Econometric Reviews

ISSN: 0747-4938 (Print) 1532-4168 (Online) Journal homepage: https://www.tandfonline.com/loi/lecr20

The Volatility of Realized Volatility

Fulvio Corsi , Stefan Mittnik , Christian Pigorsch & Uta Pigorsch

To cite this article: Fulvio Corsi , Stefan Mittnik , Christian Pigorsch & Uta Pigorsch (2008) The Volatility of Realized Volatility, Econometric Reviews, 27:1-3, 46-78, DOI: 10.1080/07474930701853616

To link to this article: https://doi.org/10.1080/07474930701853616

Published online: 07 Mar 2008.

Submit your article to this journal

Article views: 1703

View related articles

Citing articles: 229 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=lecr20

Econometric Reviews, 27(1–3):46–78, 2008
Copyright © Taylor & Francis Group, LLC
ISSN: 0747-4938 print/1532-4168 online
DOI: 10.1080/07474930701853616

THE VOLATILITY OF REALIZED VOLATILITY

Fulvio Corsi,[1] Stefan Mittnik,[2] Christian Pigorsch,[3] and Uta Pigorsch[4]

[1] *University of Lugano and Swiss Finance Institute, Lugano, Switzerland*
[2] *Department of Statistics, University of Munich, München, Germany; Center for Financial Studies, Frankfurt; and Ifo Institute for Economic Research, Munich, Germany*
[3] *Department of Economics, University of Bonn, Bonn, Germany*
[4] *Department of Economics, University of Mannheim, Mannheim, Germany*

□ *In recent years, with the availability of high-frequency financial market data modeling realized volatility has become a new and innovative research direction. The construction of "observable" or realized volatility series from intra-day transaction data and the use of standard time-series techniques has lead to promising strategies for modeling and predicting (daily) volatility. In this article, we show that the residuals of commonly used time-series models for realized volatility and logarithmic realized variance exhibit non-Gaussianity and volatility clustering. We propose extensions to explicitly account for these properties and assess their relevance for modeling and forecasting realized volatility. In an empirical application for S&P 500 index futures we show that allowing for time-varying volatility of realized volatility and logarithmic realized variance substantially improves the fit as well as predictive performance. Furthermore, the distributional assumption for residuals plays a crucial role in density forecasting.*

Keywords Density forecasting; Finance; HAR-GARCH; Normal inverse Gaussian distribution; Realized quarticity; Realized volatility.

JEL Classification C22; C51; C52; C53.

1. INTRODUCTION

Volatility plays an important role both in theoretical developments as well as in practical applications in finance. With the availability of

Received June 8, 2005; Accepted August 15, 2006
Address correspondence to Stefan Mittnik, Department of Statistics, University of Munich, Akademiestr. 1/I, München D-80799, Germany; E-mail: finmetrics@stat.uni-muenchen.de

high-frequency data research on the volatility of returns on financial assets has taken new avenues. Next to directly modeling high-frequency returns, intra-day returns are also used to construct nonparametric, lower-frequency (daily) volatility measures, termed *realized volatility*. Due to its non-latent character, realized volatility is not only used to assess the predictive performance and adequacy of existing stochastic-volatility models (see, for example, Andersen and Bollerslev, 1998), but also to explore the predictability of realized volatility. In fact, reduced-form models for realized volatility have already been considered for a variety of different markets and data sets. Andersen et al. (2003) suggest a fractionally integrated autoregressive moving average (ARFIMA) model for realized volatility to capture its distinct long-memory behavior. Persistent sample autocorrelation functions have been widely reported for various volatility measures of financial assets. Barndorff-Nielsen and Shephard (2002a) and Koopman et al. (2005) instead specify an unobserved autoregressive moving average (ARMA) component (UC) model that is based on a superposition of Ornstein–Uhlenbeck processes. Another and, due to its straightforward estimation, rather appealing model for realized volatility is the heterogeneous autoregressive (HAR) model as proposed in Corsi (2004). Although it is formally not a long-memory model, it can adequately reproduce the observed hyperbolic decay of the autocorrelation function by specifying a sum of volatility components over different horizons.

All three models have been shown to significantly improve volatility forecasts relative to conventional stochastic-volatility or generalized autoregressive conditional heteroskedasticity (GARCH) models. In both the ARFIMA and the HAR models, it is commonly assumed that innovations are Gaussian as well as identically and independently distributed (IID). In the UC-model literature no specific distributional assumption is made; but when estimating via quasi-maximum-likelihood, the Gaussian assumption enters. Moreover, the UC model also assumes white noise innovations. Although the Gaussianity assumption seems to be more acceptable when modeling the logarithm of realized volatility, we will show that it is particularly inadequate for non-logarithmic realized volatility. Empirical distributions of ARFIMA and HAR residuals tend to exhibit right skewness and fat tails. In addition, regardless of the transformation considered, we find volatility clustering in the residuals of these models, a violation of the IID assumption. Similar patterns can be expected in the UC model, since part of the time-varying variance might be attributed to the variance of the realized volatility estimator. Ignoring such properties will lead to inefficiencies when estimating realized-volatility models and result in an inferior forecasting performance. More importantly, in practical applications the presence of time-varying and non-Gaussian conditional distributions can distort risk assessment and, thus, impair risk management.

In this article, we investigate the importance of the observed volatility of realized volatility in modeling, and forecasting applications and propose two extensions of standard realized-volatility models. We allow for non-Gaussian innovations and, instead, suggest the use of the more flexible normal inverse Gaussian distribution. Furthermore, to model time-dependent conditional heteroskedasticity we also specify a GARCH specification, which can account for clustering and—to some extent—for the observed unconditional kurtosis. By doing so, we explicitly model the volatility of realized volatility which, to our knowledge, has not yet been considered in the literature.

Our assessment is twofold. Since the standard deviation of realized variance is the main variable of interest for financial applications, our assessment is primarily conducted in terms of realized volatility. In particular, we directly model realized volatility. However, it is also widely accepted that the unconditional distribution of the logarithmic transformation of realized variance is closer to Gaussianity (see, for example, Andersen et al., 2001a,b; Gonçalves and Meddahi, 2005) leading many researchers to formulate their volatility models in terms of the logarithmic transform. We therefore also consider logarithmic-realized-variance models. This allows us to investigate the adequacy of the Gaussianity assumption for these models as well as the relevance of the time-variation of volatility which we also observe for the logarithmic realized variance. The predictive performance of the different models, however, is assessed in terms of realized volatility, which is more relevant from the viewpoint of financial economics.

To assure that the observed time-variation in the volatility of realized volatility is no artefact due to misspecifications of the HAR or ARFIMA model, we also investigate the time-series behavior of the volatility of the realized volatility estimator relying on the asymptotic distribution theory of realized volatility derived in Barndorff-Nielsen and Shephard (2002a). Using different measures of integrated quarticity, the resulting series of the volatility of the realized-volatility error exhibit the same characteristics as those found for squared residuals. Similar, although less pronounced, patterns can also be observed for the corresponding measures of the volatility of logarithmic realized variance. The time-varying behavior has also been reported in Barndorff-Nielsen and Shephard (2005) who plot confidence intervals for the measurement error of realized variance. Those results suggest that any realized volatility model might be subject to heteroskedastic errors due to the time-varying volatility of the realized volatility estimator. The assessment of the relevance of the volatility of realized volatility in modeling and forecasting is therefore very important.

The remainder of the article is organized as follows. The next section briefly reviews the construction of measures of realized volatility and the volatility of the realized-volatility/logarithmic-realized variance error; it also describes the S&P 500 index futures data set used in the empirical application. Section 3 gives a short description of the standard realized-volatility models considered in this article, discusses the model extensions we propose, and presents in-sample estimation results. Section 4 presents the results of a simulation study designed to assess the efficiency implications of the proposed extensions. Section 5 reports and compares out-of-sample point and density forecasts. Section 6 concludes.

2. MEASUREMENT AND DATA

We begin by briefly reviewing the theory of quadratic variation and integrated variance and its estimation using realized variance. The asymptotic theory of this estimator is reproduced involving the notions of integrated quarticity and alternative quarticity measures, as introduced by Andersen et al. (2007, 2005) and Barndorff-Nielsen and Shephard (2003, 2004b, 2005, 2006). This allows us to compute an approximation of the volatility of the realized volatility estimator and the volatility of the logarithmic realized variance estimator. We then proceed with the description of the data set.

2.1. Construction of Volatility Measures

Let the logarithmic price of a financial asset, denoted by p_t, follow the stochastic-volatility process

$$p_t = p_0 + \int_0^t \mu(s)\,ds + \int_0^t \sigma(s)\,dW(s), \tag{1}$$

where μ and σ are càdlàg; W is a standard Brownian motion; and σ is assumed to be independent of W. Stochastic-volatility processes of this form represent a (special type of) semimartingale and are widely used in financial modeling. The *quadratic variation* process for a sequence of partitions, $\tau_0 = 0 \leq \tau_1 \leq \cdots \leq \tau_n = t$, is defined by

$$[p]_t = \operatorname*{plim}_{n\to\infty} \sum_{j=0}^{n-1} (p_{\tau_{j+1}} - p_{\tau_j})^2. \tag{2}$$

With $\sup_j\{\tau_{j+1} - \tau_j\} \to 0$ for $n \to \infty$ we obtain the *integrated variance*

$$[p]_t = \int_0^t \sigma(s)^2\,ds. \tag{3}$$

As already shown by Merton (1980) and extended by Comte and Renault (1998), Andersen and Bollerslev (1998), Andersen et al. (2001b), and by Barndorff-Nielsen and Shephard (2001), the quadratic variation and hence the integrated variance can be consistently estimated by the sum of squared returns computed over very small time intervals. These results hold even if the exact form of the drift and volatility processes are unknown (see Barndorff-Nielsen and Shephard, 2002a).

Focusing specifically on the integrated variance over one-day intervals, as is commonly done, we denote the continuously compounded within-day returns of day t with sampling frequency M by

$$r_{t,j} = p_{t-1+\frac{j}{M}} - p_{t-1+\frac{(j-1)}{M}}, \qquad j = 1, \ldots, M, \tag{4}$$

and define the *realized variance* over day t by

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2. \tag{5}$$

Then, by the theory of quadratic variation of semimartingales, (daily) realized variance converges uniformly in probability to the (daily) quadratic variation process as the sampling frequency of returns approaches infinity, i.e., for $M \to \infty$

$$RV_t \to \int_{t-1}^{t} \sigma^2(s)\, ds, \tag{6}$$

providing a consistent estimate of the integrated variance. In fact, Barndorff-Nielsen and Shephard (2002a) have shown that realized variance converges to integrated variance at rate $\sqrt{M}$.

Given a consistent estimator for the integrated variance of the stochastic-volatility model (1) the question of precision arises. The asymptotic distribution of the estimator has been derived in Barndorff-Nielsen and Shephard (2002a,b, 2003, 2004a, 2005) and is given by

$$\frac{\sqrt{M}\left(RV_t - \int_{t-1}^{t} \sigma^2(s)\, ds\right)}{\sqrt{2 \int_{t-1}^{t} \sigma^4(s)\, ds}} \xrightarrow{d} N(0, 1), \tag{7}$$

where $\int_{t-1}^{t} \sigma^4(s)\, ds$ denotes *integrated quarticity*. Note that this result does not require the exact knowledge of the drift and variance processes, μ and σ, and that the asymptotic normality holds even if the fourth moments of the returns do not exist.

Unfortunately, the computation of the asymptotic distribution is infeasible, given that the integrated quarticity is unknown. Based on the theory of power variation, Barndorff-Nielsen and Shephard (2002a, 2004b, 2006) suggest different estimators of integrated quarticity. The *realized fourth-power variation* or *realized quarticity,* defined as

$$RQ_t = \frac{M}{3}\sum_{j=1}^{M} r_{t,j}^4 \rightarrow \int_{t-1}^{t} \sigma^4(s)\,ds, \tag{8}$$

is a consistent estimator of the integrated quarticity. A more robust estimator, especially in the presence of jumps, is the *realized quad-power quarticity,*

$$RQQ_t = M\frac{\pi^2}{4}\sum_{j=4}^{M} |r_{t,j}||r_{t,j-1}||r_{t,j-2}||r_{t,j-3}| \rightarrow \int_{t-1}^{t} \sigma^4(s)\,ds. \tag{9}$$

An alternative and similarly robust measure, the *realized tri-power quarticity,*

$$RTQ_t = M\frac{\Gamma\left(\frac{1}{2}\right)^3}{4\Gamma\left(\frac{7}{6}\right)^3}\sum_{j=3}^{M} |r_{t,j}|^{\frac{4}{3}}|r_{t,j-1}|^{\frac{4}{3}}|r_{t,j-2}|^{\frac{4}{3}} \rightarrow \int_{t-1}^{t} \sigma^4(s)\,ds, \tag{10}$$

has been proposed in Andersen et al. (2007).

Based on these different quarticity measures, the asymptotic distribution of realized variance can be approximated by

$$\frac{RV_t - \int_{t-1}^{t} \sigma^2(s)\,ds}{\sqrt{\frac{2}{M}Q_t^*}} \xrightarrow{d} N(0,1), \quad Q_t^* \in (RQ_t, RQQ_t, RTQ_t), \tag{11}$$

where $\sqrt{\frac{2}{M}Q_t^*}$ provides an approximation of the standard deviation of the realized variance error.

Being interested in the volatility of the realized-volatility error, the delta method can be used to approximate the asymptotic distribution of realized volatility, i.e.,

$$\frac{\sqrt{RV_t} - \sqrt{\int_{t-1}^{t} \sigma^2(s)\,ds}}{\sqrt{\frac{Q_t^*}{2MRV_t}}} \xrightarrow{d} N(0,1), \quad Q_t^* \in (RQ_t, RQQ_t, RTQ_t). \tag{12}$$

From the different measures of integrated quarticity, we can compute three alternative approximations of the (daily) volatility of the realized volatility estimator $\sqrt{\frac{Q_t^*}{2MRV_t}}$, namely,

$$\sqrt{\frac{RQ_t}{2MRV_t}} = \sqrt{\frac{\sum_{j=1}^{M} r_{t,j}^4}{6\sum_{j=1}^{M} r_{t,j}^2}} \tag{13}$$

$$\sqrt{\frac{RQQ_t}{2MRV_t}} = \sqrt{\frac{\pi^2 \sum_{j=4}^{M} |r_{t,j}||r_{t,j-1}||r_{t,j-2}||r_{t,j-3}|}{8\sum_{j=1}^{M} r_{t,j}^2}} \tag{14}$$

$$\sqrt{\frac{RTQ_t}{2MRV_t}} = \sqrt{\frac{\Gamma(\frac{1}{2})^3 \sum_{j=3}^{M} |r_{t,j}|^{\frac{4}{3}}|r_{t,j-1}|^{\frac{4}{3}}|r_{t,j-2}|^{\frac{4}{3}}}{8\Gamma(\frac{7}{6})\sum_{j=1}^{M} r_{t,j}^2}} \tag{15}$$

As mentioned earlier, the logarithmic transformation of realized variance is well-known to provide better finite sample properties, such that we are also interested in the volatility of the logarithmic realized-variance error. In particular, Barndorff-Nielsen and Shephard (2002c) derive the approximate asymptotic distribution of the logarithmic transform of realized variance

$$\frac{\log RV_t - \log \int_{t-1}^{t} \sigma^2(s)\,ds}{\sqrt{\frac{2Q_t^*}{M(RV_t)^2}}} \xrightarrow{d} N(0,1), \qquad Q_t^* \in (RQ_t, RQQ_t, RTQ_t). \tag{16}$$

Based on the realized quarticity measure, Barndorff-Nielsen and Shephard (2006) show in a simulation study that the implied standard errors for the logarithmic transformation indeed tend to be smaller than those of the realized variance statistic given in (7), with integrated variance being also estimated by the realized quarticity measure. The different measures of the (daily) volatility of the logarithmic realized-variance error are given by

$$\sqrt{\frac{2RQ_t}{M(RV_t)^2}} = \sqrt{\frac{2}{3}\frac{\sum_{j=1}^{M} r_{t,j}^4}{\left(\sum_{j=1}^{M} r_{t,j}^2\right)^2}} \tag{17}$$

$$\sqrt{\frac{2RQQ_t}{M(RV_t)^2}} = \sqrt{\frac{\pi^2}{2}\frac{\sum_{j=4}^{M} |r_{t,j}||r_{t,j-1}||r_{t,j-2}||r_{t,j-3}|}{(\sum_{j=1}^{M} r_{t,j}^2)^2}} \tag{18}$$

$$\sqrt{\frac{2RTQ_t}{M(RV_t)^2}} = \sqrt{\frac{\Gamma(\frac{1}{2})^3 \sum_{j=3}^{M} |r_{t,j}|^{\frac{4}{3}}|r_{t,j-1}|^{\frac{4}{3}}|r_{t,j-2}|^{\frac{4}{3}}}{2\Gamma(\frac{7}{6})^3(\sum_{j=1}^{M} r_{t,j}^2)^2}} \tag{19}$$

The next subsection provides and compares the descriptive statistics of the derived volatility measures from S&P 500 index futures data.

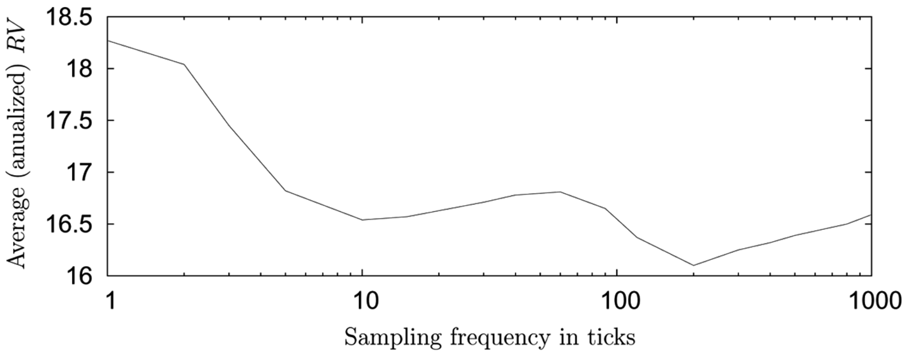

FIGURE 1 Volatility signature plot of the S&P 500 index futures constructed over the full sample period. The graph shows average annualized realized volatility constructed for different frequencies measured in number of ticks. Note that there are about 7 seconds on average between trades, such that the average annualized five-minute based realized volatility corresponds to around the 43th tick.

2.2. Data

Our empirical application is based on tick-by-tick transaction prices of S&P 500 index futures recorded at the Chicago Mercantile Exchange (CME). The sample covers the period from January 1, 1985 to December 31, 2004, a period of 5,040 trading days, and consists of 13,241,032 tick-by-tick observations.[1] It follows from the theoretical considerations discussed above that the sampling frequency for constructing the volatility measures should be as large as possible. In practice, however, very high-frequency returns are contaminated by transaction costs, bid-and-ask-bounce effects, etc., leading to biases in the variance measures. It is therefore common practice to handle this trade-off by summing returns over 5 or 30 minutes (see, for example, Andersen et al., 2007, 2001b; Barndorff-Nielsen and Shephard, 2004b).[2] Given the high liquidity of the S&P 500 index futures market, we follow Andersen and Bollerslev (1998), Andersen et al. (2001b), Maheu and McCurdy (2002), and Martens et al. (2004), among others, and use five-minute returns to construct our realized-variance and quarticity measures. Moreover, the volatility signature plot given in Figure 1—depicting the average annualized realized volatility over the full sample period constructed for different frequencies—indeed indicates that the bias induced by market microstructure effects is relatively

[1]We disregard the overnight trading of contracts at GLOBEX, the CME overnight trading platform, which started in 1994.

[2]The impact of market-microstructure effects on the realized-variance measures as well as possible data-adjustment and prefiltering procedures, allowing the full use of the tick-by-tick data, is discussed in Aït-Sahalia et al. (2005), Areal and Taylor (2002), Bandi and Russell (2005), Barndorff-Nielsen et al. (2006), Corsi et al. (2001), Curci and Corsi (2003), Hansen and Lunde (2006b), and Zhang et al. (2005) among others.

TABLE 1 Descriptive statistics

Series	Mean	SD	Median	Skewness	Kurtosis	Ljung-Box (22)	$\hat{d}$
$\sqrt{RV_t}$	0.8627	0.5935	0.7586	15.35	496.76	14,605.0	0.4765
$\log RV_t$	−0.5139	0.8775	−0.5527	0.60	4.8070	22,023.2	0.4982
$\sqrt{\frac{RQ_t}{2MRV_t}}$	0.0821	0.0893	0.0675	19.29	606.68	3,498.8	0.3862
$\sqrt{\frac{RQQ_t}{2MRV_t}}$	0.0676	0.0475	0.0570	6.62	96.89	12,551.5	0.4404
$\sqrt{\frac{RTQ_t}{2MRV_t}}$	0.0704	0.0500	0.0594	6.29	79.95	10,955.2	0.4271
$\sqrt{\frac{2RQ_t}{M(RV_t)^2}}$	0.2034	0.0915	0.1862	12.73	301.89	56.6	0.0541
$\sqrt{\frac{2RQQ_t}{M(RV_t)^2}}$	0.1648	0.0285	0.1599	1.71	9.75	33.4	0.0303
$\sqrt{\frac{2RTQ_t}{M(RV_t)^2}}$	0.1720	0.0302	0.1656	2.23	12.05	85.6	0.0621

Reported are the descriptive statistics of realized volatility, logarithmic realized variance, and the measures of the volatility of the realized volatility estimator as defined in Eqs. (13)–(15), as well as the measures of the volatility of the logarithmic realized variance as defined in Eqs. (17)–(19). $\hat{d}$ reports the Geweke-Porter-Hudak estimates of the fractional integration parameter.

small for the S&P 500 index futures, and dies out very quickly. Note that with a transaction taking place on average about every seven seconds, the average annualized realized volatility based on the five-minute intervals corresponds to around the 43th tick presented in the figure.[3] The impact of market microstructure effects on the five-minute realized volatility measure for the S&P 500 index futures over the period from 1985 to 2004 can therefore be regarded as negligible. The five-minute returns were constructed using the nearest neighbor to the five-minute tag, excluding overnight returns, and by rolling over to the most liquid contract.

Table 1 presents descriptive statistics of the computed realized volatility, $\sqrt{RV_t}$, logarithmic realized variance, $\log RV_t$, as well as the three volatility measures of the realized volatility estimator defined in (13)–(15), and the measures of the volatility of the logarithmic realized variance estimator as defined in (17)–(19). Figures 2 and 3 show the time series as well as the sample autocorrelation and partial autocorrelation functions for the four realized volatility and those for the four logarithmic realized variance related measures, respectively. Table 1 reveals that the distribution of realized volatility is fat-tailed and slightly skewed. Taking the logarithm of realized variance instead leads to a strong reduction in skewness and kurtosis, a finding which is in line with the empirical observations made in Andersen et al. (2001a,b), among others. Given that the Gaussianity assumption is more suitable for the logarithmic transform of realized variance, Andersen et al. (2001b) model the logarithmic transform of realized variance. However, the descriptive statistics show that

[3]Similarly, decreasing volatility signature plots for liquid assets based on transaction prices have also been shown in Hansen and Lunde (2006a), for example.

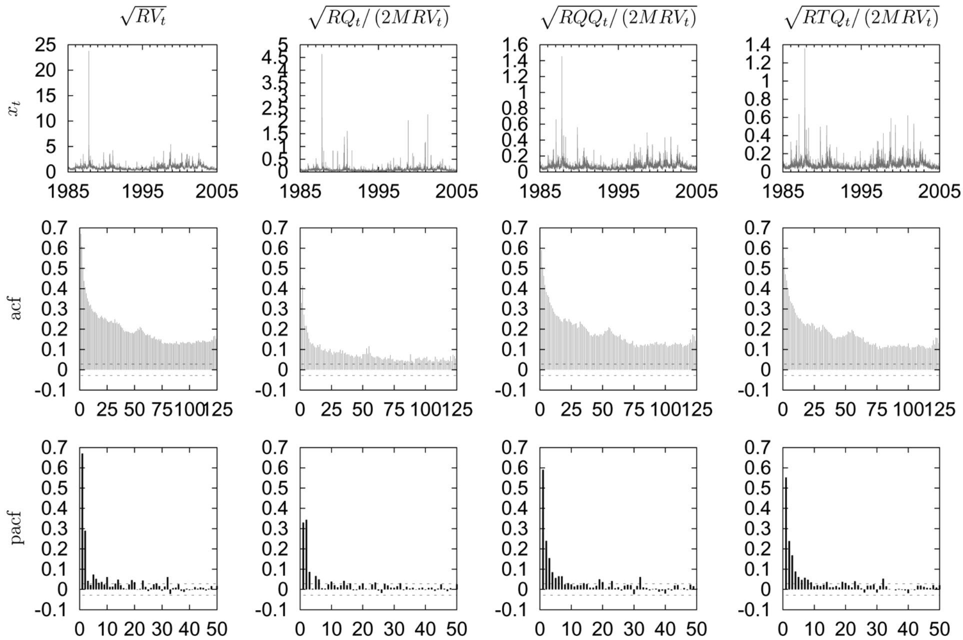

FIGURE 2 Time series (upper panel), sample autocorrelation functions (acf) (middle panel) and partial autocorrelation functions (pacf) (bottom panel) of realized volatility and the three measures of the volatility of the realized volatility estimator as defined in Eqs. (13)–(15).

skewness and kurtosis are not completely eliminated by the logarithmic transformation,[4] so that the Gaussianity assumption deserves further investigation. Gonçalves and Meddahi (2005), for example, show that other nonlinear transformations are more effective in reducing the sample skewness.

Figures 2 and 3 as well as the Ljung–Box statistics reported in Table 1 indicate strong autocorrelation in realized volatility and its logarithmic transform. In fact, both autocorrelation functions exhibit a hyperbolic decay. The Geweke–Porter–Hudak estimates of the fractional difference parameters presented in Table 1 also support this persistence. This finding is no artefact of the sampling or aggregation scheme employed in the construction of the realized volatility measure, but is rather consistent with the long-memory behavior of volatility that has been extensively reported in the GARCH literature. Although the true source of long memory in volatility is not clear (see, for example, Banerjee and Urga, 2005; Engle and Lee, 1999; Mikosch and Starica, 2004), its existence has been

[4]The Kolmogorov–Smirnov test rejects the null of Gaussianity (p-value $= 0.0087$). Our results differ from those reported in Thomakos and Wang (2003), who also perform tests on Gaussianity but use a much shorter sample period.

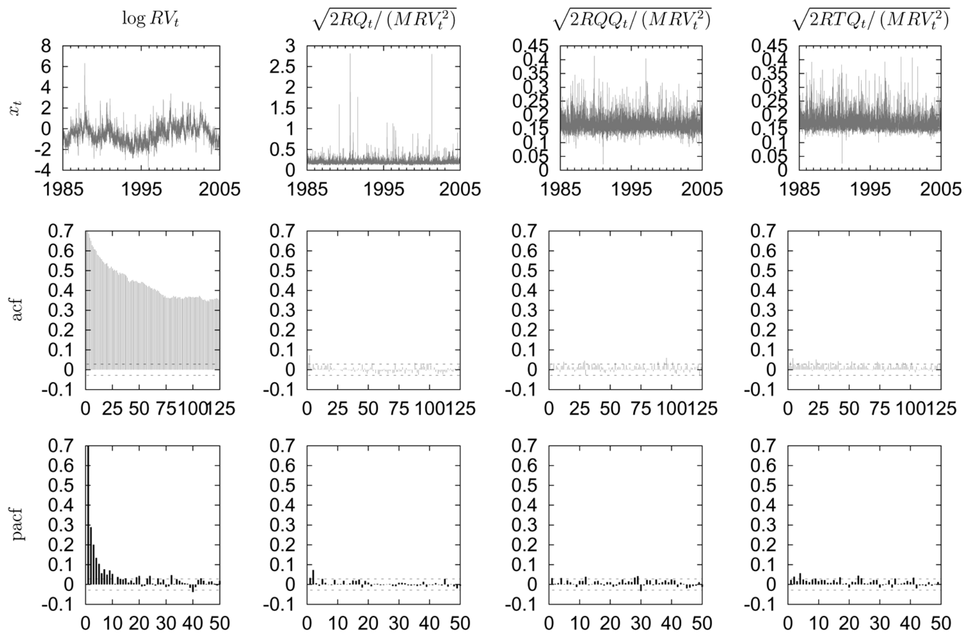

FIGURE 3 Time series (upper panel), sample autocorrelation functions (acf) (middle panel) and partial autocorrelation functions (pacf) (bottom panel) of logarithmic realized variance and the three measures of the volatility of the logarithmic realized variance estimator as defined in Eqs. (17)–(19).

widely recognized and captured in different volatility models, such as, for example, the FIGARCH or ARFIMA models.

The unconditional distributions of all three measures for volatility of the realized volatility exhibit skewness and leptokurtosis, both of which are most pronounced for the realized-quarticity-based measure (13). This can be explained by the construction of this measure, with the fourth power yielding high values for large (absolute) intraday returns. The Geweke–Porter–Hudak estimates in Table 1, and the autocorrelation functions presented in Figure 2 also show that all volatility of realized volatility series exhibit long memory. Moreover, the volatility of realized volatility assumes high values when realized volatility is high (cf. Barndorff-Nielsen and Shephard, 2005). Most importantly, all three measures exhibit time-variation and volatility clustering. Similar patterns are also found for the volatility of the log-transformed realized volatility estimator. However, fluctuations over time are somewhat lower. Although the Ljung–Box-Test indicates significant autocorrelation up to at least one month for two of the three series (the critical value at the 5% significance level is 33.92), the autocorrelation functions presented in Figure 3, and the corresponding Geweke–Porter–Hudak estimators suggest that there exists no long memory in the volatility of the logarithmic realized variance series.

3. MODELING REALIZED VOLATILITY

Volatility modeling plays a prominent role in financial econometrics and risk management. With the availability of high-frequency data the volatility literature has developed in several directions, one of which focuses on modeling and predicting alternative measures of realized volatility. In this section, we first summarize recent approaches to modeling realized volatility and then discuss the extensions under investigation.

3.1. Conventional Realized-Volatility Models

Our data analysis and the related empirical literature suggest that the persistence of realized volatility is a distinct feature a realized-volatility model should capture. As pointed out earlier, this finding is not only peculiar for the realized volatility measure. Using different volatility proxies numerous empirical studies—starting with Andersen and Bollerslev (1997), Ding et al. (1993), Ding and Granger (1996), and Granger et al. (2000)—show the existence of long memory in volatility. Although at the empirical level the evidence of a strong volatility persistence has been unanimously recognized, at the theoretical level there is much less consensus on the mechanism generating this phenomenon. This is due to the fact that alternative long-memory models are consistent with the data and, thus, empirically indistinguishable.[5] As a consequence, the source of long memory in realized volatility is still an open issue. We therefore focus our attention on models that are commonly employed in the existing realized-volatility literature.

[5]In fact, there exists a large number of different approaches to explain long memory. Historically, the first class of long-memory models has been the fractionally integrated process proposed by Granger and Joyeux (1980) and Hosking (1981) (for comprehensive surveys see Beran, 1994; Robinson, 2003). With another seminal article showing the link between long memory and the aggregation of an infinite number of stationary processes, Granger (1980) also started an alternative strand of literature, which tries to approximate long-memory dependence through a multicomponent approach, as in Andersen and Bollerslev (1997), Barndorff-Nielsen and Shephard (2001), Engle and Lee (1999), Gallant et al. (1999), Lux and Marchesi (1999), and Müller et al. (1997). A profoundly different view on the source of long memory is instead offered by, among others, Diebold and Inoue (2001), Gourieroux and Jasiak (2001), Granger and Hyung (2004), Granger and Teräsvirta (1999), and Mikosch and Starica (2004), who provide theoretical justification and Monte Carlo evidence that models with structural breaks and regime-shifting may exhibit spurious long memory. In addition, other approaches for reproducing long-memory dependence, such as the multifractals and cascade models of Calvet and Fisher (2002, 2004) and Mandelbrot et al. (1997), or the error duration model of Parke (1999), have been proposed (see also Banerjee and Urga, 2005; Davidson and Teräsvirta, 2002 for recent reviews on long-memory models).

To capture the long memory in realized volatility or logarithmic realized variance, Andersen et al. (2003) specify the autoregressive fractionally integrated moving average, in short, ARFIMA(p, d, q) model

$$\phi(L)(1-L)^d(y_t-\mu)=\psi(L)u_t, \tag{20}$$

with d denoting the fractional difference parameter, $\phi(L)=1-\phi_1 L-\cdots-\phi_p L^p$ and $\psi(L)=1+\psi_1 L+\cdots\psi_q L^q$. Typically, u_t is assumed to be a Gaussian white noise process, and y_t denotes either the realized variance (see, for example, Koopman et al., 2005 or Oomen, 2004) or its logarithmic transform (as first advocated in Andersen et al., 2003). Retaining the Gaussianity assumption (regardless of the transformation of realized volatility considered), several articles have adopted and extended this model by including, for example, leverage effects (or other nonlinearities) and exogenous variables. The results reported in the literature for different markets and data sets show significant improvements in the point forecasts of volatility when using ARFIMA rather than GARCH-type models.[6] In the context of interval forecasting the distributional assumptions for the error terms should, however, be important.

An alternative to ARFIMA modeling of realized volatility, though formally not a long-memory model, has been suggested by Corsi (2004). Extending the heterogeneous ARCH model of Müller et al. (1997), the long-memory pattern is reproduced by a sum of (a small number of) volatility components constructed over different time horizons. Defining the k-period realized-volatility component by the sum of the single-period realized volatilities, i.e.,[7]

$$(\sqrt{RV})_{t+1-k:t}=\frac{1}{k}\sum_{j=1}^{k}\sqrt{RV_{t-j}}, \tag{21}$$

the HAR model of Corsi (2004), including the daily, weekly, and monthly realized-volatility components, is given by

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

[6]See, for example, Andersen et al. (2003), Koopman et al. (2005), Martens et al. (2004), Martens and Zein (2004), Oomen (2004), Pong et al. (2004), and Thomakos and Wang (2003), among others.

[7]Note that based on Jensen's inequality, the volatility components cannot exactly be interpreted as the realized-volatility over the specific time interval. However, our definition allows to interpret the HAR model as a restricted AR(22) model. Also, when employing the "true" daily, weekly, and monthly realized volatilities—as defined by the square root of the sum of the realized variances—we obtain similar empirical results.

In Corsi (2004), u_t is also assumed to be Gaussian white noise. Employing the volatility-component structure (22), simulations reported in Corsi (2004) show that the HAR model is able to reproduce the observed hyperbolic decay of the sample autocorrelations of realized volatility. Moreover, the HAR model's in- and out-of-sample performance is strong and even slightly better than that of an ARFIMA model for realized volatility. A good predictive performance has also been reported in Andersen et al. (2007), who extend the HAR model by including different jump measures. They also consider the HAR model for the logarithm of realized variance, yielding similar results. The logarithmic version of the HAR model considered here is given by

$$\log RV_t = \alpha_0 + \alpha_d \log RV_{t-1} + \alpha_w (\log RV)_{t-5:t-1} + \alpha_m (\log RV)_{t-22:t-1} + u_t \tag{23}$$

with the multiperiod logarithmic realized-variance components defined by

$$(\log RV)_{t+1-k:t} = \frac{1}{k} \sum_{j=1}^{k} \log RV_{t-j}. \tag{24}$$

Note, that we follow the above representation, i.e., we formulate the HAR model as a restricted AR(22) model.[8] Hence, the HAR model is nested within the general ARFIMA class, and (in-sample) model comparison is therefore straightforward.

In view of the similar performance of ARFIMA and HAR models and given the straightforward estimation of the latter, the HAR model might be preferable in practice. In contrast to the HAR model, the estimation of ARFIMA models is nontrivial. The simplest approach is to first estimate the fractional difference parameter, using, for example, the semiparametric estimator of Geweke and Porter-Hudak (1983), and then fit an ARMA model to the filtered series. However, the joint estimation of the ARMA parameters and the fractional difference parameter has been shown to generally improve the accuracy of the estimate of d,[9] though complicating the estimation since the long-memory autocovariance matrix needs to be estimated. Well-known methods for a joint maximum-likelihood estimation of ARFIMA parameters include the approaches of Hosking (1981) and Sowell (1992).[10]

The ARFIMA parameter estimates reported below are jointly estimated using exact maximum-likelihood with the Geweke–Porter–Hudak estimate

[8]As a consequence the logarithmic realized-variance components cannot be directly interpreted as the logarithm of the multiperiod realized variance.

[9]Agiakloglou et al. (1993) report poor small-sample properties of the Geweke–Porter–Hudak estimator.

[10]See Doornik and Ooms (2003) for a recent review on this topic.

TABLE 2 Estimation results

	Parameter estimates										
ARFIMA$(0, d, 3)$	d	ψ_1	ψ_2	ψ_3						AIC	BIC
	0.3483	0.09389	0.1798	0.0452							
	(0.0217)	(0.0258)	(0.0181)	(0.0174)						5484	5517
	Mean Eq.				Distribution		Variance Eq.				
HAR model	α_0	α_d	α_w	α_m	α	β	ω	α_1	β_1		
I_S	0.1066	0.4983	0.2132	0.1659			0.1985			5702	5735
	(0.0198)	(0.0015)	(0.0059)	(0.0191)			(0.0003)				
II_S	0.0657	0.2339	0.4541	0.2130			0.0040	0.7464	0.2428	−67	−21
	(0.0068)	(0.0170)	(0.0189)	(0.0177)			(0.0002)	(0.0053)	(0.0066)		
III_S	0.2180	0.2540	0.2285	0.2645	1.0313	0.6740	0.0933			−1306	−1260
	(0.0075)	(0.0068)	(0.0077)	(0.0078)	(0.0498)	(0.0479)	(0.0034)				
IV_S	0.0868	0.2322	0.3965	0.2565	1.6918	1.054	0.0034	0.8143	0.1237	−2316	−2257
	(0.0073)	(0.0142)	(0.0227)	(0.0184)	(0.1088)	(0.0975)	(0.0003)	(0.0117)	(0.0110)		
I_L	−0.0297	0.3483	0.3637	0.2318			0.2604			7496	7528
	(0.0093)	(0.0120)	(0.0227)	(0.0208)			(0.0035)				
II_L	−0.0386	0.3089	0.4005	0.2331			0.0430	0.7369	0.0963	7239	7284
	(0.0090)	(0.0187)	(0.0282)	(0.0234)			(0.0050)	(0.0241)	(0.0083)		
III_L	−0.0350	0.2997	0.3713	0.2626	1.4706	0.4150	0.2557			7018	7063
	(0.0084)	(0.0139)	(0.0229)	(0.0198)	(0.0879)	(0.0564)	(0.0068)				
IV_L	−0.0377	0.2938	0.4000	0.2469	1.6499	0.4478	0.0040	0.7611	0.0798	6899	6958
	(0.0085)	(0.0170)	(0.0268)	(0.0218)	(0.1076)	(0.0648)	(0.0081)	(0.0401)	(0.0132)		

The different HAR-model specifications are as follows: I is a standard HAR model with Gaussian innovations; II also includes GARCH effects; III is a standard HAR model with (standardized) NIG innovations; and IV corresponds to the HAR–GARCH model with (standardized) NIG innovations. The indices S and L denote the HAR models formulated for realized volatility and logarithmic realized variance, respectively. The numbers in parentheses are the standard errors. The AIC $= -2L + k$, and BIC $= -2L + k \log T$, where L denotes the log likelihood, k the number of parameters in the model, and T is the number of observations.

serving as starting value. The AIC and BIC criteria as well as the correlograms of the residuals suggest an ARFIMA$(0, d, 3)$ model for the realized volatility of S&P 500 index futures. Table 2 presents the parameter estimates of the ARFIMA and the standard HAR models, with the latter also being estimated via maximum likelihood.

Figures 4 and 5 show the results of the residual analysis for the two models. The time series plots and the sample autocorrelation and partial autocorrelation functions of the squared residuals clearly illustrate that the residuals of both models exhibit volatility clustering. In both cases, ARCH-LM tests indicate strong autoregressive conditional heteroskedasticity. This is in line with the time-variation observed in the measures of the volatility of the realized-volatility error. Moreover, the QQ-Plot and the kernel density estimates in Figures 4 and 5 convincingly illustrate the inadequacy of the normality assumption for both models. For the logarithmic realized variance we find the same form of time-dependent heteroskedasticity for

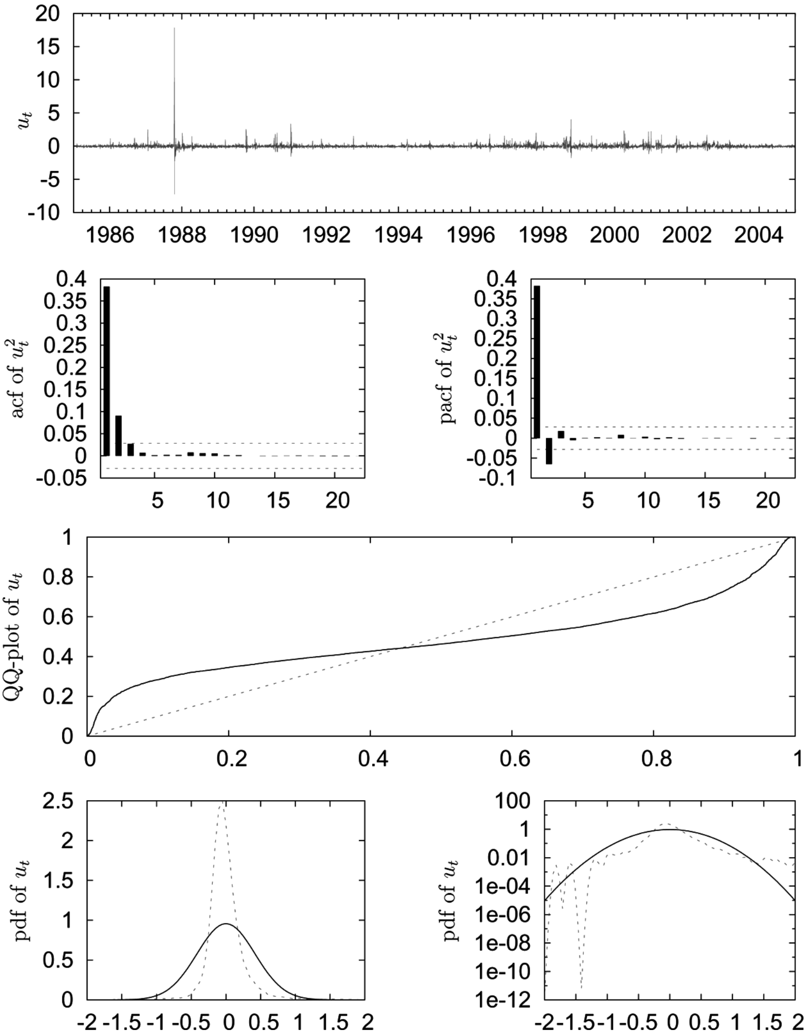

FIGURE 4 Residual analysis of the ARFIMA$(0, d, 3)$ model for realized volatility with Gaussian innovations. Shown are the time series of the residuals (upper panel), the sample autocorrelation functions (acf) and partial autocorrelation functions (pacf) of the squared residuals (second panel), the quantile–quantile plot (third panel) and on the bottom panel the kernel density estimates of the residuals (dashed line) and the estimated normal density (solid line) in level (left) and log scales (right).

both models.[11] However, the volatility clustering is less pronounced, a finding that is consistent with the characteristics of the measures for the volatility of logarithmic realized variance. Also, as expected, the non-Gaussianity is less pronounced, but still existent.

[11]For brevity we do not present the corresponding figures here.

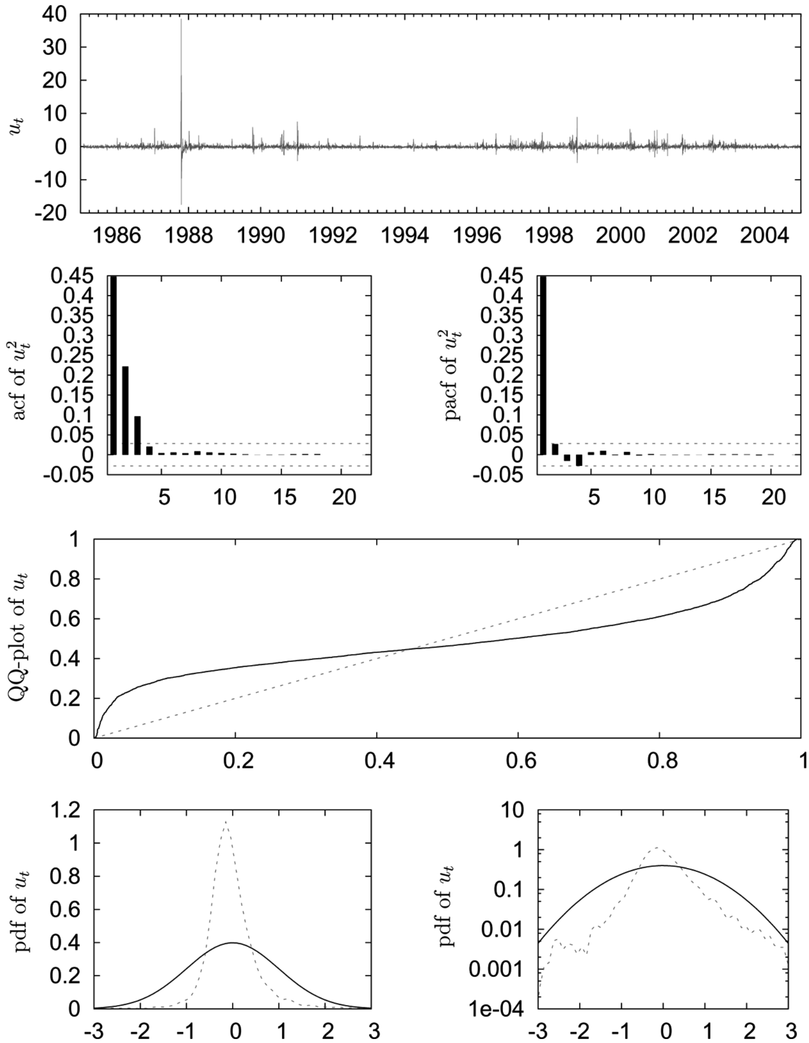

FIGURE 5 Residual analysis of the pure HAR model for realized volatility with Gaussian innovations. Shown are the time series of the residuals (upper panel), the sample autocorrelation functions (acf) and partial autocorrelation functions (pacf) of the squared residuals (second panel), the quantile–quantile plot (third panel) and on the bottom panel the kernel density estimates of the residuals (dashed line) and the estimated normal density (solid line) in level (left) and log scales (right).

3.2. Long Memory, Volatility Clustering, and Fat Tails

Motivated by the empirical results we extend the realized-volatility models in this section. Since HAR and ARFIMA models behave similarly in terms of forecasting and model misspecifications, we focus our discussion solely on extended HAR models. The proposed modifications can be straightforwardly adopted in an ARFIMA framework—though the estimation will be even more challenging. We expect the results for the extended HAR and ARFIMA models to be compatible. In fact, in Corsi (2004) it was shown that good fitting as well as good point forecast performance coincide in the HAR model. In particular, the empirical results indicate both, a slightly improved in-sample and out-of-sample performance of the HAR over the ARFIMA model.

To account for the observed volatility clustering in realized volatility, we extend the HAR model by including a GARCH component, giving rise to the HAR-GARCH(p, q) model

$$y_t = \alpha_0 + \alpha_d y_{t-1} + \alpha_w (y)_{t-5:t-1} + \alpha_m (y)_{t-22:t-1} + \sqrt{h_t}\epsilon_t \tag{25}$$

$$h_t = \omega + \sum_{j=1}^{q} \alpha_j u_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{26}$$

$$\epsilon_t \mid \Omega_{t-1} \sim (0, 1), \tag{27}$$

where Ω_{t-1} denotes the σ-field generated by all the information available up to time $t-1$, and y is either $\sqrt{RV}$ or $\log RV$. The error term, $u_t = \sqrt{h_t}\epsilon_t$, follows a conditional density with time-varying variance.

Although the incorporation of the GARCH specification can produce fatter unconditional tails, the normality assumption does not allow for the observed skewness. To deal with the non-Gaussianity of the error terms we specify a standardized normal inverse Gaussian (NIG) distribution for the (unconditional) iid innovations ϵ_t. The NIG distribution, introduced into the financial literature by Barndorff-Nielsen (1997, 1998), which is rather flexible and able to reproduce a range of symmetric and asymmetric distributions. Its density is given by

$$f(x; \alpha, \beta, \mu, \delta) = \frac{\alpha}{\pi} \frac{K_1\left(\alpha\delta\sqrt{1 + (\frac{x-\mu}{\delta})^2}\right)}{\sqrt{1 + (\frac{x-\mu}{\delta})^2}} \exp\left\{\delta\left(\sqrt{\alpha^2 - \beta^2} + \beta\left(\frac{x-\mu}{\delta}\right)\right)\right\} \tag{28}$$

where $K_i(x)$ is the modified Bessel function of the third kind and of order $i \in \mathbb{R}$; denotes the location parameter, $\delta > 0$ the scale, $\alpha > 0$ and $\beta \in [-\alpha, \alpha]$ the shape parameters, with $\beta = 0$ indicating a symmetric distribution. Mean and variance are given by

$$E[x] = \mu + \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}} \qquad \text{and} \qquad \mathrm{Var}[x] = \frac{\delta\alpha^2}{\sqrt{\alpha^2 - \beta^2}^3}. \tag{29}$$

To derive the *standardized* NIG distribution with zero mean and unit variance, we solve the resulting equations and derive the values of μ and δ in terms of α and β, and obtain

$$\mu = -\frac{\beta(\alpha^2 - \beta^2)}{\alpha^2} \qquad \text{and} \qquad \delta = \frac{(\alpha^2 - \beta^2)^{3/2}}{\alpha^2}. \tag{30}$$

By combining the HAR model (22) with a standardized NIG distribution and a GARCH specification, we obtain a quasi-long-memory model that should be able to capture both non-Gaussianity and time-dependent conditional heteroskedasticity.

Note that these extensions can also be adopted in the ARFIMA framework and have, in part, been considered in Baillie et al. (1996), who propose an ARFIMA–GARCH model to analyze inflation.

Maximum-likelihood estimates for various HAR specifications are presented in Table 2. Specifically, we extend the conventional HAR model with Gaussian innovations (Model I) by including a GARCH(1,1) specification (Model II); Model III corresponds to Model I but with zero mean NIG-distributed errors; and Model IV includes both modifications, i.e., we allow for NIG-distributed innovations and conditional heteroskedasticity. The corresponding results of the HAR model formulated for realized volatility are indexed by S, those for logarithmic realized variance are indexed by L.

The results show that the GARCH extension substantially improves the goodness of fit, as measured by the AIC and BIC criteria, especially for realized volatility. Both criteria as well as the ARCH-LM test suggest a GARCH(1,1) specification, which is also the preferred choice when modeling the volatility of asset returns. Comparing the parameter estimates of the mean equation of Model II to those of the standard HAR Model I, we observe an increase in the parameter of the weekly volatility component, α_w, while, at the same time, the influence of realized volatility lagged by one day decreases when including the GARCH specification.

It is well-known that for a GARCH process, the kurtosis of the dependent variable is determined by both the kurtosis of the error distribution and the persistence in the GARCH equation, i.e., by $\alpha_1 + \beta_1$ in a GARCH$(1, 1)$ process. Bai et al. (2003) have shown that the commonly reported parameter estimates, which are in the range of $0.85 < \hat{\alpha}_1 + \hat{\beta}_1 < 1$, are not sufficient for generating the observed kurtosis when assuming normally distributed errors. Given that our persistence estimates under the Gaussianity assumption lies within the (0.83, 1)-interval and that the kurtosis of realized volatility is much stronger than is commonly found for asset returns, simply adding a GARCH specification will not suffice to capture the observed kurtosis of realized volatility. A more heavy-tailed distribution for the innovations is required. Given that the logarithmic transformation of realized variance reduces skewness and kurtosis, the need for a more flexible distribution might be less important than for the realized volatility measure.

According to the goodness-of-fit measures reported in Table 2 replacing the Gaussian by the NIG distribution greatly improves the models' fit both for the conventional and the GARCH specification (Models I and II)—with the improvement being again more pronounced

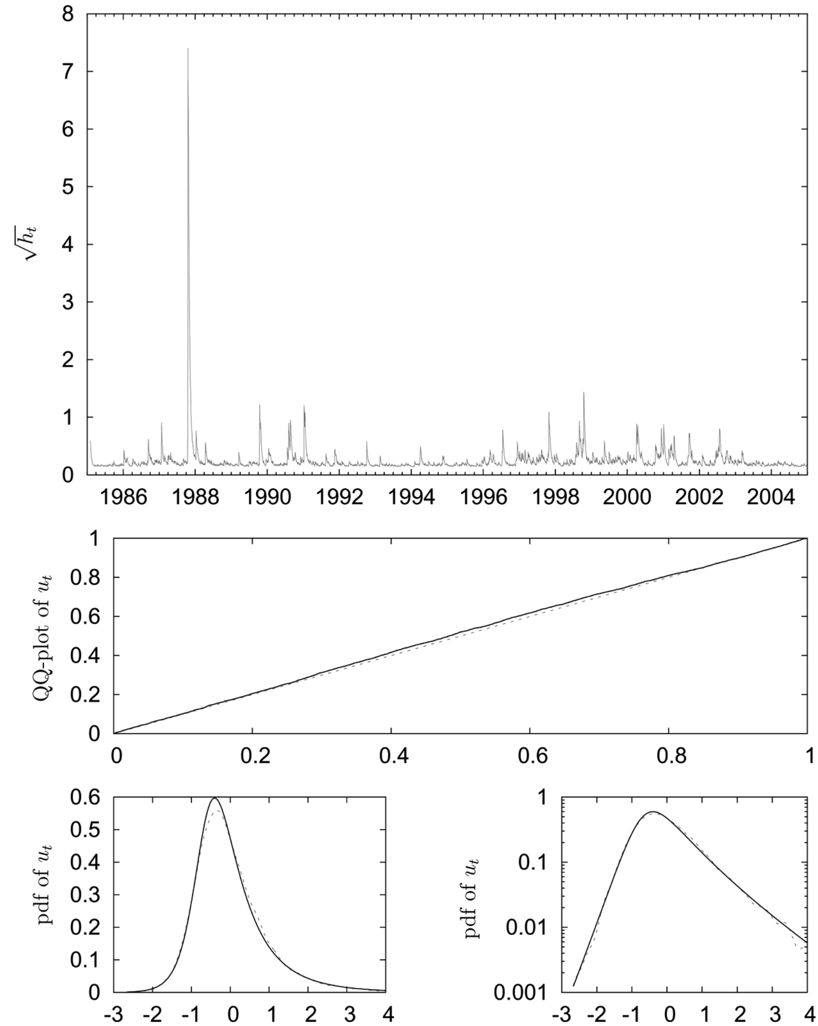

FIGURE 6 Diagnostics of the HAR–GARCH(1, 1)–NIG model for realized volatility. The upper panel presents the GARCH-filtered volatility of realized volatility series; the middle panel shows the quantile–quantile plot of the residuals, the bottom panel presents the kernel density estimates of the residuals (dashed line) and the estimated NIG density (solid line) in level (left) and log scales (right).

for realized volatility than for the logarithmic transform. The parameters in the mean equation of the HAR–NIG model differ from those in the standard HAR model. Although the persistence—measured by the sum of the autoregressive coefficients—reduces, the model-implied unconditional mean still matches the sample mean of the realized variance measures.

For both measures the overall preferred model turns out to be the HAR–GARCH specification with NIG-distributed innovations, suggesting

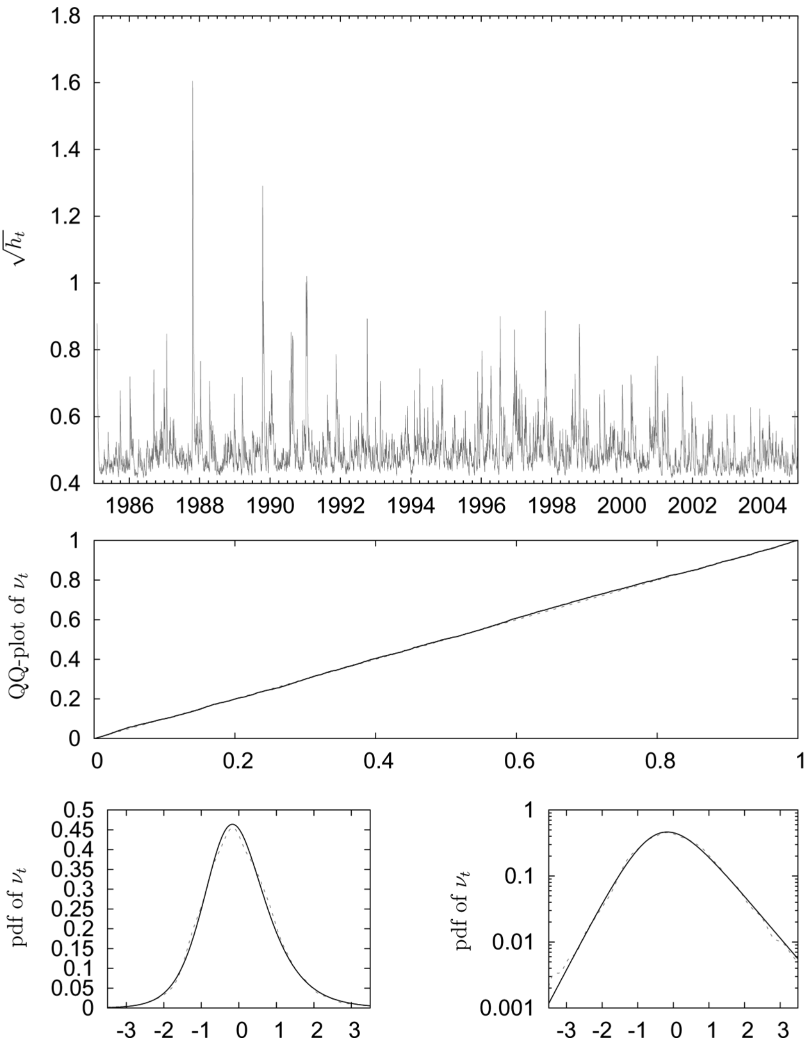

FIGURE 7 Diagnostics of the HAR–GARCH(1, 1)–NIG model for logarithmic realized variance. The upper panel presents the GARCH-filtered volatility of logarithmic realized variance series; the middle panel shows the quantile–quantile plot of the residuals, the bottom panel presents the kernel density estimates of the residuals (dashed line) and the estimated NIG density (solid line) in level (left) and log scales (right).

that both extensions to the standard long-memory model are important. This finding is in line with the GARCH literature for asset returns. For example, Verhoeven and McAleer (2004) and Mittnik et al. (1998, 2000) show that GARCH models with skewed and leptokurtic errors outperform their Gaussian counterparts.

Introducing the NIG distribution affects the GARCH-parameter estimates. For realized volatility the persistence is much less than under the Gaussian assumption since excess kurtosis can, in part, be captured by the shape parameter of the NIG distribution, α. Note also that, in comparison to the HAR–NIG model without GARCH specification, the shape parameters indicate less kurtosis.

Figures 6 and 7 demonstrate the adequacy of the HAR–GARCH(1, 1)–NIG model for realized volatility and logarithmic realized variance, respectively. Both skewness and the tail behavior of the innovations are well captured by the NIG distribution. The GARCH-filtered volatility series (top panels in Figures 6 and 7) show the clustering in the volatility of volatility. In fact, the time series pattern of the filtered series is, though less pronounced, similar to the characteristics observed in the measures of the volatility of realized volatility discussed in Section 2.2. This also applies to the observation that for logarithmic realized variance volatility clustering is less severe.

4. GAINS IN EFFICIENCY

Ignoring the presence of heteroskedasticity and non-Gaussianity in innovations leads to inefficient parameter estimates when estimating the (standard) HAR and ARFIMA models proposed in the literature. Inaccurate estimates do not only hamper their interpretation but also affect forecasting accuracy.

To assess the effects of explicitly allowing for heteroskedasticity and non-Gaussianity on efficiency, we conduct a simulation study. We generate 1,000 series, each with sample size 5,000, from a HAR-GARCH(1,1) model with standardized NIG-distributed innovations, using the estimates reported in Table 2. For each replication we consider the first 500, 1,250, 2,500, and, finally, all 5,000 data points which corresponds to about 2, 5, 10, and 20 years of daily data, respectively, with the latter approximating the sample size of our data set. From the simulated data we estimate the HAR specifications I–IV discussed in the previous section.

Tables 3 and 4 report the root mean square error of the parameter estimates for the four models for realized volatility and logarithmic realized variance, respectively, from the 1,000 simulation runs. Focusing first on the results for the parameters of the HAR mean equation, we see that the inclusion of the GARCH specification yields greater improvements in efficiency than just allowing for NIG-distributed innovations, indicating

TABLE 3 Efficiency results for realized-volatility models

		Mean eq.				Distribution		Variance eq.		
Obs.	Model	α_0	α_d	a_w	α_m	α	β	ω	α_1	β_1
500	I_S	0.0763	0.0818	0.1403	0.1238			0.0606		
	II_S	0.0723	0.0672	0.1216	0.1185			0.0061	0.1953	0.1055
	III_S	0.0940	0.0536	0.1245	0.1029	0.5634	0.4313	0.0650		
	IV_S	0.0493	0.0455	0.0825	0.0816	0.5355	0.4272	0.0031	0.1017	0.0471
1250	I_S	0.0392	0.0592	0.0952	0.0794			0.0589		
	II_S	0.0338	0.0428	0.0722	0.0679			0.0023	0.0802	0.0508
	III_S	0.0652	0.0358	0.0955	0.0660	0.5945	0.4501	0.0558		
	IV_S	0.0229	0.0278	0.0491	0.0470	0.2376	0.2032	0.0011	0.0383	0.0276
2500	I_S	0.0256	0.0409	0.0650	0.0565			0.0571		
	II_S	0.0223	0.0302	0.0499	0.0476			0.0013	0.0501	0.0370
	III_S	0.0559	0.0285	0.0837	0.0481	0.6082	0.4607	0.0535		
	IV_S	0.0149	0.0203	0.0333	0.0310	0.1466	0.1262	0.0007	0.0252	0.0195
5000	I_S	0.0179	0.0343	0.0537	0.0438			0.0529		
	II_S	0.0141	0.0219	0.0361	0.0335			0.0008	0.0325	0.0251
	III_S	0.0504	0.0223	0.0775	0.0387	0.6270	0.4743	0.0513		
	IV_S	0.0099	0.0139	0.0237	0.0220	0.1112	0.0938	0.0004	0.0166	0.0131

All entries report the root mean square error of parameter estimates for the different models. They are based on 1,000 simulations from the HAR–GARCH–NIG model as given in Table 2. "Obs." denotes the number of simulated observations of each simulation run and "Model" corresponds to the different models: I is a standard HAR model with Gaussian innovations; II also includes GARCH effects; III is a standard HAR model with (standardized) NIG innovations; and IV corresponds to the HAR–GARCH model with (standardized) NIG innovations.

that the incorporation of conditional heteroskedasticity is more relevant. Notably, although the HAR model with NIG errors is generally more efficient than the standard HAR specification, it has difficulties in estimating the constant and the parameter of the weekly volatility component. This is in line with our discussion in the previous section. However, when allowing for a GARCH specification, these problems vanish. As expected, given that it matches the data generation process, the HAR–GARCH–NIG model exhibits the strongest gains in parameter efficiency. The results suggest that allowing for volatility clustering leads to substantial efficiency gains. The implications of the NIG extension on efficiency is somewhat ambiguous. The forecasting experiment reported in the next section will shed additional light on this issue.

Turning our attention to the parameters of the variance equation in Table 3 we find that, for larger sample sizes, the efficiency results for the (wrongly specified) Gaussian HAR–GARCH model are surprisingly good. The results for the parameters of the NIG distribution should not be taken too seriously in the case of the standard HAR model, given the trade-off between the kurtosis induced by the GARCH specification and by the NIG distribution. However, for the HAR–GARCH specification we can conclude

TABLE 4 Efficiency results for logarithmic-realized-volatility models

		Mean eq.				Distribution		Variance eq.		
Obs.	Model	α_0	α_d	a_w	α_m	α	β	ω	α_1	β_1
500	I_L	0.0669	0.0595	0.1052	0.0956			0.2099		
	II_L	0.0645	0.0572	0.1017	0.0934			0.0725	0.3255	0.0617
	III_L	0.0635	0.0551	0.0978	0.0884	0.6787	0.4241	0.2105		
	IV_L	0.0602	0.0539	0.0938	0.0866	0.7530	0.7152	0.0645	0.2912	0.0529
1250	I_L	0.0303	0.0374	0.0609	0.0533			0.2088		
	II_L	0.0288	0.0357	0.0582	0.0521			0.0397	0.1814	0.0352
	III_L	0.0286	0.0338	0.0575	0.0486	0.2764	0.1459	0.2090		
	IV_L	0.0267	0.0326	0.0545	0.0482	0.3085	0.1712	0.0343	0.1563	0.0309
2500	I_L	0.0172	0.0274	0.0435	0.0382		0.2106			
	II_L	0.0168	0.0255	0.0406	0.0362			0.0200	0.0941	0.0244
	III_L	0.0165	0.0255	0.0406	0.0347	0.2097	0.1042	0.2105		
	IV_L	0.0157	0.0240	0.0377	0.0331	0.1746	0.1001	0.0159	0.0749	0.0204
5000	I_L	0.0110	0.0200	0.0314	0.0261			0.2105		
	II_L	0.0106	0.0184	0.0296	0.0249			0.0116	0.0571	0.0168
	III_L	0.0111	0.0176	0.0302	0.0252	0.1862	0.0866	0.2103		
	IV_L	0.0103	0.0166	0.0274	0.0231	0.1281	0.0697	0.0100	0.0481	0.0143

All entries report the root mean square error of parameter estimates for the different models. They are based on 1,000 simulations from the HAR-GARCH-NIG model as given in Table 2. "Obs." denotes the number of simulated observations of each simulation run and "Model" corresponds to the different models: I is a standard HAR model with Gaussian innovations; II also includes GARCH effects; III is a standard HAR model with (standardized) NIG innovations; and IV corresponds to the HAR–GARCH model with (standardized) NIG innovations.

that a sufficiently large sample size is required to accurately estimate the distributional parameters.

5. PREDICTION

In order to assess the relevance of allowing for conditional heteroscedasticity of realized volatility for out-of-sample forecasts we consider the period from December 13, 1988 to December 30, 2004 of the futures data, providing us with 4,040 forecasts. We estimate all four model specifications from the first 1,000 observations (January 1, 1985 to December 12, 1988) and recursively construct one-step-ahead realized-volatility forecasts. In each recursion we reestimate, expanding the data set by one observation.

Since the standard deviation of realized variance is the relevant variable in many applications, e.g., for risk assessment and management applications, we evaluate the predictive performance of the different models in terms of their ability to forecast realized volatility. Moreover, to assure comparability in the forecasting performance of the realized volatility/logarithmic realized variance models, we explicitly account for the bias induced by the transformation of the logarithmic realized-variance forecasts. In particular, the realized-volatility forecasts of the logarithmic-realized-variance models are computed as

$$\sqrt{\widehat{RV_{t\,|\,t-1}}} = E\big(\sqrt{RV_{t\,|\,t-1}}\big) = \exp\left\{\frac{1}{2}\log \widehat{RV_{t\,|\,t-1}}\right\} M\left(\frac{1}{2}h_t\right), \tag{31}$$

with $\log \widehat{RV_{t\,|\,t-1}}$ denoting the one-step-ahead logarithmic realized-variance forecast for day t constructed from information available up to time $t-1$, and $M(\cdot)$ denoting the moment generating function of the assumed distribution. For the standardized normal distribution $M(\zeta) = \exp\{\frac{\zeta^2}{2}\}$ $\forall \zeta \in \mathbb{R}$, and for the standardized NIG distribution

$$M(\zeta) = \exp\left\{\frac{(\alpha^2-\beta^2)}{\alpha^2}\left(-\beta\zeta + (\alpha^2-\beta^2)\left(1 - \frac{\sqrt{\alpha^2-(\beta+\zeta)^2}}{\sqrt{\alpha^2-\beta^2}}\right)\right)\right\}, \tag{32}$$

for all $\zeta \in [-(\alpha+\beta), \alpha-\beta]$.

To evaluate the predictive performance of the different volatility models we follow Andersen and Bollerslev (1998) and Andersen et al. (2003), among others, and compute the R^2 statistic from the Mincer–Zarnowitz regressions of observed realized volatility on the corresponding forecasts. In addition to the R^2 statistic, we also report the root mean square forecast error (RMSE), the mean absolute error (MAE), and the root mean squared percentage error (RMSPE). Assigning more weight to large volatility-forecast errors, the RMSE will be of particular interest under a risk-management perspective.[12]

The results of the forecasting exercise are presented in Table 5. The first four columns report the criteria for the one-step-ahead point forecast evaluation criteria for the different HAR models. The results show that, regardless of the transformation considered, all four statistics yield more or less the same performance rankings for the four models. The best model for forecasting daily volatility is clearly the HAR model with GARCH specification and Gaussian distributed innovations, and we find the HAR–GARCH–NIG model to perform second best. Surprisingly, the forecasts obtained from the HAR model with NIG assumption perform worst.

Comparing the evaluation criteria across all eight specifications shows that generally the logarithmic-realized-variance models perform marginally better than their realized-volatility counterparts. The Gaussian HAR–GARCH model formulated for logarithmic realized variance provides the best point forecasts according to all criteria, with the sole exception of the RMSPE favoring the same model but expressed for realized volatility.

[12]Note, however, that under the null hypothesis of the Mincer–Zarnowitz test for unbiasedness of forecasts, the RMSE is a homogeneous function of the regression coefficient.

TABLE 5 One-step-ahead forecast evaluation

Model	R^2	RMSE	MAE	RMSPE	$S(f_{t\mid t-1}, y_t)$
I_S	0.4848	0.3161	0.1926	0.3298	−1725.9
II_S	0.5109	0.3084	0.1849	0.2925	−98.8
III_S	0.5034	0.3173	0.1996	0.3422	538.8
IV_S	0.5074	0.3091	0.1870	0.3034	941.3
I_L	0.4972	0.3108	0.1869	0.3165	865.9
II_L	0.5228	0.3033	0.1822	0.2934	931.5
III_L	0.5091	0.3126	0.1884	0.3272	1022.7
IV_L	0.5169	0.3052	0.1842	0.2961	1068.7

"Model" represents the different model specifications: I is a standard HAR model with Gaussian innovations; II also includes GARCH effects; III is a standard HAR model with (standardized) NIG innovations; and IV corresponds to the HAR–GARCH model with (standardized) NIG innovations. The indices S and L denote the HAR models formulated for realized volatility and logarithmic realized variance, respectively. The reported R^2 are the regression coefficients of realized volatility on a constant and volatility forecasts. $S(f_{t\mid t-1}, y_t)$ are the logarithmic scores as defined in (34) of the one-step-ahead density forecasts.

In summary, the forecasting evaluations show that allowing for time-varying volatility of realized volatility/logarithmic realized variance improves the accuracy of volatility point forecasts and should, therefore, be important for risk management applications, such as Value-at-Risk calculations. In contrast, permitting skewness and leptokurtosis in the innovation distribution does not seem to help in point forecasting. This is somehow in line with the conclusions drawn from our efficiency simulations discussed in Section 4. Moreover, forecasting realized volatility based on the logarithmic-realized-variance models results in an improvement in forecast accuracy.

Although volatility point forecasting is, in general, of primary interest, interval or density prediction can also be of interest. First, when forecasting returns, the uncertainty associated with the volatility estimates carries over to the uncertainty of return forecasts. Second, the increasing availability of volatility-based derivatives renders volatility density forecasts to be more and more relevant (see, for example, Corradi et al., 2005, 2006, who develop an estimator of the predictive density of the integrated variance—measured in terms of realized variance).

Because time-variation in the volatility of realized volatility implies time-varying conditional densities, we also evaluate the accuracy of the one-step-ahead density forecasts using a method proposed by Diebold et al. (1998), which is based on Rosenblatt (1952). With $\{f_{t\mid t-1}(y_t)\}_{t=1}^{n}$ denoting a sequence of n density forecasts constructed from a parametric forecasting model using information available up to time $t-1$, the probability-integral

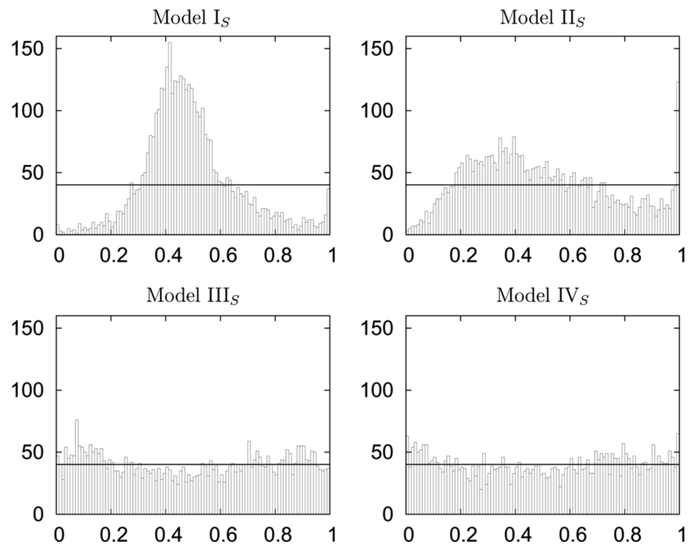

FIGURE 8 Probability-integral transforms of density forecasts based on the four realized-volatility models: Model I_S refers to the standard HAR model with Gaussian innovations; Model II_S also includes GARCH effects; Model III_S is a standard HAR model with standardized NIG innovations; and Model IV_S corresponds to the HAR–GARCH(1, 1) model with standardized NIG innovations.

transform z_t is defined as

$$z_t = \int_{-\infty}^{y_t} f_{t\,|\,t-1}(u)du. \tag{33}$$

The quantity z_t represents the cumulative distribution function based on the forecasted density evaluated at the realization y_t, i.e., at $\sqrt{RV_t}$. If the density forecast is correct, then the sequence of probability-integral transforms, $\{z_t\}_{t=1}^{n}$, is IID uniformly distributed on the unit interval (see Diebold et al., 1998).

Figures 8 and 9 present the histograms of the corresponding probability-integral transforms of the four models for realized and logarithmic realized variance, respectively. It turns out that for both transformations of realized variance the z_t of the models with Gaussian innovations are far from being uniformly distributed, although the incorporation of the volatility of realized volatility leads to some improvement. The inability to capture skewness is clearly illustrated by the graphs. In contrast, the NIG-based HAR models provide very accurate density forecasts with the HAR–GARCH specification being somewhat superior. These results strongly favor the NIG extension of HAR and HAR–GARCH models. Moreover, the probability-integral transforms of the logarithmic-realized-variance models are closer to being uniformly distributed than the corresponding series of the realized-volatility models.

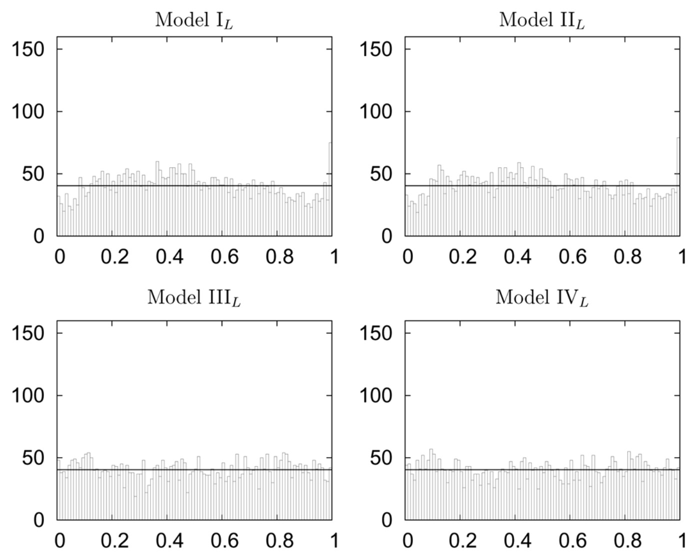

FIGURE 9 Probability-integral transforms of density forecasts based on the four logarithmic-realized-variance models: Model I_L refers to the standard HAR model with Gaussian innovations; Model II_L also includes GARCH effects; Model III_L is a standard HAR model with standardized NIG innovations; and Model IV_L corresponds to the HAR–GARCH(1, 1) model with standardized NIG innovations.

To better quantify these visual results, we additionally compute the logarithmic score of the density forecasts defined as (see, for example, Matheson and Winkler, 1976)

$$S(\{f_{t|t-1}\}_{t=1}^{n}, \{y_t\}_{t=1}^{n}) = n^{-1} \sum_{t=1}^{n} \log f_{t\,|\,t-1}(y_t), \tag{34}$$

with $f_{t\,|\,t-1}$ denoting again the predicted density, and y_t the realized value of the variable to be forecast. Using the logarithmic scoring rule allows us to explicitly rank the competing density forecasts. A high score, implying a high likelihood for the observed y_t-realization, indicates a better predictive performance.

The logarithmic scores are presented in the last column of Table 5.[13] The results confirm the conclusions drawn from the histograms of the probability-integral transform series in Figures 8 and 9, i.e., the HAR–GARCH(1, 1)–NIG models provide the best density forecasts; and the density forecast based on the logarithmic-realized-variance models outperform all of their realized-volatility counterparts.

[13]Since the density forecasting performance is also evaluated in terms of realized volatility, the logarithmic scores of the logarithmic-realized-variance models are given by $S(\{f_{t\,|\,t-1}\}_{t=1}^{n}, \{\sqrt{RV_t}\}_{t=1}^{n}) = n^{-1}\sum_{t=1}^{n} \log\left(\frac{f_{t\,|\,t-1}(\log RV_t)}{2\sqrt{RV_t}}\right)$.

The fact that NIG-based HAR/HAR–GARCH models provide better realized volatility density forecasts, but at the same time perform worse than the Gaussian HAR/HAR–GARCH models in point forecasting may appear counterintuitive. However, it is widely reported in the (point) forecasting literature that more parsimonious, though potentially misspecified models, may generate more accurate point forecasts (see, e.g., Clements and Hendry, 1998 or Lütkepohl, 1993).[14]

6. CONCLUSION

We have shown that the commonly used reduced-form realized-volatility models, such as the ARFIMA or HAR models are better characterized by non-Gaussian innovations and time-varying volatility that might be partly attributed to the time-variation in the volatility of the realized volatility estimator. Specifically, we favor a (semi)-heavy-tailed distribution, such as the NIG distribution, and more importantly a model that allows for GARCH-type clustering in the volatility of realized volatility or logarithmic realized variance. In-sample estimation results show an overwhelming superiority of the HAR–GARCH model with NIG distributed innovations. It appears to be important to incorporate the volatility of volatility and to specify an adequate distribution of the error terms, when modeling realized volatility or the logarithm of realized variance. Both extensions have been considered in Bollerslev et al. (2006) as part of a highly accurate simultaneous three-equation model for returns and realized variations.

Investigating the implications of the two proposed extensions for the efficiency of the parameter estimates we conclude that the time-varying volatility of realized volatility/logarithmic realized variance is of importance and a GARCH-type extension should be incorporated. Our forecasting experiments suggest that for accurate point forecasts the GARCH specification is more important than allowing for a fat-tailed and possibly skewed distribution. For risk management applications, such as Value-at-Risk calculations, the time-variation of the volatility of volatility should, therefore, be incorporated, whereas the specification of a non-Gaussian distribution is not necessarily required. However, whenever interval or density forecasts of realized volatility are the main focus—for example when pricing or assessing the risk of volatility derivatives—a flexible distribution, such as the NIG, should be specified in addition to a

[14] Much less is known about the effects of model parsimony on the accuracy of density forecasts. A small simulation study comparing the point and density forecast accuracy of the pure Gaussian–HAR model and the HAR–NIG model, with the latter being the correct data generating process corroborate our empirical observation. In particular, we find strong superiority of the HAR–NIG model in density forecasting, but slightly less accurate point forecasts than those produced by the Gaussian–HAR model.

GARCH component. Moreover, accounting for the uncertainty associated with the realized-volatility forecasts may also improve the accuracy of return forecasts.

The relevance of our findings for extended realized-volatility models, for example models including leverage effects (see Martens et al., 2004), will be subject of future research. We expect, however, that, although the incorporation of the leverage effect may lead to some reduction in the skewness of the innovations, the specification of a non-Gaussian distribution might still be necessary in order to reproduce the observed excess kurtosis. We also expect that the time-variation in volatility of realized volatility will remain, since part of the time-variation can be attributed to the volatility of the realized volatility estimator.

ACKNOWLEDGMENTS

The work of Corsi was supported by the National Centre of Competence in Research Financial Valuation and Risk Management (NCCR FINRISK) of the Swiss National Science Foundation. Stefan Mittnik and Christian Pigorsch acknowledge financial support from the Deutsche Forschungsgemeinschaft (SFB 386) and the Fulbright Commission. Uta Pigorsch acknowledges financial support from the German Academic Exchange Service (doctoral scholarship).

We thank two anonymous referees for helpful comments, as well as Ole E. Barndorff-Nielsen, Neil Shephard, and the participants at the 2005 NSF/NBER Time Series Conference in Heidelberg, at the Eurandom 2006 Workshop on Risk Measures & Risk Management for High-Frequency Data in Eindhoven, the 2006 Karlsruhe Econometrics Workshop on Risk Assessment: Decisions in Banking and Finance, and at the International Conference on High-Frequency Finance in Konstanz.

REFERENCES

Agiakloglou, C., Newbold, P., Wohar, M. (1993). Bias in an estimator of the fractional difference parameter. *Journal of Time Series Analysis* 14:235–246.

Aït-Sahalia, Y., Mykland, P. A., 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.

Andersen, T. G., Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns. *The Journal of Finance* 52:975–1005.

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., Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics* 61:43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P. (2001b). The distribution of realized 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.

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.

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

Bai, X., Russell, J. R., Tiao, G. C. (2003). Kurtosis of GARCH and stochastic volatility models with non-normal innovations. *Journal of Econometrics* 114:349–360.

Baillie, R. T., Chung, C.-F., Tieslau, M. A. (1996). Analysing inflation by the fractionally integrated ARFIMA–GARCH model. *Journal of Applied Econometrics* 11:23–40.

Bandi, F. M., Russell, J. R. (2005). Microstructure noise, realized volatility, and optimal sampling. Forthcoming. *Review of Economic Studies*.

Banerjee, A., Urga, G. (2005). Modelling structural breaks, long memory and stock market volatility: an overview. *Journal of Econometrics* 129:1–2.

Barndorff-Nielsen, O. E. (1997). Normal inverse gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics* 24:1–13.

Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type. *Finance and Stochastics* 2:41–68.

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

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

Barndorff-Nielsen, O. E., Shephard, N. (2002b). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics* 17:457–477.

Barndorff-Nielsen, O. E., Shephard, N. (2002c). *Lévy Processes – Theory and Applications*. Boston: Birkhauser, chap. Modelling by Lévy Processess for Financial Econometrics, pp. 283–318.

Barndorff-Nielsen, O. E., Shephard, N. (2003). Realised power variation and stochastic volatility models. *Bernoulli* 9:243–265.

Barndorff-Nielsen, O. E., 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., Shephard, N. (2004b). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* 2:1–37.

Barndorff-Nielsen, O. E., Shephard, N. (2005). How accurate is the asymptotic approximation to the distribution of realised variance? In: Andrews, D., Powell, J., Ruud, P., Stock, J., eds. *Identification and Inference for Econometric Models. A Festschrift for Tom Rothenberg*. Cambridge: Cambridge University Press, pp. 306–331.

Barndorff-Nielsen, O. E., Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4:1–30.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., Shephard, N. (2006). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. Tech. rep., Oxford University.

Beran, J. (1994). *Statistics for Long-Memory Processes*. Chapman & Hall, New York.

Bollerslev, T., Kretschmer, U., Pigorsch, C., Tauchen, G. (2006). A discrete-time model for daily S&P 500 returns and realized variations: jumps and leverage effects. Forthcoming. *Journal of Econometrics*.

Calvet, L., Fisher, A. (2002). Multifractality in asset returns: theory and evidence. *Review of Economics and Statistics* 84:381–406.

Calvet, L., Fisher, A. (2004). How to forecast long-run volatility: regime switching and the estimation of multifractal processes. *Journal of Financial Econometrics* 2:49–83.

Clements, M. P., Hendry, D. F. (1998). *Forecasting Economic Time Series*. Cambridge: Cambridge University Press.

Comte, F., Renault, E. (1998). Long-memory in continuous-time stochastic volatility models. *Mathematical Finance* 8:291–323.

Corradi, V., Distaso, W., Swanson, N. R. (2005). Predictive density estimators for daily volatility based on the use of realized measures. Forthcoming. *Journal of Econometrics*.

Corradi, V., Distaso, W., Swanson, N. R. (2006). Predictive inference for integrated volatility. Tech. rep., Queen Mary, University of London and Rutgers University.

Corsi, F. (2004). A simple long memory model of realized volatility. Tech. rep., University of Southern Switzerland.

Corsi, F., Zumbach, G., Müller, U. A., Dacorogna, M. (2001). Consistent high-precision volatility from high-frequency data. *Economic Notes* 30:183–204.

Curci, G., Corsi, F. (2003). A discrete sine transform approach for realized volatility measurement. Tech. Rep. 44, National Centre Of Competence in Research Financial Valuation and Risk Management.

Davidson, J., Teräsvirta, T. (2002). Long memory and nonlinear time series. *Journal of Econometrics* 110:105–112.

Diebold, F. X., Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics* 105:131–159.

Diebold, F. X., Gunther, T. A., Tay, A. S. (1998). Evaluating density forecasts with applications to financial risk management. *International Economic Review* 39:863–883.

Ding, Z., Granger, C. W. J. (1996). Modeling volatility persistence of speculative returns: a new approach. *Journal of Econometrics* 73:185–215.

Ding, Z., Granger, C. W. J., Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1:83–106.

Doornik, J. A., Ooms, M. M. (2003). Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. *Computational Statistics and Data Analysis* 42:333–348.

Engle, R. F., Lee, G. (1999). A permanent and transitory component model of stock return volatility. *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*, pp. 475–497.

Gallant, A., Hsu, C., Tauchen, G. (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *The Review of Economics and Statistics* 81:617–631.

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. (2005). Bootstrapping realized volatility. Tech. rep., University of Montreal.

Gourieroux, C., Jasiak, J. (2001). *Financial Econometrics: Problems, Models, and Methods*. Princeton University Press.

Granger, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics* 14:227–238.

Granger, C. W. J., Joyeux, R. (1980). An introduction to long-range time series models and fractional differencing. *Journal of Time Series Analysis* 1:15–30.

Granger, C. W. J., Teräsvirta, T. (1999). A simple nonlinear time series model with misleading linear properties. *Economics Letters* 62:161–165.

Granger, C. W. J., Hyung, N. (2004). Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *Journal of Empirical Finance* 11:399–421.

Granger, C. W. J., Spear, S., Ding, Z. (2000). Stylized facts on the temporal and distributional properties of daily data from speculative markets. In: Chan, W. S., Lin, W. K., Tong, H., eds. *Statistics and Finance: An Interface*. London: Imperial College Press, pp. 97–120.

Hansen, P. R., Lunde, A. (2006a). Consistent ranking of volatility models. *Journal of Econometrics* 131:97–121.

Hansen, P. R., Lunde, A. (2006b). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics* 24:127–161.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika* 68:165–176.

Koopman, S. J., Jungbacker, B., Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance* 12:445–475.

Lütkepohl, H. (1993). *Introduction to Multiple Time Series Analysis*. Berlin: Springer-Verlag.

Lux, T., Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature* 397:498–500.

Maheu, J. M., McCurdy, T. (2002). Nonlinear features of realized FX volatility. *Review of Economics and Statistics* 84:668–681.

Mandelbrot, B., Fisher, A., Calvet, L. (1997). A multifractal model of asset returns. Cowles foundation discussion papers, Cowles Foundation, Yale University.

Martens, M., Zein, J. (2004). Predicting financial volatility: high-frequency time-series forecasts vis-à-vis implied volatility. *Journal of Futures Markets* 24:1005–1028.

Martens, M., Dijk, D. v., de Pooter, M. (2004). Modeling and forecasting S&P500 volatility: long memory, structural breaks and nonlinearity. Tech. rep., Erasmus University Rotterdam.

Matheson, J., Winkler, R. L. (1976). Scoring rules for continuous probability distributions. *Management Science* 22:1087–1096.

Merton, R. C. (1980). On estimating the expected return on the market : an exploratory investigation. *Journal of Financial Economics* 8:323–361.

Mikosch, T., Starica, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *The Review of Economics and Statistics* 86:378–390.

Mittnik, S., Paolella, M. S., Rachev, S. T. (1998). Unconditional and conditional distributional models for the Nikkei index. *Asia-Pacific Financial Markets* 5:99–128.

Mittnik, S., Paolella, M. S., Rachev, S. T. (2000). Diagnosing and treating the fat tails in financial return data. *Journal of Empirical Finance* 7:389–416.

Müller, U. A., Dacorogna, M. M., Davé, R. D., Olsen, R. B., Pictet, O. V., von Weizsäcker, J. E. (1997). Volatilities of different time resolutions – analyzing the dynamics of market components. *Journal of Empirical Finance* 4:213–239.

Oomen, R. C. (2004), Modelling realized volatility when returns are serially correlated. Tech. Rep. SP II 2004-11, Social Science Research Center, Berlin.

Parke, W. (1999). What is fractional integration?. *The Review of Economics and Statistics* 81:632–638.

Pong, S., Shackleton, M. B., Taylor, S. J., Xu, X. (2004). Forecasting currency volatility: a comparison of implied volatilities and AR(FI)MA models. *Journal of Banking & Finance* 28:2541–2563.

Robinson, P. (2003). *Time Series with Long Memory.* Oxford University Press.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Annals of Mathematical Statistics* 23:470–472.

Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics* 53:165–188.

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

Verhoeven, P., McAleer, M. (2004). Fat tails and asymmetry in financial volatility models. *Mathematics and Computers in Simulation* 64:351–361.

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.