatility in logarithmic form against lagged standardized returns. This conclusion has, of course, to be taken with care in light of the limited number of daily observations in our study.

It is common in the financial literature to examine the parametric modeling of volatility through GARCH or stochastic volatility (SV) models. More precisely, GARCH volatilities may be used to standardize daily returns, and may be compared with realized volatility results. Following Benz and Truck (2009), we specify the AR(1)-GARCH(1,1) model:

$$R_t = \beta_0 + \beta_0 R_{t-1} + \epsilon_t \tag{7}$$

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 h_{t-1} \tag{8}$$

with R_t the daily returns, and ϵ_t the error term in equation (7). Equations (7) and (8) are estimated by Quasi Maximum Likelihood (QML) (Gourieroux *et al.* (1984)) using the BHHH algorithm (Berndt *et al.* (1974)).

Insert Table 4 about here

Estimation results of the AR(1)-GARCH(1,1) model are presented in Table 4. Residual tests for the chosen specification provide evidence that any autocorrelation in the residuals and squared residuals has been removed[21]. The distribution of GARCH-standardized returns is more normal than the distribution of realized volatility-standardized returns (see Table 3 and Figure 7). This result is unusual in the financial economics literature, as GARCH-standardized returns are generally more fat-tailed than realized volatility-standardized returns. The natural leptokurticity of GARCH models is

[20] European options with various strike prices have indeed been introduced in October 2006 on ECX (see Chevallier *et al.* (2009)).

[21] To conserve space, the autocorrelation and partial autocorrelation functions of the residuals and squared residuals are not reproduced here, and may be obtained upon request to the authors.

generally argued to be insufficient to accommodate the empirical excess kurtosis of financial time-series[22].

Insert Figure 9 about here

Figure 9 plots the time series of the AR(1)-GARCH(1,1) model. We observe that GARCH estimates are significantly smoother than realized estimates. In light of our empirical study, GARCH modeling appears more suitable to reach normality once returns have been standardized. This result highlights the critical role which may be played by jumps in the time-series of ECX CO_2 emissions 2008 futures.

In the next section, we investigate the properties of the conditional distribution of futures returns and realized volatility.

5 Modeling realized volatility dynamics

In this section, we are interested in modeling the conditional distribution of volatility. This investigation has practical applications for forecasting purposes, and may also be of interest for traders who need accurate volatility estimates for derivatives pricing.

We first investigate the autocovariance in the realized variance, the realized volatility, and the logarithm of volatility series.

Insert Figure 10 about here

Figure 10 plots the autocorrelation function (ACF) and partial autocorrelation function (PACF) estimated for the naive estimator[23]. We detect the presence of serial correlation for realized variance and realized volatility at least with one lag. For the log-transformation of the volatility series, the estimated autocorrelation does not appear to decay exponentially, but rather hyperbolically. This may be an indication of the presence of an unit root.

Insert Table 5 about here

The test statistics provided in the first column of Table 5 indicate the rejection of the unit-root hypothesis in all cases. In what follows, we focus on the existence of long memory in the data generating process.

Because the tick-by-tick time-series of ECX CO_2 emissions futures is very short to investigate the presence of long memory, we consider two estimation procedures for the fractional integration coefficient, as in ABDL (2001) and Areal and Taylor (2002). First, let S_T be the variance of the sum of

[22] Log-likelihood based on fat-tailed distributions (generalized error distribution (GED), Student, etc.) is commonly used to accommodate this high degree of kurtosis. We did not find however any improvement in our estimation by using a similar approach.

[23] Similar plots were obtained for the two other estimators, and thus are not reported here to conserve space.

T consecutive observations of, say, logarithm of the realized volatility $log(RV_t^{1/2})$. For long memory processes, the variances S_T follow a scaling law such that:

$$T^{-(2d+1)} S_T \to C \tag{9}$$

as $T \to \infty$ with $d > 0$, and C is a constant[24].

Insert Figure 11 about here

Figure 11 plots the sample variances S_T of the partial sums of the realized logarithmic standard deviations against the logarithm of the aggregation level for T. The regression coefficient corresponds to $2d+1$, and thus leads to an implicit value of the fractional integration coefficient reported in Table 5.

The second methodology to estimate this coefficient is the Geweke-Porter-Hudak's (henceforth GPH, 1983) method (see Brockwell and Davis (1991) for a formal presentation, or Corsi (2009) for a discussion). The GPH estimate is based on the regression of the logarithm of the periodogram estimate of the spectral density against $ln(\omega)$ over a range of frequencies ω with:

$$w^{2d} f(\omega) \to C \tag{10}$$

as $T\omega \to 0$ and C a constant. Again, the estimates are comprised in the range of $[0, 0.5]$, which indicates the presence of long memory.

In view of these strong indications of long memory in the log time-series, we choose to rely on Corsi's (2009) parsimonious HAR-RV model for at three main reasons. First, recall that our dataset contains only 240 trading days. This is clearly too few for ARFIMA modeling, despite the presence of long memory[25]. Second, Pong *et al.* (2008) show that long memory may not be distinguished from short memory below 250 trading days. Second, the HAR-RV model succeeds in reproducing the long memory features of the time-series, while being easier to estimate particularly on a shorter time-horizon. Third, the heterogeneous behavior assumed between economic agents may be justified by the fact that traders, utilities and financial institutions operating on the EU ETS have different investment horizons. The HAR-RV model is used in ABD (2007), Corsi *et al.* (2008), and Liu and Maheu (2009) among others. The economic intuition behind this model is that different groups of investors have different investment horizons, and consequently behave differently (see Müller *et al.* (1997) for the presentation of the HARCH original model relying on the Heterogeneous Hypothesis).

The original HAR-RV model by Corsi (2009) is formally a constrained AR(22) model, slightly different from ABDL (2001) and Corsi *et al.* (2008)[26]. The HAR-RV model using daily, weekly and monthly

[24] In comparison, setting $d = 0$ is a feature of short memory.

[25] Note that ARFIMA estimation does not appear suitable alternatives for the one-year ECX emissions futures with tick-by-tick data, since the estimation of formal long memory models would require several years of data.

[26] ABDL (2001) formally use an AR(5). In this article, we adopt an intermediate specification by selecting a simplified HAR-RV model with only a weekly component, thus leading to a constrained AR(5) specification. Note that our choice is also econometrically motivated by the $Q(20)$ test statistics reported in Table 3.

realized-volatility components may be defined as follows:

$$\sqrt{RV_t} = \alpha_0 + \alpha_d \sqrt{RV_{t-1}} + \alpha_w (\sqrt{RV})_{t-5:t-1} + \alpha_m (\sqrt{RV})_{t-22:t-1} + u_t \tag{11}$$

or in logarithmic form:

$$\log RV_t = \alpha_0 + \alpha_d \log RV_{t-1} + \alpha_w (\log RV)_{t-5:t-1} + \alpha_m (\log RV)_{t-22:t-1} + u_t \tag{12}$$

Following ABD (2007), the HAR-RV model for forecasting with the horizon h may be defined in general form by using the multiperiod realized variation (the sum of the corresponding one-period measures):

$$RV_{t,t+h} = h^{-1}[RV_{t+1} + RV_{t+2} + \ldots + RV_{t+h}] \tag{13}$$

and by definition, $RV_{t,t+1} \equiv RV_{t+1}$. The HAR-RV model proposed by Corsi (2009) is a specific case of equation (13) for which $h = 1$, thereby assuming that traders have investment horizons corresponding to one-day ahead, one-week ahead, and one-month ahead forecasts.

As demonstrated below, the ECX CO_2 emissions 2008 futures contract only requires a weekly component, thus simplifying Corsi's initial model. For each estimator and for RV, $RV^{1/2}$ and $log(RV^{1/2})$, we estimate the following specification:

$$RV_{t,t+h} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{t-5,t} + u_t \tag{14}$$

Insert Table 6 about here

Insert Table 7 about here

Insert Table 8 about here

Estimates are reported in Tables 6 to 8. From Table 6, we may observe that the HAR-RV model performs poorly in fitting the daily realized variance, as shown by the low R^2 from 0.0003 (regression (9)) to 0.0109 (regression (1)). These results are in line with previous literature on realized volatility, where the "raw" realized variance is difficult to model. The results displayed in Table 7 show the same pattern for the daily realized volatility, where the values obtained for the R^2 range from 0.0653 (regression (8)) to 0.1211 (regression (1)). This improvement from realized variance to realized volatility is common in other empirical studies (see for instance ABDL (2001, 2003)). The best results are generally achieved using the logarithmic transformation. Table 9 shows indeed a dramatic improvement in the results obtained. The R^2 values obtained for the daily realized volatility in logarithmic form range from 0.2798 (regression (2)) to 0.3691 (regression (4)). These values are comparable to ABD (2007) for FX markets and S&P 500 futures. We may conclude that the fit of the HAR-RV model for the

log-series of the ECX CO_2 emissions 2008 futures data is much better than the fit for realized variance or realized volatility. The dramatic improvement in the fit of realized volatility models when using the log-transformation is well documented in the literature (see ABDL (2001, 2003), ABD (2007), and Corsi (2009) among others). A better in-sample fit leads to a better out-of-sample forecasting accuracy.

In the next section, we provide a forecasting exercice using the HAR-RV model versus the GARCH specification.

6 Forecasting

In this section, we use Mincer-Zarnowitz regression techniques, as in ABD (2003, 2005), to investigate the forecasting power of our competing models[27]. To compare the forecasting accuracy of the HAR-RV model versus the GARCH model estimated in the previous section, we run the following regressions:

$$(v_{t+1}) = b_0 + b_1(v_{t+1|t,HAR-RV}) + b_2(v_{t+1|t,GARCH}) + u_{t+1} \tag{15}$$

$$(v_{t+1})^{1/2} = b_0 + b_1(v_{t+1|t,HAR-RV})^{1/2} + b_2(v_{t+1|t,GARCH})^{1/2} + u_{t+1} \tag{16}$$

$$\log(v_{t+1})^{1/2} = b_0 + b_1 \log(v_{t+1|t,HAR-RV})^{1/2} + b_2 \log(v_{t+1|t,GARCH})^{1/2} + u_{t+1} \tag{17}$$

Due to the limited historical dataset for ECX CO_2 emissions futures, we only consider *one-step-ahead* forecasts[28]. The HAR-RV model is estimated with a daily and a weekly component for the three estimators.

Insert Figure 12 about here

The corresponding forecasts for the daily realized variance, the daily realized volatility, and the daily realized volatility in logarithmic form versus actual observations are displayed in Figure 12[29].

If the forecasting properties of the HAR-RV model are satisfactory, the b_0 coefficient should be equal to zero, the b_1 coefficient should be equal to one, and the introduction of an alternative model (here a GARCH model) through the coefficient b_2 should not increase significantly the R^2 of the regression. Thus, we are especially interested in the stability of the b_0 and b_1 coefficients, as well as in the increase of the R^2 between models. The b_2 coefficient depends on the scaling of the different variables, and thus is subject to a wide variability.

Insert Table 9 about here

[27] These are also known as "encompassing regressions".

[28] *i.e.* at each period t we use the data observed until $t-1$, and base our forecasts on the parameters of the model estimated over the period $[0, t-1]$. The first forecast is made using 100 observations, the second forecast 101 observations, and so on.

[29] Note that contrary to Figure 2, we decided to keep in our forecasting exercise the outlier on October 13, 2008, possibly due to the "credit crunch" effect as discussed in Section 3.2.

The main results of our forecasting exercise are presented in Table 9. The model which provides the best results is the logarithmic model. This result is not surprising, since the logarithmic model estimates were characterized by the highest values for the R^2 in Table 8. Our results confirm the robustness of the HAR-RV model. As predicted, we observe that the b_0 coefficients are close to zero, while the b_1 coefficients are close to one in all regressions (RV_t, $RV_t^{1/2}$, $\log(RV_t^{1/2})$). Besides, the GARCH estimates do not seem to improve significantly the R^2 of the regressions, especially in the case of RV_t. For $RV_t^{1/2}$ and $\log(RV_t^{1/2})$, we only observe a slight increase of the R^2, but the GARCH coefficient is only significant at the 10% level for the log-series. This property of GARCH models is widely documented in previous literature. Indeed, GARCH forecasts track much better the broad temporal movements in the volatilities for lower frequency variations, and their accuracy tends to perform poorly at higher frequencies.

Accordingly, our forecasting results do not seem to indicate that the mixture of the HAR-RV and GARCH models improves significantly the forecast accuracy of our estimates. For all regressions, the b_1 coefficients are lower than one, and the values of the R^2 do not seem significantly higher.

Overall, we demonstrate in this section the accuracy of the HAR-RV model, as well as the inaccuracy of GARCH forecasts and their inability to adapt to high-frequency movements. As noted in ABDL (2003)[30], this is due to the superiority of realized measures in estimation. As such, superior estimates of present conditions translate into superior forecasts of the future[31].

7 Conclusion

This article constitutes the first attempt to use realized measures of volatility for a specific energy commodity, namely the ECX CO_2 emissions futures contract of maturity December 2008. We proceed as is standard in the realized volatility literature to assess the distributional and dynamic properties of realized volatility for this contract. Besides, this article constitutes one of the first attempts to analyze the properties of CO_2 prices in the EU ETS using high-frequency data.

Our main findings may be summarized as follows: *(1)* the unconditional distribution of daily returns are near normal; *(2)* any attempt to standardize these returns using realized measures and to a lesser extent GARCH estimates does not lead the distribution to Gaussianity; *(3)* we thereby strongly reject the *mixture-of-distribution-hypothesis* developed by Clark (1973) and Tauchen and Pitts (1983); *(4)* the dynamics of realized volatility is well captured using the HAR-RV model with a daily and a weekly component, which outperforms significantly the GARCH specification; and (5) the predictive accuracy of the HAR-RV model outperforms unambiguously other models of conditional volatility based on daily data.

This work may be extended in several directions. First, the ECX CO_2 emissions futures tick-by-tick

[30] "*We have identified the quadratic variation and its empirical counterpart, the realized volatility, as the key objects of interest for volatility measurement, and we consequently assess our various volatility forecasts relative to this measure. It is perhaps not surprising that models built directly for the realized volatility produce forecasts superior to those obtained from less direct methods*, [...]" (ABDL, 2003, p. 613).

[31] Note the forecasts presented here only constitute a statistical metrics, and not an economic metrics such as the value of CO_2 allowances used for option pricing or portfolio management.

data set considered here only covers one-year with about 240 trading days and 700 transactions per day, thereby multiplying parameter and model uncertainties. These uncertainties could be reduced using bootstrap methods as developed very recently in Gonçalves and Meddahi (2009). These authors mainly resort to the wild bootstrap method to increase the number of available intraday data each day, without suffering from the so-called "*microstructure-noise*" bias.

Second, the inclusion of jumps within realized volatility measures appears necessary to fit the characteristics of CO_2 futures highlighted in previous literature. Daskalakis *et al.* (2009) use a jump-diffusion model to approximate the random behavior of CO_2 prices. Benz and Truck (2009) analyze the spot price behavior with a Markov-switching model. Lin and Lin (2007) model CO_2 prices as a result of mean-reversion with varying trends, combined with state-dependent price jumps and volatility structure, and show that mean-reversion fares better in forecasting futures prices.

Third, the use of realized volatility for ECX CO_2 emissions futures contracts may be useful for option pricing (see Stentoft (2008) for a first application to option stock markets) with a high-frequency measure of volatility. This may be of great help on such an emerging commodity market, as on the EU ETS any attempt to price derivatives is subject to strong uncertainties.

Fourth, the "*maturity effect*" encountered when selecting the sampling frequency here may be checked on other markets for more robust conclusions, and statistical tests may be used to determine the optimal sampling frequency. Indeed, if realized volatility is significantly different at different moments in the life of a futures contract, hedge ratios should be modified accordingly.

References

Aït-Sahalia, Y., Mykland, P., Zhang, L., 2005. How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies* 18, 351-416.

Alberola, E., Chevallier, J., 2009. European carbon prices and banking restrictions: Evidence from Phase I (2005-2007). *The Energy Journal* 30, 107-136.

Alberola, E., Chevallier, J., Chèze, B., 2008. Price drivers and structural breaks in European carbon prices 2005-2007. *Energy Policy* 36, 787-797.

Andersen, T.G., Benzoni, L., 2009. Realized volatility. In: Andersen, T.G., Davis, R.A., Kreiß, J.-P., Mikosch, Th. (Eds.) *Handbook of Financial Time Series*, Springer.

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

Andersen, T.G., Bollerslev, T., 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., Ebens, H., 2001. The distribution of stock return volatility. *Journal of Financial Economics* 61, 43-76.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001. The distribution of exchange rate volatility. *Journal of the American Statistical Association* 96, 42-55.

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

Andersen, T.G., Bollerslev, T., Meddahi, N.,, 2005. Correcting the errors: volatility forecast evaluation using high-frequency data and realized volatilities. *Econometrica* 73, 279-296.

Andreou E, Ghysels E., 2002. Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics* 17, 579-600.

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

Awartani, B.M.A., Corradi, V., Distaso, W., 2009. Testing market microstructure effect with an application to the Dow Jones Industrial Average stocks. *Journal of Business and Economic Statistics*, forthcoming.

Barndorff-Nielsen, O., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B* 64, 253-280.

Benz, E., Klar, J., 2008. Liquidity and price discovery in the European CO_2 futures market: an intraday analysis. Working Paper, Bonn Graduate School of Business.

Benz, E., Truck, S., 2009. Modeling the Price Dynamics of CO_2 Emission Allowances. *Energy Economics* 1, 4-15.

Berndt, E.K., Robert, E.,Hall, B.H., Hausman, J.A., 1974. Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement* 3, 653–665.

Brockwell, P.J., Davis, R.A., 1991. *Time Series: Theory and Methods*. Springer Series in Statistics, Springer-Verlag, New York.

Busch, T., Hoffman, V.H., 2007. Emerging carbon constraints for corporate risk management. *Ecological Economics* 62, 518-528.

Bunn, D., Fezzi, C., 2007. Interaction of European Carbon Trading and Energy Prices. *Fondazione Eni Enrico Mattei* Working Paper 123.

Cai, J., Cheung, Y.L., Lee, R.S.K., Melvin, M., 2001. "Once-in-a-generation" yen volatility in 1998: fundamentals, intervention, and order flow. *Journal of International Money and Finance* 20, 327-347.

Cartea, Á., Börger, R.H., Kiesel, R., Schindlmayr, G., 2007. A multivariate commodity analysis and applications to risk management. Birkbeck Working Papers in Economics and Finance BWPEF 0709, Birkbeck, University of London.

Chen, W.W., Deo, R.S., 2004. Power transformations to induce normality and their applications. *Journal of the Royal Statistical Association B* 66, 117-130.

Chevallier, J., Ielpo, F., Mercier, L., 2009. Risk aversion and institutional information disclosure on the European carbon market: a case-study of the 2006 compliance event. *Energy Policy* 37, 15-28.

Clark, P.K., 1973. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* 41, 135-156.

Christiansen, A.C., Arvanitakis, A., Tangen, K., Hasselknippe, H., 2005. Price determinants in the EU emissions trading scheme. *Climate Policy* 5, 15-30.

Corsi, F., 2009. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics* 7,

174-196.

Corsi, F., Mittnik, S., Pigorsch, C., Pigorsch, U., 2008. The volatility of realized volatility. *Econometric Reviews* 27, 46-78.

Dacorogna, M.M., Gençay, R., Müller, U.A., Olsen, R.B., Pictet, O.V., 2001. *An Introduction to High-Frequency Finance*. Academic Press, San Diego, CA.

Daskalakis, G., Psychoyios, D., Markellos, R.N., 2009. Modeling CO_2 emission allowance prices and derivatives: evidence from the European trading scheme. *Journal of Banking and Finance*, forthcoming.

Duong, H.N., Kalev, P.S., 2008. The Samuelson Hypothesis in futures markets: an analysis using intraday data. *Journal of Banking and Finance* 32, 489-500.

Fleming, J., Paye, B.S., 2005. High-frequency returns, jumps and the mixture of normals hypothesis. Unpublished manuscript.

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

Gonçalves, S., Meddahi, N., 2008. Box-Cox transforms for realized volatility. Unpublished manuscript.

Gonçalves, S., Meddahi, N., 2009. Bootstrapping realized volatility. *Econometrica* 77, 283-306.

Gourieroux, C., Monfort, A., Trognon, A., 1984. Pseudo Maximum Likelihood Methods: Theory. *Econometrica* 52, 680–700.

Guillaume, D.M., Dacorogna, M.M., Davé, R.R., Müller, U.A., Olsen, R.B., Pictet, O.V., 1997. From the bird's eye to the microscope: a survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics* 1, 95-129.

Hansen, P.R., Lunde, A., 2006. Realized variance and market microstructure noise. *Journal of Business and Economic Statistics* 24, 127-218.

Illueca, M., Lafuente, J.A., 2006. new evidence on expiration-day effects using realized volatility: an intraday analysis for the Spanish stock exchange. *Journal of Futures Markets* 26, 923-938.

Jondeau, E., Poon, S.-H., Rockinger, M., 2007. *Financial Modeling Under Non-Gaussian Distributions*. Springer-Verlag, Springer.

Liu, C., Maheu, J.M., 2008. Are there structural breaks in realized volatility? *Journal of Financial Econometrics* 6, 326-360.

Lin, Y.N., Lin, A.Y., 2007. Pricing the Cost of Carbon Dioxide Emission Allowance Futures. *Review of Futures Markets* 16, 1-16.

Luu, C.J., Martens, M., 2003. Testing the mixture-of-distributions hypothesis using "realized" volatility. *Journal of Futures Markets* 23, 661-679.

McAleer, M., Medeiros, M.C., 2008. Realized volatility: a review. *Econometric Reviews* 27, 10-45.

Müller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V., 1997. Volatilities of different time resolutions – Analyzing the dynamics of market components. *Journal of Empirical Finance* 4, 213-239.

Oberndorfer, U., 2009. EU emission allowances and the stock market: evidence from the electricity industry. *Ecological Economics* 68, 1116-1126.

Paolella, M.S., Taschini, L., 2008. An Econometric Analysis of Emission Trading Allowances. *Journal of Banking and Finance* 32, 2022-2032.

Pong, S., Shackleton, M.B., Taylor, S.J., 2008. Distinguishing short and long memory volatility specifications. *Econometrics Journal* 11, 617-637.

Samuelson, P.A., 1965. Proof that Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review* 6, 41-49.

Stentoft, L., 2008. Option pricing using realized volatility. CREATES Research Paper 2008-13, Aarhus Business School.

Tauchen, G.E., Pitts, M., 1983. The price variability-volume relationship on speculative markets. *Econometrica* 51, 485-505.

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

Tucker, M., 2001. Trading carbon tradable offsets under Kyoto's clean development mechanism: the economic advantages to buyers and sellers of using call options. *Ecological Economics* 37, 173-182.

Zhang, L., Mykland, P.A., 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.

Zhou, B., 1996. High frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics* 14, 45-52.

Zivot, E., 2005. Analysis of High Frequency Financial Data: Models, Methods and Software. University of Washington, Unpublished Manuscript.

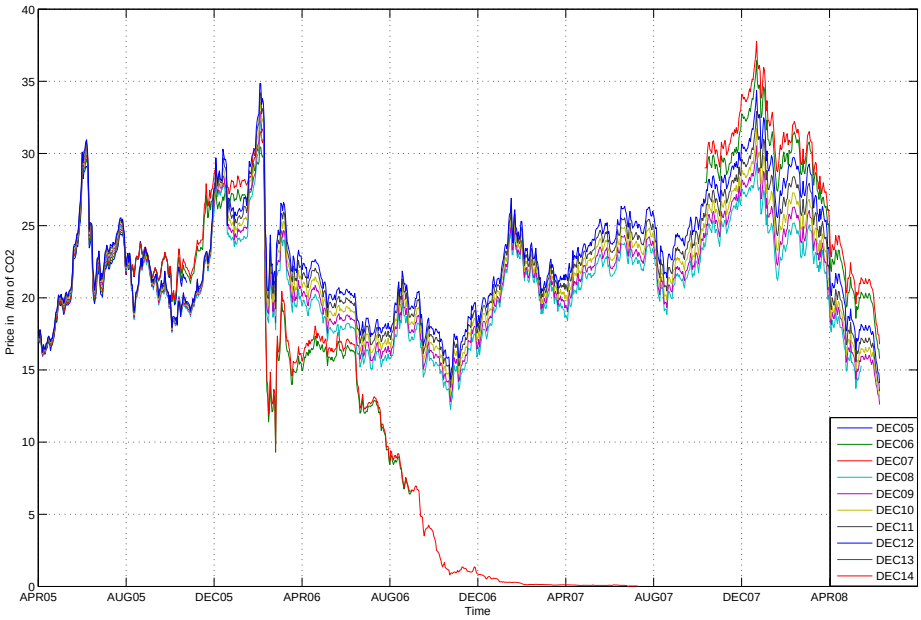

Figure 1: CO_2 futures prices of maturities December 2005 through 2014 from April 22, 2005 to January 16, 2009
Source: ECX

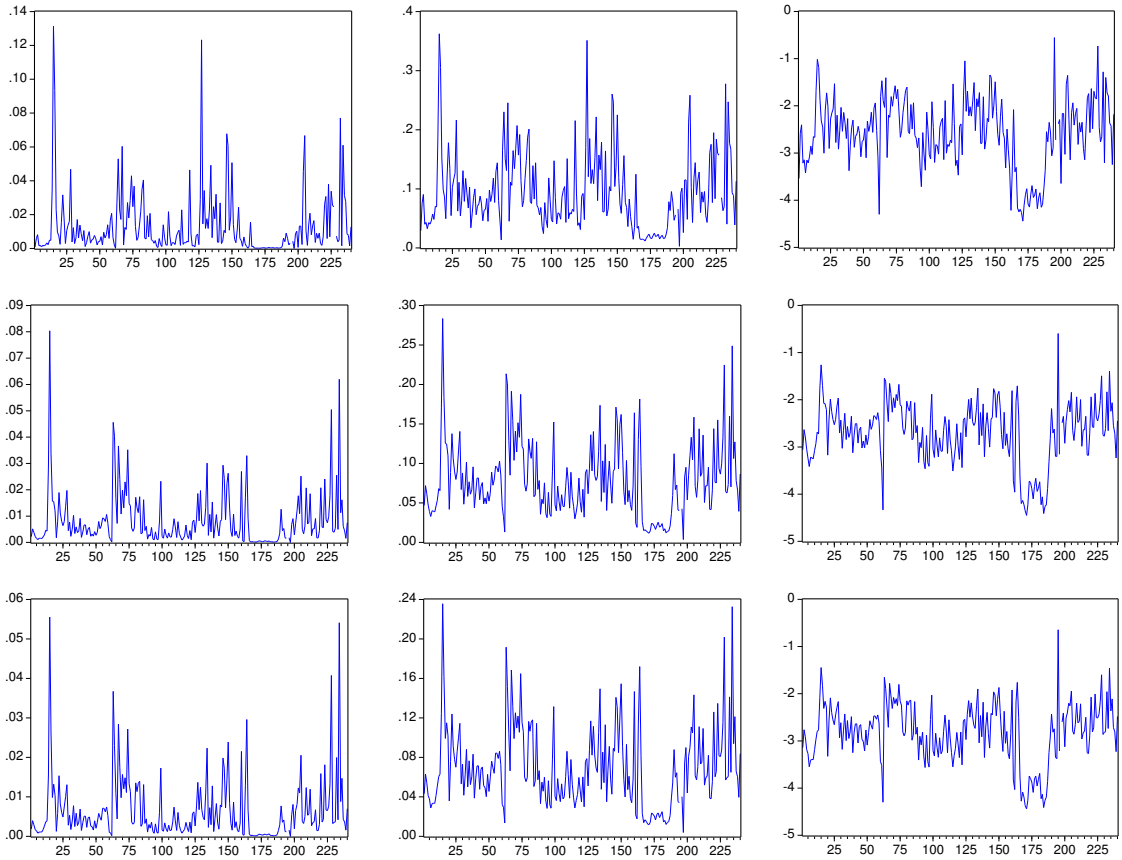

Figure 2: Daily realized variance (RV_t, left panel), daily realized volatility in standard deviation form ($RV_t^{1/2}$, middle panel), and daily realized volatility in logarithmic form ($log(RV_t^{1/2})$, right panel) for the three estimators (naive on the first row, Zhang *et al.* (2005) sub-sampling estimator on the second row, and Bartlett kernel-based estimator on the third row).

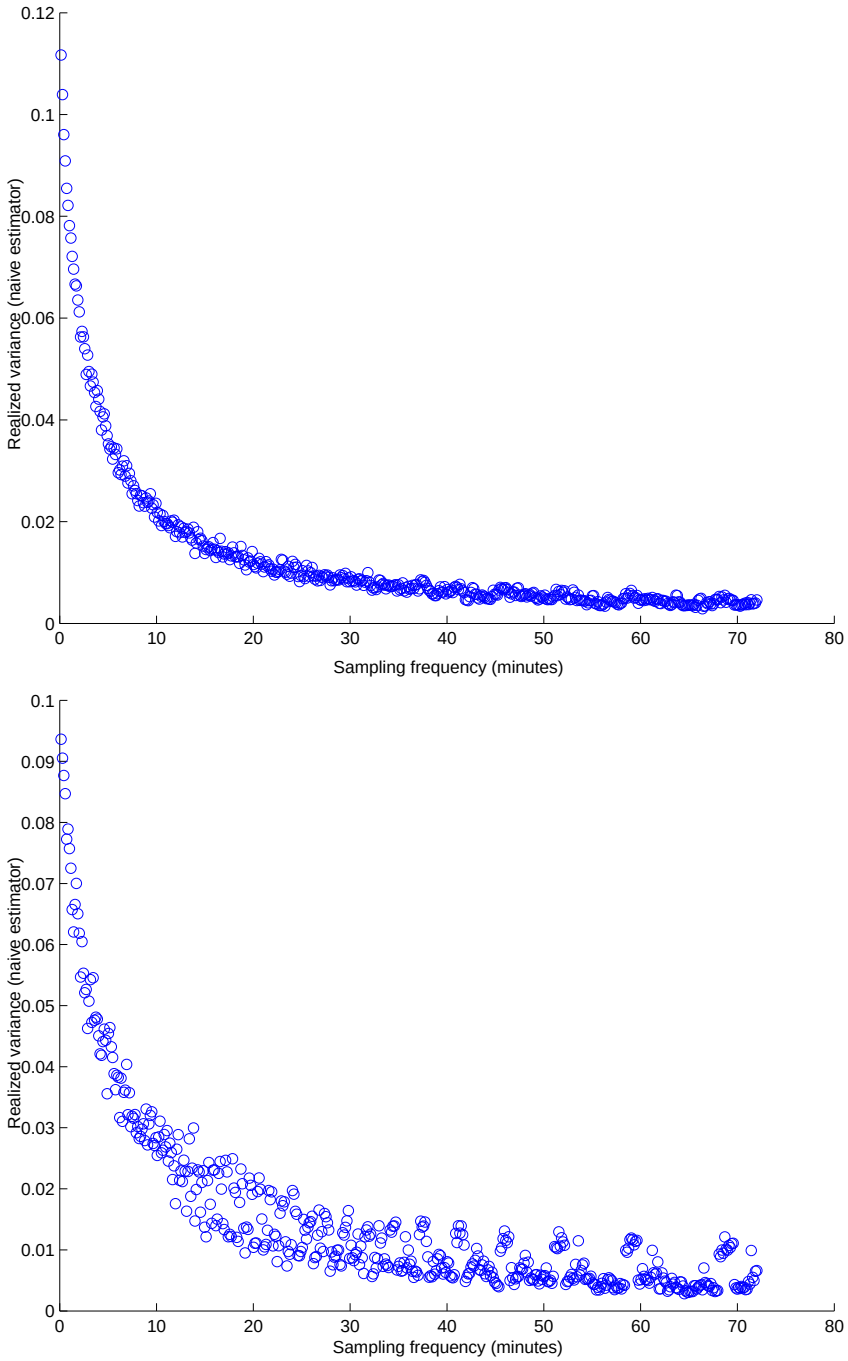

Figure 3: Volatility signature plots for the full (top) and November-December (bottom) samples using the naive estimator for realized variance.

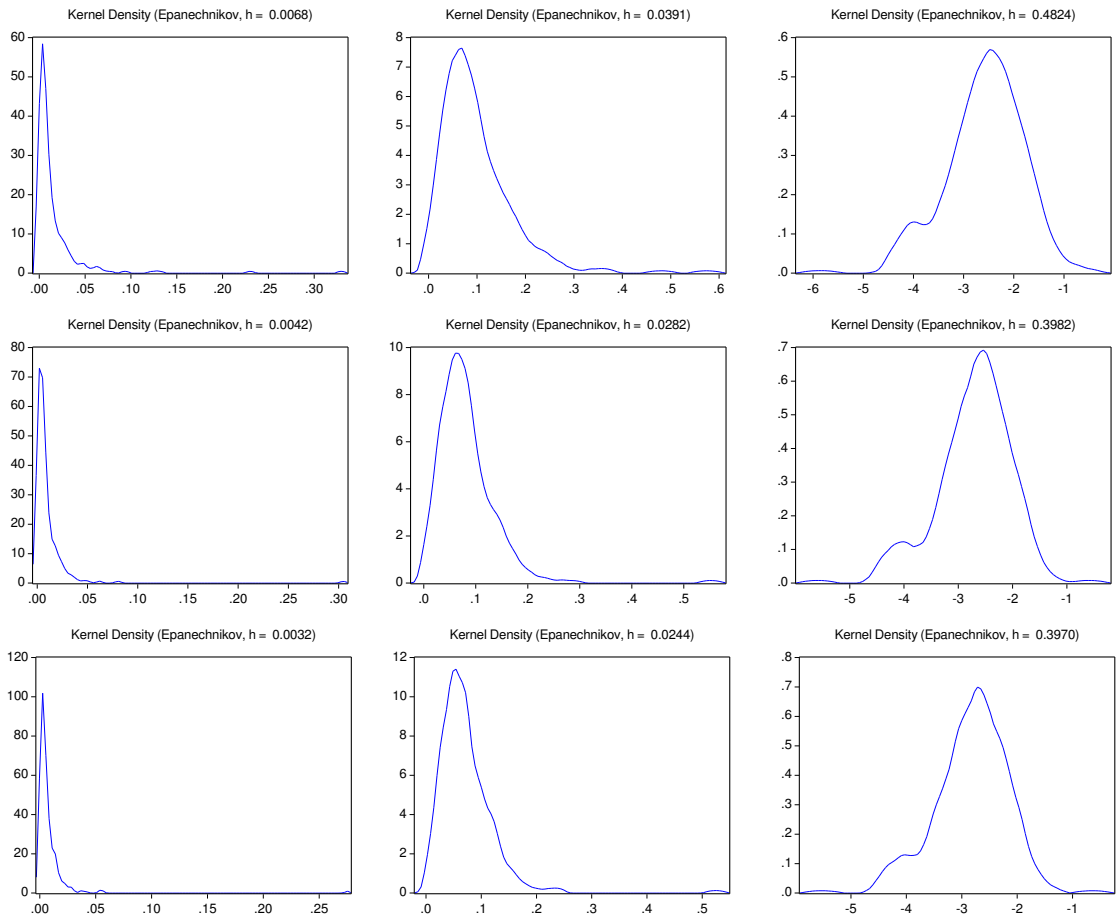

Figure 4: Centered kernel density estimates of the unconditional distribution for the daily realized variance (RV_t, left panel), the daily realized volatility in standard deviation form ($RV_t^{1/2}$, middle panel), and the daily realized volatility in logarithmic form ($log(RV_t^{1/2})$, right panel) based on 15-minute returns. The first row is for the naive estimator, the second row is for the Zhang *et al.* (2005) sub-sampling estimator, and the third row is for the Bartlett kernel-based estimator.

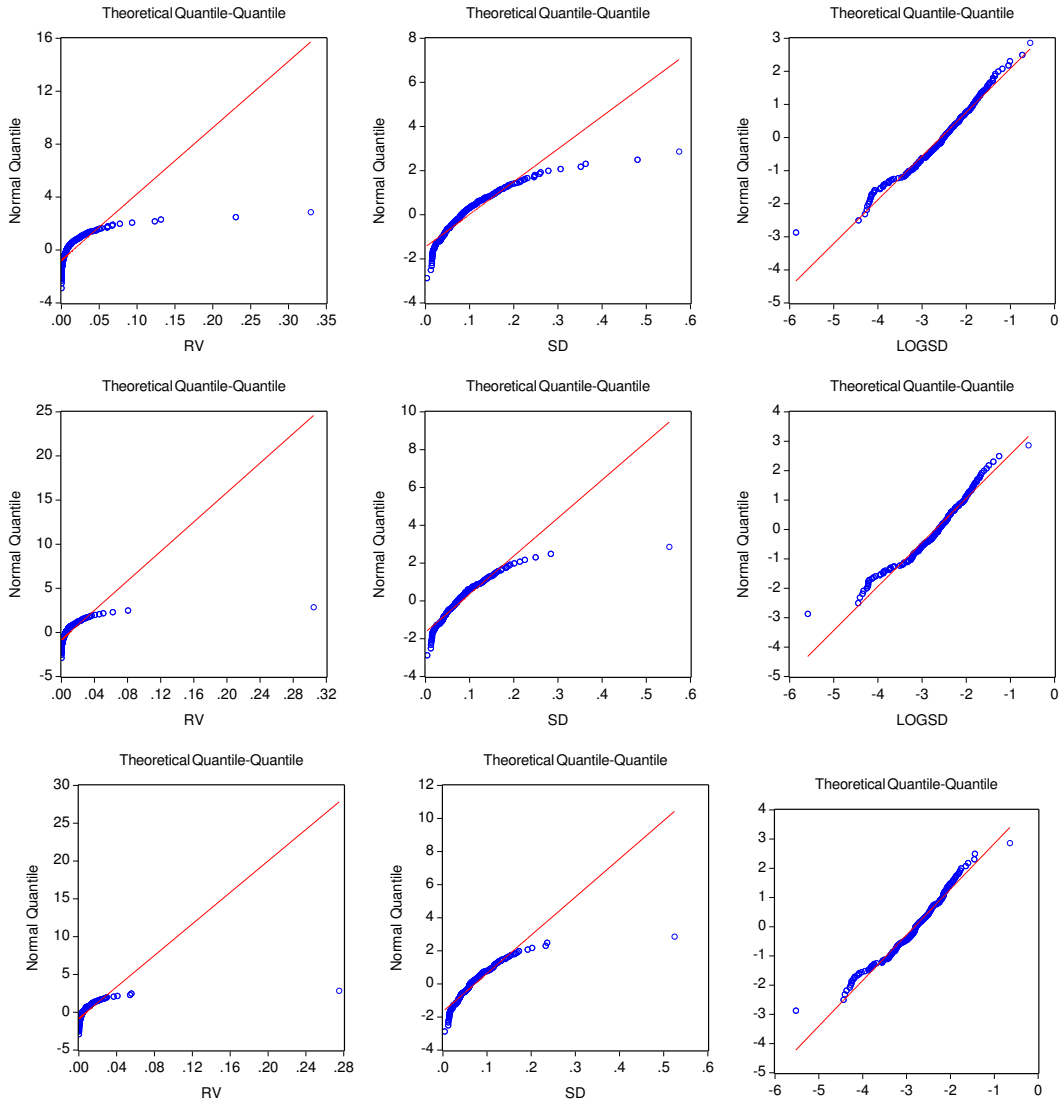

Figure 5: QQ plots for the realized variance (left panel), realized standard deviation (middle panel) and log of the standard deviation (right panel) for the three estimators (naive on the first row, Zhang *et al.* (2005) sub-sampling estimator on the second row, and Bartlett kernel-based estimator on the third row).

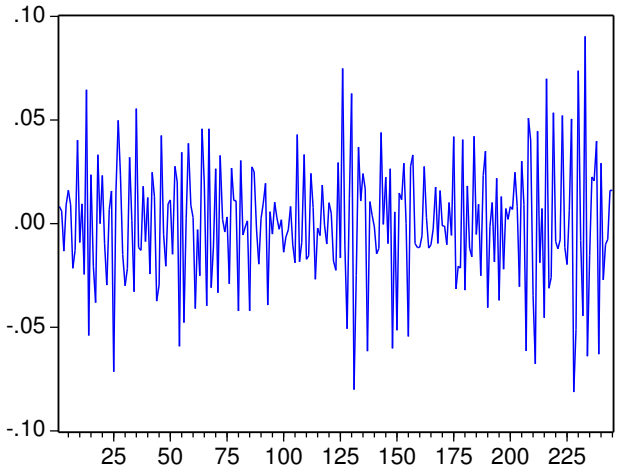

Figure 6: Time-series of daily raw returns.

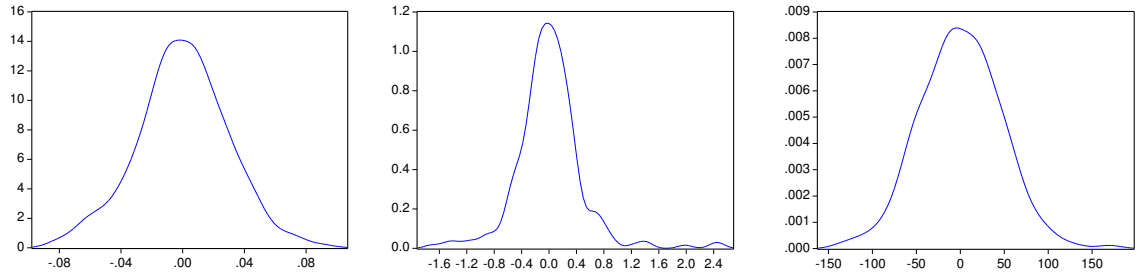

Figure 7: Smoothed Gaussian kernel distribution of daily returns (left panel), realized volatility (naive estimator) standardized returns (middle panel) and GARCH standardized returns (right panel).

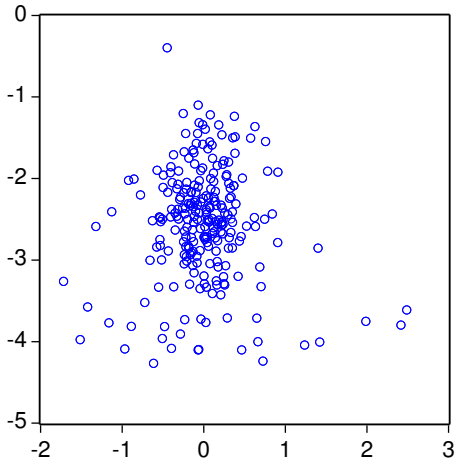

Figure 8: Scatterplot of the logarithmic realized volatility against lagged standardized returns.

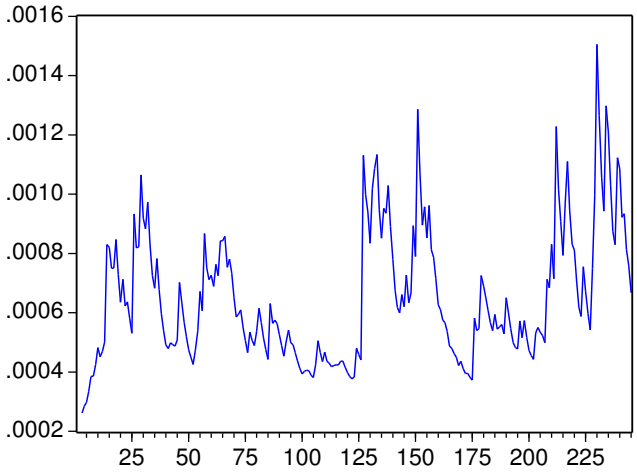

Figure 9: Time series of GARCH volatility measure.

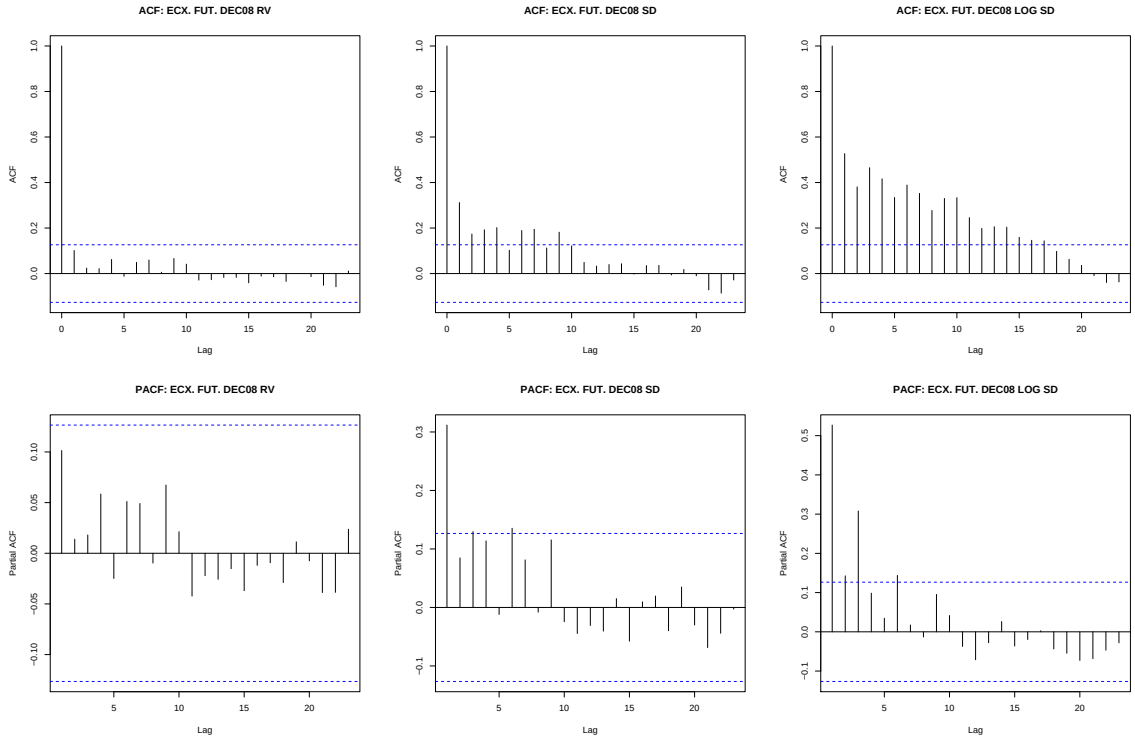

Figure 10: Autocorrelation and partial autocorrelation functions of the daily realized variance (RV_t, left panel), daily realized volatility in standard deviation form ($RV_t^{1/2}$, middle panel), and daily realized volatility in logarithmic form ($log(RV_t^{1/2})$, right panel) for the naive estimator.

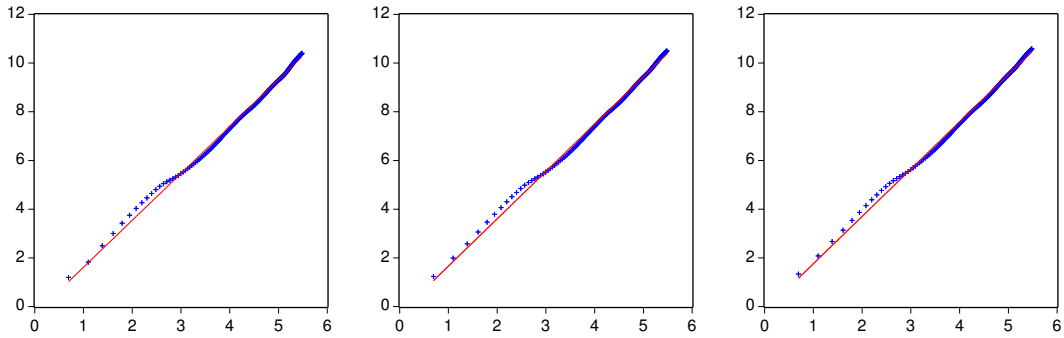

Figure 11: Scaling plot of the sample variances S_T of the partial sums of the realized logarithmic standard deviations against the logarithm of the aggregation level.

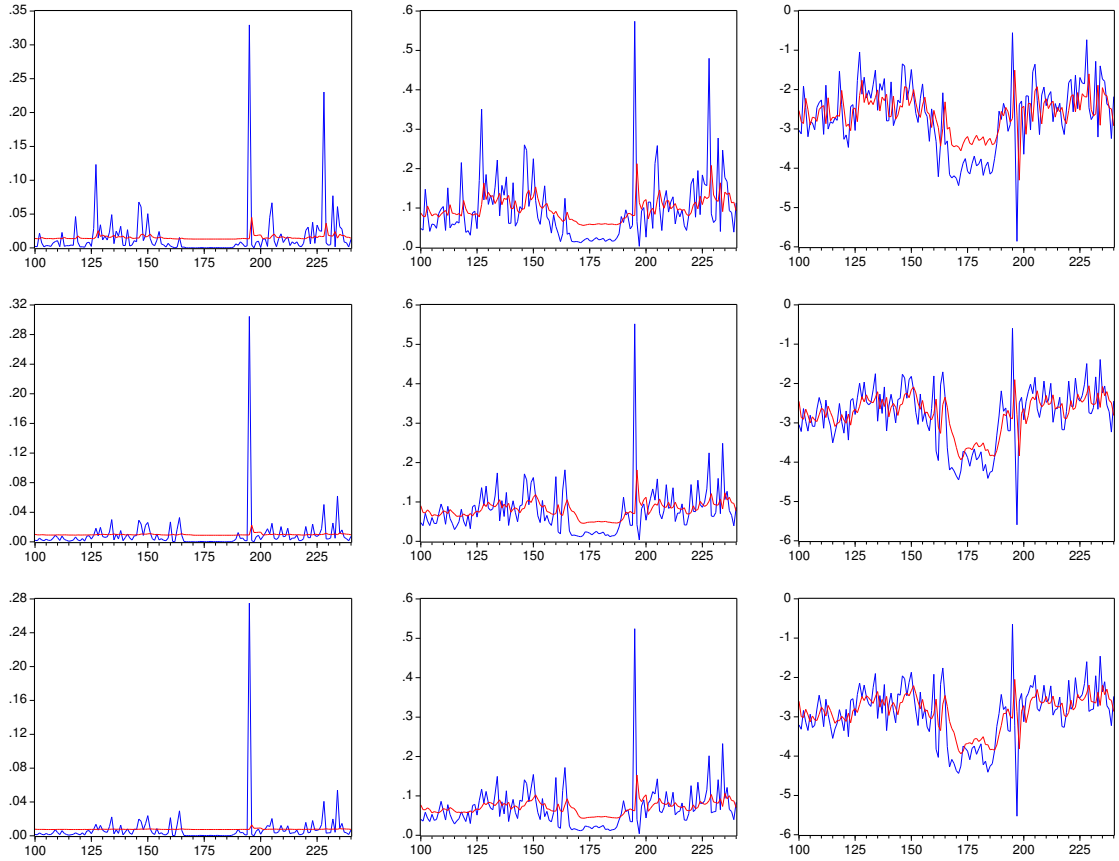

Figure 12: Forecasting of the daily realized variance (RV_t, left panel), the daily realized volatility in standard deviation form ($RV_t^{1/2}$, middle panel), and the daily realized volatility in logarithmic form ($log(RV_t^{1/2})$, right panel) with the HAR-RV model for the three estimators (naive on the first row, Zhang *et al.* (2005) sub-sampling estimator on the second row, and Bartlett kernel-based estimator on the third row).

	Mean	SD	Median	Min	Max	Skewness	Kurtosis	Ljung-Box (20)
Naive estimator								
RV_t	0.0130	0.0184	0.0064	0.0001	0.1313	3.2097	16.6500	7.1942
$RV_t^{1/2}$	0.0948	0.0636	0.0800	0.0029	0.0636	1.2998	5.1144	82.886
$log(RV_t^{1/2})$	-2.5652	0.7369	-2.5218	-4.4419	-0.5553	-0.3279	2.9927	420.63
Zhang *et al.* (2005) subsampling estimator								
RV_t	0.0085	0.0104	0.0052	0.0001	0.0804	2.9814	15.7588	3.3318
$RV_t^{1/2}$	0.0798	0.0467	0.0718	0.0037	0.2835	1.1008	4.7061	74.181
$log(RV_t^{1/2})$	-2.6966	0.6551	-2.6313	-4.4489	-0.5944	-0.4657	3.4233	376.51
Bartlett kernel-based estimator								
RV_t	0.0065	0.0079	0.0037	0.0001	0.0555	3.0012	15.4313	2.2043
$RV_t^{1/2}$	0.0702	0.0403	0.0609	0.0040	0.2356	1.1365	4.8489	59.803
$log(RV_t^{1/2})$	-2.8119	0.6264	-2.7932	-4.4386	-0.6454	-0.3229	3.3850	334.17

Table 1: Descriptive statistics of the daily realized variance (RV_t), daily realized volatility ($RV_t^{1/2}$), and daily realized volatility in logarithmic form ($log(RV_t^{1/2})$) for the naive, subsampling, and kernel-based estimators.

Note: The number of trading days is 240. SD stands for standard deviation. Ljung-Box test statistics are computed for a maximum number of 20 lags.

	Lilliefors	Crámer-Von Mises	Jarque-Bera	Watson	Anderson-Darling
Naive estimator					
$RV^{1/2}$	0.127955	1.350926	1080.817	1.089644	8.055166
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$log(RV^{1/2})$	0.062920	0.205757	22.16161	0.164762	1.347298
	(0.0522)	(0.0045)	(0.000015)	(0.0095)	(0.0017)
Zhang *et al.* (2005) subsampling estimator					
$RV^{1/2}$	0.128204	1.047318	4607.472	0.870926	9.000989
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$log(RV^{1/2})$	0.079671	0.353115	34.84085	0.286651	2.259627
	(0.0036)	(0.0001)	(0.0000)	(0.0002)	(0.0000)
Bartlett kernel-based estimator					
$RV^{1/2}$	0.120181	1.171580	8198.267	0.994852	9.903061
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$log(RV^{1/2})$	0.075590	0.264016	25.42408	0.219758	1.671065
	(0.0013)	(0.0009)	(0.0000)	(0.0016)	(0.0003)

Table 2: Normality test statistics for the realized standard deviation and logarithmic transformation with the three estimators.

Note: The values reported in parentheses are the p-values.

	Mean	SD	Skewness	Kurtosis	Jarque-Bera	$Q(20)$	$Q^2(20)$
Daily returns R_t	0.0000337	0.029600	-0.047258	3.242590	0.691953	75.609	51.660
RV-standardized daily returns	0.001904	0.498409	0.893659	8.846009	381.4887	66.923	152.95
GARCH-standardized daily returns	0.3078	46.3145	0.1034	3.4476	2.4622	72.154	19.500

Table 3: Descriptive statistics of continuously compounded daily returns, realized volatility (naive estimator) standardized returns, and GARCH standardized daily returns.

Note: The number of trading days is 240. SD stands for standard deviation, $Q(20)$ and $Q^2(20)$ stand for the Ljung-Box Q test statistics and the Ljung-Box $Q^2(20)$ test statistic computed up to 20 lags for returns and squared returns, respectively.

	Daily returns
Mean equation	
β_0	0.000045
	(0.0015)
β_1	-0.3881***
	(0.0677)
Variance equation	
α_0	0.0000945
	(0.0000668)
α_1	0.1839**
	(0.0945)
α_2	0.6973***
	(0.1572)
R^2	0.1300
Adj. R^2	0.1155

Table 4: AR(1)-GARCH(1,1) model estimates for daily returns

Note: The dependent variable is the daily return. Robust standard errors in parenthesis. *** indicates significance at 1%, ** at 5% and * at 10% levels.

	ADF test	$d(GPH)$	$\hat{d}$ from regression
Naive estimator			
RV_t	-13.9122	0.4376	–
$RV_t^{1/2}$	-11.1715	0.3318	–
$log(RV_t^{1/2})$	-4.2934	0.6849	0.4634
Zhang *et al.* (2005) subsampling estimator			
RV_t	-14.6932	0.4399	–
$RV_t^{1/2}$	-11.3561	0.3247	–
$log(RV_t^{1/2})$	-4.4725	0.6964	0.4588
Bartlett kernel-based estimator			
RV_t	-15.0757	0.4306	–
$RV_t^{1/2}$	-11.8635	0.3066	–
$log(RV_t^{1/2})$	-3.7696	0.6520	0.4711

Table 5: ADF test statistics up to 14 lags, $d(GPH)$ Geweke-Porter-Hudak estimates of the fractional integration parameter, and $\hat{d}$ coefficients estimated from regressions for the daily realized variance (RV_t), the daily realized volatility in standard deviation form ($RV_t^{1/2}$), and the daily realized volatility in logarithmic form ($log(RV_t^{1/2})$) with the naive, subsampling and kernel-based estimators.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
β_0	0.0130	0.0137	0.0130	0.0090	0.0093	0.0090	0.0074	0.0075	0.0074
	(0.0028)	(0.0022)	(0.0028)	(0.0019)	(0.0015)	(0.0019)	(0.0016)	(0.0013)	(0.0016)
β_1	0.0810	0.1013		0.0323	0.04683		0.0139	0.0211	
	(0.0746)	(0.0645)		(0.0741)	(0.0648)		(0.0739)	(0.0649)	
β_2	0.0762		0.1556	0.0594		0.0916	0.0283		0.0424
	(0.1505)		(0.1315)	(0.1580)		(0.1395)	(0.1619)		(0.1435)
R^2	0.0109	0.0102	0.0059	0.0026	0.0021	0.0018	0.0005	0.0004	0.0003
Adj. R^2	0.0024	0.0061	0.0017	-0.0059	-0.0020	-0.0024	-0.0080	-0.0037	-0.0039
Log-lik.	484.29	494.19	483.69	564.60	575.98	564.50	595.53	607.52	595.51
AIC	-4.0960	-4.1187	-4.0995	-4.7796	-4.8031	-4.7873	-5.0428	-5.0671	-5.0511
SC	-4.0519	-4.0896	-4.0701	-4.7354	-4.7740	-4.7578	-4.9986	-5.0380	-5.0217

Table 6: OLS estimates for the daily realized variance (RV_t) with the HAR-RV model (three estimators, naive: columns (1) to (3); subsampling: columns (4) to (6); kernel: columns (7) to (9)).

Note: The model estimated is $RV_t = \beta_0 + \beta_1 RV_{t-1} + \beta_2 RV_{t-6,t-1} + u_t$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
β_0	0.0490	0.0680	0.0497	0.0388	0.0577	0.0390	0.0366	0.0538	0.0367
	(0.0107)	(0.0076)	(0.0108)	(0.0088)	(0.0061)	(0.0089)	(0.0081)	(0.0054)	(0.0081)
β_1	0.1904	0.3118		0.1522	0.2960		0.1181	0.2554	
	(0.0770)	(0.0615)		(0.0773)	(0.0619)		(0.0775)	(0.0627)	
β_2	0.3174		0.5013	0.3782		0.5281	0.3787		0.4964
	(0.1234)		(0.0996)	(0.1255)		(0.1004)	(0.1298)		(0.1046)
R^2	0.1211	0.0975	0.0980	0.1207	0.0877	0.1060	0.0971	0.0653	0.0881
Adj. R^2	0.1136	0.0937	0.0941	0.1131	0.0839	0.1022	0.0894	0.0613	0.0842
Log-lik.	290.55	293.69	287.49	359.07	362.22	357.12	382.36	386.22	381.19
AIC	-2.4472	-2.4409	-2.4297	-3.0304	-3.0144	-3.0223	-3.2286	-3.2152	-3.2271
SC	-2.4030	-2.4118	-2.4003	-2.9862	-2.9853	-2.9929	-3.1844	-3.1861	-3.1977

Table 7: OLS estimates for the daily realized volatility in standard deviation form ($RV_t^{1/2}$) with the HAR-RV model (three estimators, naive: columns (1) to (3); subsampling: columns (4) to (6); kernel: columns (7) to (9)).

Note: The model estimated is $RV_t^{1/2} = \beta_0 + \beta_1 RV_{t-1}^{1/2} + \beta_2 RV_{t-6,t-1}^{1/2} + u_t$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
β_0	-0.6329	-1.2134	-0.6452	-0.6663	-1.1683	-0.6866	-0.7345	-1.2884	-0.7512
	(0.1850)	(0.1479)	(0.1884)	(0.1893)	(0.1493)	(0.1960)	(0.2039)	(0.1581)	(0.2101)
β_1	0.2480	0.5275		0.3299	0.5678		0.3058	0.542864	
	(0.0789)	(0.0549)		(0.0776)	(0.0534)		(0.0781)	(0.0545)	
β_2	0.5041		0.7473	0.4226		0.7449	0.4325		0.7322
	(0.1043)		(0.0713)	(0.1022)		(0.0711)	(0.1045)		(0.0733)
R^2	0.3477	0.2798	0.3200	0.3691	0.3226	0.3200	0.3429	0.2946	0.2995
Adj. R^2	0.3421	0.2768	0.3171	0.3637	0.3197	0.3171	0.3373	0.2916	0.2965
Log-lik.	-220.98	-228.18	-225.88	-190.05	-200.23	-198.87	-184.02	-194.03	-191.54
AIC	1.9062	1.9851	1.9394	1.6430	1.6923	1.7095	1.5917	1.6404	1.6471
SC	1.9504	2.0142	1.9688	1.6872	1.7214	1.7389	1.6358	1.6695	1.6766

Table 8: OLS estimates for the daily realized volatility in logarithmic form ($log(RV_t^{1/2})$) with the HAR-RV model (three estimators, naive: columns (1) to (3); subsampling: columns (4) to (6); kernel: columns (7) to (9)).

Note: The model estimated is $log(RV_t^{1/2}) = \beta_0 + \beta_1 log(RV_{t-1}^{1/2}) + \beta_2 log(RV_{t-6,t-1}^{1/2}) + u_t$.

	b_0	b_1	b_2	R^2
Daily realized variance (RV_t)				
HAR-RV	0.006327	0.5678		0.0011
	(0.02075)	(1.3777)		
GARCH daily	0.01301		1970.17	0.0028
	(0.00434)		(3120.91)	
HAR-RV +	0.00699	0.4156	1788.50	0.0033
GARCH daily	(0.0208)	(1.4074)	(3190.7)	
Daily realized volatility in standard deviation form ($RV_t^{1/2}$)				
HAR-RV	-0.00654	1.0408***		0.1139
	(0.0240)	(0.2419)		
GARCH daily	0.05527		45.8000***	0.0703
	(0.0130)		(13.879)	
HAR-RV +	-0.0069	0.8526***	23.403	0.1281
GARCH daily	(0.0238)	(0.2735)	(15.322)	
Daily realized volatility in logarithmic form ($log(RV_t^{1/2})$)				
HAR-RV	0.1479	1.0656***		0.4704
	(0.2517)	(0.0942)		
GARCH daily	2.3640***		0.6945***	0.1917
	(0.8599)		(0.1188)	
HAR-RV +	1.2419*	0.9724***	0.1854*	0.4800
GARCH daily	(0.7032)	(0.1090)	(0.1113)	

Table 9: Estimates of the Mincer-Zarnowitz regression (equations 15 to 17) using forecasts for the daily realized variance, the daily realized volatility, and the daily realized volatility in logarithmic form obtained from the naive estimator.

Note: The values reported in parentheses are robust standard errors.