No. 2005/33

The Volatility of Realized Volatility

Fulvio Corsi, Uta Kretschmer,
Stefan Mittnik, and Christian Pigorsch

Center for Financial Studies
an der Universität Frankfurt ▪ Mertonstraße 17-21 ▪ D-60325 Frankfurt am Main
Telefon: +49 (0)69 798-30050
Fax: +49 (0)69 798-30077 ▪ http://www.ifk-cfs.de ▪ E-Mail: info@ifk-cfs.de

Center for Financial Studies

The *Center for Financial Studies* is a nonprofit research organization, supported by an association of more than 120 banks, insurance companies, industrial corporations and public institutions. Established in 1968 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia.

The CFS Working Paper Series presents the result of scientific research on selected topics in the field of money, banking and finance. The authors were either participants in the Center´s Research Fellow Program or members of one of the Center´s Research Projects.

If you would like to know more about the *Center for Financial Studies*, please let us know of your interest.

Prof. Dr. Jan Pieter Krahnen | Prof. Volker Wieland, Ph.D.

Center for Financial Studies
an der Universität Frankfurt ▪ Mertonstraße 17-21 ▪ D-60325 Frankfurt am Main
Telefon: +49 (0)69 798-30050
Fax: +49 (0)69 798-30077 ▪ http://www.ifk-cfs.de ▪ E-Mail: info@ifk-cfs.de

CFS Working Paper No. 2005/33

The Volatility of Realized Volatility*

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

This Version: November 28, 2005

Abstract:
Using unobservable conditional variance as measure, latent–variable approaches, such as GARCH and stochastic–volatility models, have traditionally been dominating the empirical finance literature. In recent years, with the availability of high–frequency financial market data modeling realized volatility has become a new and innovative research direction. By constructing "observable" or realized volatility series from intraday transaction data, the use of standard time series models, such as ARFIMA models, have become a promising strategy for modeling and predicting (daily) volatility. In this paper, we show that the residuals of the commonly used time–series models for realized volatility exhibit non–Gaussianity and volatility clustering. We propose extensions to explicitly account for these properties and assess their relevance when modeling and forecasting realized volatility. In an empirical application for S&P500 index futures we show that allowing for time–varying volatility of realized volatility leads to a substantial improvement of the model's fit as well as predictive performance. Furthermore, the distributional assumption for residuals plays a crucial role in density forecasting.

JEL Classification: C22, C51, C52, C53

Keywords: Finance, Realized Volatility, Realized Quarticity, GARCH, Normal Inverse Gaussian Distribution, Density Forecasting

* Uta Kretschmer acknowledges financial support from the German Academic Exchange Service (doctoral scholarship). Stefan Mittnik and Christian Pigorsch acknowledge financial support from the Deutsche Forschungsgemeinschaft (SFB 386) and the Fulbright Commission. We thank Neil Shephard and the participants at the 2005 NSF/NBER Time Series Conference in Heidelberg for helpful comments.

[1] Institute of Finance, University of Lugano, Via G. Buffi 13, CH-6904 Lugano, Switzerland.
E-mail: fulvio.corsi@lu.unisi.ch.

[2] Department of Economics, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany.
E-mail: uta.kretschmer@uni-bonn.de.

[3] Corresponding author, Email: finmetrics@stat.uni-muenchen.de. Department of Statistics, University of Munich, Akademiestr 1/I, D-80799 München, Germany; Center for Financial Studies, Frankfurt; Ifo Institute for Economic Research, Munich.

[4] Department of Statistics, University of Munich, Akademiestr 1/I, D-80799 München, Germany.
E-mail: pigorsch@stat.uni-muenchen.de.

Non–technical Summary

The phenomenon that the volatility of financial asset returns commonly exhibits serial correlation has lead to enormous research activities on modeling and forecasting volatility. For daily data, asset return volatility is typically modeled via GARCH–type or stochastic–volatility processes which treat volatility as a latent variable. The increasing availability of high–frequency data gives rise to a more straightforward approach to modeling volatility by constructing daily time series. This allows us to treat volatility as an "observed" rather than latent variable to which we can apply standard time–series modeling techniques. Also, because summing over high–frequency intra–daily squared returns yields a consistent estimator of the actual (daily) volatility, the so–called *realized volatility*, the forecasting performance of estimated models can be assessed more adequately as is the case when using squared daily returns as volatility measure. The autoregressive fractionally integrated moving average (ARFIMA) and, due to its simple estimation, the heterogeneous autoregressive (HAR) model are the most popular models in the current literature. They are capable of capturing the observed long–memory of volatility, and empirically outperform GARCH and stochastic–volatility models in terms of their forecasting ability.

In this paper, we investigate more closely the properties of realized volatility and the common realized volatility models. Using S&P500 index futures data we find that the (non–transformed as well as the logarithmic) realized volatility series exhibit non–Gaussianity and conditional heteroskedasticity features that have been neglected in the current research on realized volatility. To accommodate these properties, we extend models for realized volatility by replacing the Gaussian with the more flexible *normal inverse Gaussian* (NIG) distribution to allow for fat–tailedness and skewness. In addition, we specify a GARCH process to account for the clustering in the squared residuals of the realized volatility model. Our empirical in–sample results suggest that both extensions strongly improve the model's fit. Of particular importance for risk assessment and risk management, however, is the predictive performance of volatility models. Evaluating this for the proposed models over different horizons, we find that explicitly modeling the volatility of realized volatility leads to better point and interval forecasts. Moreover, the additional specification of the normal inverse Gaussian distribution further improves the density forecasts.

Nichttechnische Zusammenfassung

Die Beobachtung, dass die Volatilität von Aktienrenditen serielle Korrelation aufweist, hat zu starkem Forschungsinteresse auf dem Gebiet der Modellierung und Prognose der Volatilität geführt. Bei Verwendung von Tagesdaten wird die Volatilität üblicherweise mit Hilfe von GARCH Modellen oder Stochastischen Volatilitätsmodellen beschrieben, welche die Volatilität als unbeobachtbare Variable behandeln. Die zunehmende Verfügbarkeit von hochfrequenten Finanzmarktdaten ermöglicht jedoch die Konstruktion von täglichen Volatilitätsmaßen, welche nun als "beobachtbare" und nicht mehr nur als latente Variable behandelt werden können. Gängige Verfahren aus der Zeitreihenmodellierung können daher ohne weiteres zur Modellierung der Volatilität angewandt werden. Da die Summe über quadrierte Intra–Tagesrenditen ein konsistenter Schätzer für die wahre (tägliche) Volatilität ist, kann die sogenannte *realisierte Volatilität* auch zur—im Vergleich zu dem bisher üblichen täglichen Volatilitätmaß, der quadrierten Tagesrendite—exakteren Evaluierung der Prognosegüte geschätzter Modelle genutzt werden. Die bekanntesten Modelle für die realisierte Volatilität sind das autoregressive fraktionell integrierte moving average (ARFIMA) Modell und aufrund seiner einfachen Schätzung das heterogene autoregressive (HAR) Modell. Empirische Studien haben gezeigt, dass diese Modelle nicht nur das lange Gedächtnis der Volatilität reproduzieren können, sondern auch zu wesentlichen Verbesserungen in der Prognosegüte gegenüber den GARCH– oder stochastischen Volatilitätsmodellen führen.

In dieser Studie werden die Eigenschaften der realisierten Volatilität und der realisierten Volatilitätsmodelle genauer untersucht. Unter Verwendung von S&P500 Index Futures Daten wird gezeigt, dass sowohl die (untransformierte als auch die logarithmierte) realisierte Volatilität nicht normalverteilt ist und bedingte Heteroskedastizität aufweist. Diese Eigenschaften, insbesondere die der zeitvariierenden Volatiliät der realisierten Volatilität, wurden in der bisherigen Literatur vernachlässigt. Zur Berücksichtigung dieser Eigenschaften werden in dieser Studie zwei Erweiterungen der existierenden Modelle betrachtet. Um Schiefe und Kurtosis zu berücksichtigen, wird die Normalverteilung durch eine flexiblere Verteilung, der *Normal Invers Gauss'schen* Verteilung, ersetzt. Des weiteren werden die Volatilitätscluster in den quadrierten Residuen der realisierten Volatilitätsmodelle durch eine GARCH Spezifikation modelliert. Die empirischen Ergebnisse dieser Studie zeigen, dass beide Erweiterungen zu einer signifikanten Verbesserung in der Modellgüte führen. Wesentlich relevanter für die Einschätzung und das Management von Risiken ist jedoch die Prognosegüte von Volatilitätsmodellen. Eine Prognoseevaluierung der erweiterten Modelle über verschiedene Prognosehorizonte zeigt, dass die explizite Modellierung der Volatilität der realisierten Volatilität zu besseren Punkt– und Intervallprognosen führt. Die Berücksichtigung der Normal Invers Gauss'schen Verteilung verbessert zudem die Güte von Dichteprognosen.

1 Introduction

Volatility plays an important role both in theoretical developments in finance as well as in practical applications. With the availability of high–frequency data the 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 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 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 paper, 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 generalized autoregressive conditional heteroskedasticity (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.

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. The time–varying behavior has also been reported in Barndorff-Nielsen and Shephard (2005b) 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 paper is organized as follows. The next section briefly reviews the construction of measures of realized volatility and the volatility of the realized volatility error; and then describes the S&P500 index futures data set used in the empirical application. Section 3 presents a short description of the standard realized–volatility models considered in this paper, discusses the model extensions we propose, and presents in–sample estimation results. In Section 4 we perform a simulation study to assess efficiency implications of the proposed extensions. Section 5 presents 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. (2002, 2005) and Barndorff-Nielsen and Shephard (2003, 2004b, 2005a,b). This allows us to compute an approximation to the volatility of the realized volatility 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, $\tau_0 = 0 \leq \tau_1 \leq \ldots \leq \tau_n = t$, is defined by

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

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

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

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, 2005b) 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 the *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, 2005a) suggest different estimators of integrated quarticity. The *realized fourth-power variation* or *realized quarticity*, defined as

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

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

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

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

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

has been proposed in Andersen et al. (2005).

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 derive an approximate 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,

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

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

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

2.2 Data

Our empirical application is based on tick–by–tick transaction prices of S&P500 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 common practice to handle this trade–off by summing returns over 5 or 30 minutes (see, for example, Andersen et al., 2001b, 2005; Barndorff-Nielsen and Shephard, 2004b).[2] Given the high liquidity of the S&P500 index futures market our realized–variance and quarticity measures are based on five–minute returns, an interval for which we assume that market–microstructure effects are 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, $\sqrt{RV_t}$, and the three volatility measures of the realized volatility estimator defined in (13)-(15). Figure 1 shows plots of the four series as well as their sample autocorrelation and partial autocorrelation functions. Table 1 reveals that the distribution of realized volatility is fat tailed and slightly skewed. Similar but much less pronounced patterns are also found for the logarithmic transform of realized volatility.[3] Figure 1 as well as the Ljung–Box statistics reported in Table 1 indicate strong autocorrelation in realized volatility. The autocorrelation function exhibits a hyperbolic decay, a finding that is in line with the widely reported long–memory behavior of volatility and also observed in the variance/volatility implied by estimated GARCH models.

The unconditional distributions of all three measures of the volatility of the realized volatility estimator 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 larger values for high (absolute) intra–day returns and focusing on one period at a time. Figure 1 also shows that all volatility–of–volatility series have long–memory, and that their values are high when realized volatility is high (cf. Barndorff-Nielsen and Shephard, 2005b). Most importantly, all three measures exhibit clear time–variation and volatility clustering. Similar patterns are also found for the volatility of the log–transformed realized volatility estimator.[4]

---

[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); Zhang et al. (2005), Areal and Taylor (2002), Bandi and Russell (2005), Corsi et al. (2001), and Curci and Corsi (2003), among others.

[3] Sample skewness and kurtosis of the logarithmic realized volatility are 0.5950 and 4.8070, respectively; and the Kolmogorov–Smirnov test rejects the null of Gaussianity (p–value=0.0087). Our results for logarithmic realized volatility are in line with Andersen et al. (2001a,b), who note that the assumption of Gaussianity is better in the case of logarithmic realized volatility, but differ from those reported in Thomakos and Wang (2003), who perform tests on Gaussianity, however, using a much shorter sample period.

[4] However, fluctuations over time are somewhat smaller, and the volatility measures have an asymptotic lower bound, as has been shown by Barndorff-Nielsen and Shephard (2005b). Results for log–transformed realized volatility are available upon request.

3 Modeling Realized Volatility

Volatility modeling plays a prominent role in the financial–econometrics and risk–management literature. 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 briefly summarize the recently developed approaches to modeling realized volatility before discussing the modeling extensions examined.

3.1 Conventional Realized–Volatility Models

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

$$\phi(L)(1-L)^d(\sqrt{RV}_t - \mu) = \psi(L)u_t, \tag{16}$$

with d denoting the fractional difference parameter, $\phi(L) = 1-\phi_1 L - \ldots - \phi_p L^p$ and $\psi(L) = 1+\psi_1 L + \ldots \psi_q L^q$. Typically, u_t is assumed to be a Gaussian white noise process. Keeping this assumption (regardless of the transformation of realized volatility considered), several papers have adopted and extended this model by including, for example, leverage effects (and other nonlinearities) or 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.[5] In the context of interval forecasting, we expect the distributional assumptions for the error terms to be of particular importance.

An alternative to ARFIMA, 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 horizons. Defining the k–period realized volatility component by the sum of the single–period realized volatilities, i.e.,[6]

$$\left(\sqrt{RV}\right)_{t+1-k:t} = \frac{1}{k}\sum_{j=1}^{k}\sqrt{RV_{t-j}}, \tag{17}$$

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

$$\sqrt{RV}_t \;=\; \alpha_0 + \alpha_d\sqrt{RV}_{t-1} + \alpha_w\left(\sqrt{RV}\right)_{t-5:t-1} + \alpha_m\left(\sqrt{RV}\right)_{t-22:t-1} + u_t. \tag{18}$$

In Corsi (2004), u_t is assumed to be Gaussian white noise, giving rise to the same objections as raised against the ARFIMA model class. Employing the volatility–component structure

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

[6] 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, employing the "true" daily, weekly and monthly realized volatilities—as defined by the square root of the sum of the realized variances—yields similar empirical results.

(18), 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 forecasting performance is strong and similar to that of ARFIMA models. A good predictive performance has also been reported in Andersen et al. (2005), who extend the HAR model by including different jump measures.

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 non–trivial. 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,[7] 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).[8]

Below, we jointly estimate parameters using exact maximum–likelihood with the Geweke–Porter–Hudak estimate 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&P500 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 2 and 3 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, which is also 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 2 and 3 convincingly illustrate the inadequacy of the normality assumption in both realized–volatility models. Skewness and kurtosis of the ARFIMA residuals are 17.45 and 734.71, respectively, and those of the HAR model are 15.08 and 609.43.[9]

3.2 Long-memory, Time–dependent Heteroskedasticity 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. But we expect the results for the extended HAR and ARFIMA models to be compatible.

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

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

[9]Note that we find the same form of time–dependent heteroskedasticity for both models when using logarithmic realized volatility. The non–Gaussianity is less pronounced, but still existent.

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

$$\sqrt{RV}_t = \alpha_0 + \alpha_d \sqrt{RV}_{t-1} + \alpha_w \left(\sqrt{RV}\right)_{t-5:t-1} + \alpha_m \left(\sqrt{RV}\right)_{t-22:t-1} + \sqrt{h_t} u_t \quad (19)$$

$$h_t = \omega + \sum_{j=1}^{q} \alpha_j u_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \quad (20)$$

$$u_t | \Omega_{t-1} \sim (0,1). \quad (21)$$

The error term, $\sqrt{h_t} u_t$, follows a conditional density with time–varying variance. To deal with the non–Gaussianity of the error terms we specify a standardized normal inverse Gaussian (NIG) distribution for the (unconditional) iid innovations u_t. Although the incorporation of the GARCH specification can produce fatter unconditional tails, the normality assumption does not allow for the observed skewness. We will adopt the NIG distribution, which is rather flexible and able to reproduce a range of symmetric and asymmetric distributions, such as the normal and the Cauchy. Its density is given by

$$f(x;\alpha,\beta,\mu,\delta) = \frac{\alpha}{\pi} \frac{K_1\left(\alpha\delta\sqrt{1+\left(\frac{x-\mu}{\delta}\right)^2}\right)}{\sqrt{1+\left(\frac{x-\mu}{\delta}\right)^2}} \exp\left\{\delta\left(\sqrt{\alpha^2-\beta^2}+\beta\left(\frac{x-\mu}{\delta}\right)\right)\right\} \quad (22)$$

where $K_i(x)$ is the modified Bessel function of the second kind with index i; $\mu \in \mathbb{R}$ denotes the location parameter, $\delta > 0$ the scale, $\alpha > 0$ the shape, and $\beta \in (-\alpha, \alpha)$ the skewness parameter. Mean and variance are given by

$$\mathrm{E}[x] = \mu + \frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}} \qquad \text{and} \qquad \mathrm{Var}[x] = \frac{\delta\alpha^2}{\sqrt{\alpha^2-\beta^2}^3}. \quad (23)$$

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

$$\mu = -\frac{\beta\left(\alpha^2-\beta^2\right)}{\alpha^2} \qquad \text{and} \qquad \delta = \frac{\left(\alpha^2-\beta^2\right)^{3/2}}{\alpha^2}. \quad (24)$$

By combining the HAR model (18) 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 be easily 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 results show that the GARCH extension substantially improves the goodness of fit, as measured by the AIC and BIC criteria. 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 Gaussianity assumption lies within the (0.85, 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. A more heavy–tailed distribution for the innovations is required.

The goodness–of–fit measures reported in Table 2 show that replacing the Gaussian by the NIG distribution greatly improves the models' fit both for the conventional and the GARCH specification (Models I and II). 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 volatility.

The overall, and overwhelmingly preferred specification turns out to be the HAR–GARCH model with NIG–distributed innovations, suggesting that both extensions to the standard long–memory model for realized volatility 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.

The introduction of the NIG distribution affects the GARCH–parameter estimates. The persistence is much less than under the Gaussian assumption since excess kurtosis can in part be captured by the NIG's shape parameter, α. Note also that, in comparison to the HAR–NIG model without GARCH specification, the shape parameter is much larger, indicating less kurtosis.

Figure 4 demonstrates the adequacy of the HAR–GARCH(1,1)–NIG model. Both skewness and the tail behavior of the innovations are well captured by the NIG. The GARCH–filtered volatility series (top panel in Figure 4) shows the clustering in the volatility of realized 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.

4 Gains in Efficiency

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

Table 3 reports the root mean square error of the parameter estimates for the four models 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 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. Adding the GARCH specification, however, 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 clustering in the volatility of volatility leads to substantial efficiency gains, whereas the relevance 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, in view of the existing trade–off between the kurtosis induced by the GARCH specification and the kurtosis of the NIG distribution, captured by the shape parameter α. However, for the HAR–GARCH specification we can conclude that a sufficiently large sample size is required to accurately estimate the distributional parameters.

5 Prediction

In order to assess the relevance of the volatility 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 construct realized–volatility forecasts up to 22 days ahead. We then recursively re–estimate by expanding the data set by one observation and again predict up to 22 days ahead. The specific forecasting horizons we consider are one day, one week, two weeks, and one month.

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 this regression coefficient, 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.[10]

The results of the forecasting exercise are presented in Table 4. The first four columns report the evaluation criteria for h–day–ahead forecasts of *daily* volatility. Specifically, we forecast daily volatility up to 22 days ahead and compute the R^2 of a projection of the h–day–ahead daily realized volatility on a constant and the h–day–ahead daily volatility forecast. Correspondingly, RMSE, MAE and RMSPE values are based on daily forecast errors. The results show that all four statistics yield more or less the same performance rankings for the four models. Strikingly, the best model for forecasting daily volatility is clearly the HAR model with GARCH specification and Gaussian distributed innovations. The model consistently provides the best forecasts over all forecasting horizons. Considering shorter horizons of up to one week, we find the HAR–GARCH–NIG model to perform second best. However, for longer horizons the standard HAR model performs slightly better. In fact, the performance of the two models is very similar. In contrast, the forecasts obtained from the HAR model with NIG assumption perform worst.

The last four columns of Table 4 report criteria for evaluating the performance of h–day–ahead *average* realized volatility rather than the h–day–ahead *daily* volatility forecasts. Note that the h–day–ahead average volatility actually is of main interest for regulatory risk assessment. The R^2 statistic in the last four columns is based on a regression of h–day–ahead average realized volatility on a constant and the h–day–ahead predictions. All other criteria are based on the h–day–ahead average forecast errors. Since we model and predict daily realized volatility, we have to transform these daily forecasts to obtain the average volatility forecasts. This induces a bias for which we did not correct.

According to all evaluation criteria, the HAR–GARCH model with Gaussian distributed innovations is still the overall preferred model. This is consistent with the results for the daily volatility forecasts. The standard HAR model and the HAR–GARCH model with NIG–distributed innovations have a similar performance, with the latter being slightly better and favoring the inclusion of the GARCH specification.

In summary, the forecasting evaluations show that allowing for time–varying volatility of realized volatility improves the accuracy of volatility (point) forecasts. The improvements are more pronounced for larger forecast horizons. 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.

Time–varying volatility of realized volatility gives rise to a time–varying conditional density of realized volatility. In such a setting, however, the simple evaluation of point forecasts is not sufficient and the uncertainty associated with the forecasts should be assessed. The evaluation of forecast intervals is important in applications, since the uncertainty of the volatility estimates carries over to the uncertainty of return forecasts. Therefore, we also evaluate the accuracy of the density forecasts using a method proposed by Diebold et al. (1998) which is based on Rosenblatt (1952). They show that for a correct density forecast the probability–integral transform, z_t, defined as the cumulative density function of the forecast errors

$$z_t = \int_{\infty}^{y_t} p_t\left(u|x_{t-1}\right) du \qquad (25)$$

[10]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.

should be iid uniformly distributed on the unit interval.

Figure 5 presents the corresponding probability integral transforms of the four models. It turns out that the z_t of the models with Gaussian innovations are far from uniformly distributed, although the incorporation of the volatility of realized volatility leads to some improvement. The inability to capture the 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–GARCH models.

6 Conclusion

We have shown that the commonly used reduced–form realized–volatility models, such as the ARFIMA or HAR, exhibit non–Gaussian distributed innovations and time–varying volatility that might be partly attributed to the time–variation in the volatility of the realized volatility estimator. We, therefore, reject the common Gaussian iid assumption for the residuals in these models and favor a specification with NIG–distributed innovations and, more importantly, one that allows for the clustering in the volatility of realized volatility by also including a GARCH component. 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 realized volatility and to specify an adequate distribution of the error terms, when modeling realized volatility. Both extensions are also considered in Bollerslev et al. (2005) as part of a highly accurate three–equation auxiliary model for returns and realized variations that can also be used for indirect inference.

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 is of importance and a GARCH–type extension should be incorporated. Our forecasting experiments suggest that the specification of a fat–tailed and possibly skewed distribution, such as the NIG, does not seem to improve point forecasts of volatility. However, the analysis of transformed residuals suggests that distributional assumptions should matter for tail–quantile forecasts. In fact, our density forecast results show that, whenever interval or density forecasts are the main focus, a flexible distribution, such as the NIG, should be specified in addition to GARCH.

Table 1: Descriptive Statistics

Series	Mean	Std.Dev.	Median	Skewness	Kurtosis	Ljung–Box(22)
$\sqrt{RV_t}$	0.8627	0.5935	0.7586	15.35	496.76	14,605.0
$\sqrt{\frac{RQ_t}{2MRV_t}}$	0.0821	0.0893	0.0675	19.29	606.68	3,498.8
$\sqrt{\frac{RQQ_t}{2MRV_t}}$	0.0676	0.0475	0.0570	6.62	96.89	12,551.5
$\sqrt{\frac{RTQ_t}{2MRV_t}}$	0.0704	0.0500	0.0594	6.29	79.95	10,955.2

Reported are the descriptive statistics of realized volatility and the measures of the volatility of the realized volatility estimator as defined in equations (13)-(15).

Table 2: Estimation Results

Model	Parameter Estimates									AIC	BIC
ARFIMA(0,d,3)	d	ψ_1	ψ_2	ψ_3							
	0.3483	0.09389	0.1798	0.0452						1.0882	1.0947
	(0.0217)	(0.0258)	(0.0181)	(0.0174)							
HAR Model	Mean Eq.				Distribution		Variance Eq.				
	α_0	α_d	α_w	α_m	α	β	ω	α_1	β_1		
I	0.1066	0.4983	0.2132	0.1659			0.1985			1.1315	1.1380
	(0.0198)	(0.0015)	(0.0059)	(0.0191)			(0.0003)				
II	0.0657	0.2339	0.4541	0.2130			0.0040	0.7464	0.2428	-0.0133	-0.0042
	(0.0068)	(0.0170)	(0.0189)	(0.0177)			(0.0002)	(0.0053)	(0.0066)		
III	0.2180	0.2540	0.2285	0.2645	1.0313	0.6740	0.0933			-0.2592	-0.2501
	(0.0075)	(0.0068)	(0.0077)	(0.0078)	(0.0498)	(0.0479)	(0.0034)				
IV	0.0868	0.2322	0.3965	0.2565	1.6918	1.054	0.0034	0.8143	0.1237	-0.4597	-0.4480
	(0.0073)	(0.0142)	(0.0227)	(0.0184)	(0.1088)	(0.0975)	(0.0003)	(0.0117)	(0.0110)		

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 numbers in parentheses are the standard errors.

Table 3: Efficiency Results

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

All entries report 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.

Table 4: Forecast Evaluation

h	Model	h–day–ahead Daily Vola. R^2	RMSE	MAE	RMSPE	h–day–ahead Average Vola. R^2	RMSE	MAE	RMSPE
1 day	I	0.4848	0.3161	0.1926	0.3298				
	II	0.5109	0.3084	0.1849	0.2925				
	III	0.5034	0.3173	0.1996	0.3422				
	IV	0.5074	0.3091	0.1870	0.3034				
5 days	I	0.4122	0.3431	0.2216	0.3799	0.5839	0.5816	0.3773	0.2476
	II	0.4244	0.3369	0.2043	0.3137	0.6033	0.5726	0.3521	0.2123
	III	0.3999	0.3574	0.2326	0.4076	0.5844	0.6174	0.4102	0.2770
	IV	0.4141	0.3448	0.2184	0.3664	0.5966	0.5807	0.3687	0.2348
10 days	I	0.3274	0.3697	0.2413	0.4234	0.5817	0.8099	0.5474	0.2529
	II	0.3494	0.3608	0.2155	0.3285	0.6032	0.7954	0.4974	0.2074
	III	0.3057	0.3823	0.2519	0.4458	0.5690	0.8726	0.6015	0.2835
	IV	0.3250	0.3730	0.2399	0.4126	0.5884	0.8206	0.5429	0.2450
22 days	I	0.2029	0.4015	0.2711	0.4839	0.5292	1.2385	0.8914	0.2757
	II	0.2591	0.3914	0.2373	0.3548	0.5693	1.2142	0.7953	0.2170
	III	0.1633	0.4128	0.2804	0.4993	0.5036	1.3294	0.9683	0.3002
	IV	0.1937	0.4069	0.2712	0.4737	0.5340	1.2661	0.8958	0.2696

h denotes the forecast horizon and "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 reported R^2 are the regression coefficients of realized volatility on a constant and volatility forecasts.

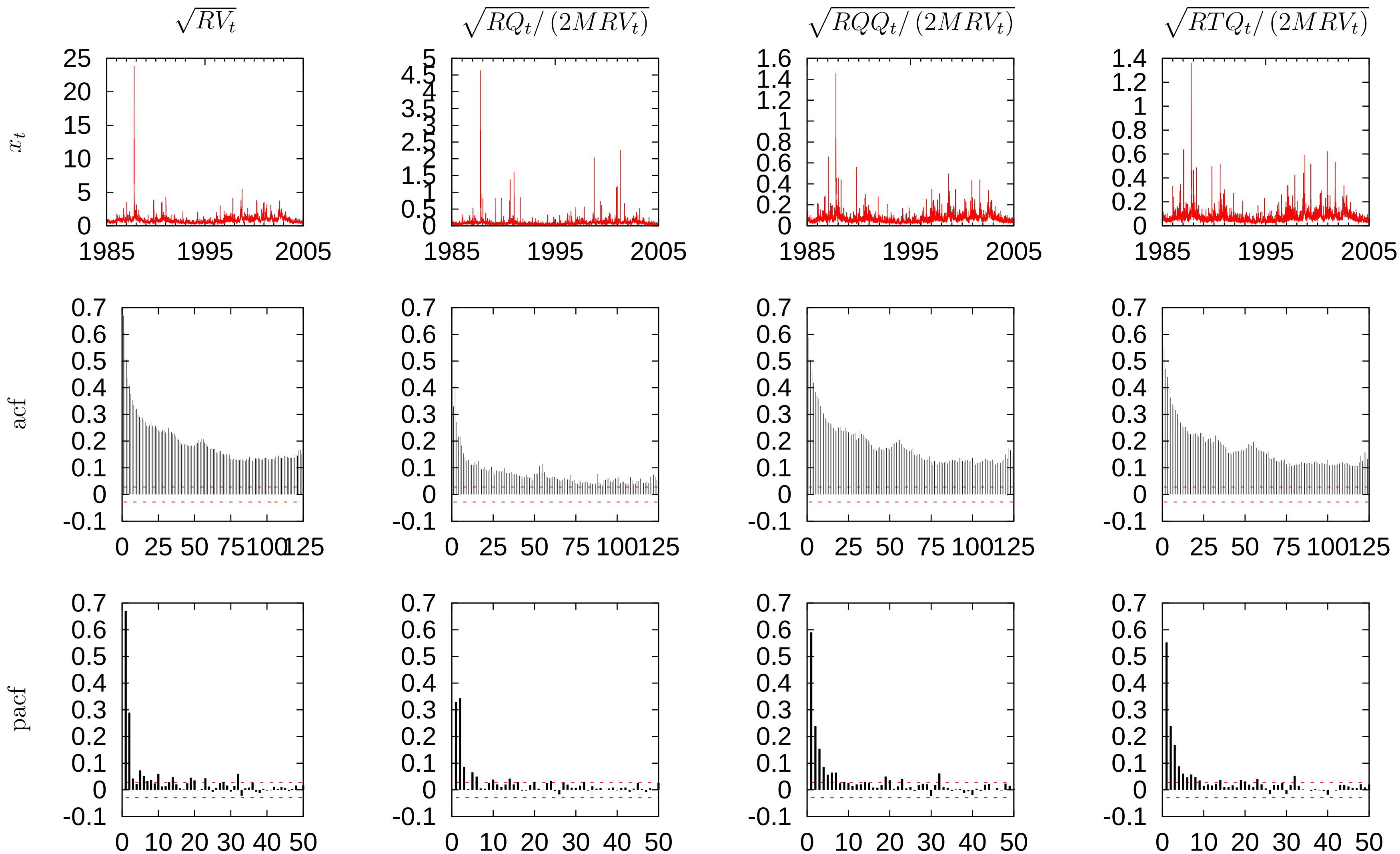

Figure 1: Time series (upper panel), sample autocorrelation functions (acf) (middle panel) and partial autocorrelation functions (pacf) (bottom panel) of realized volatility and the three measures of the volatility of the realized volatility estimator as defined in equations (13)-(15).

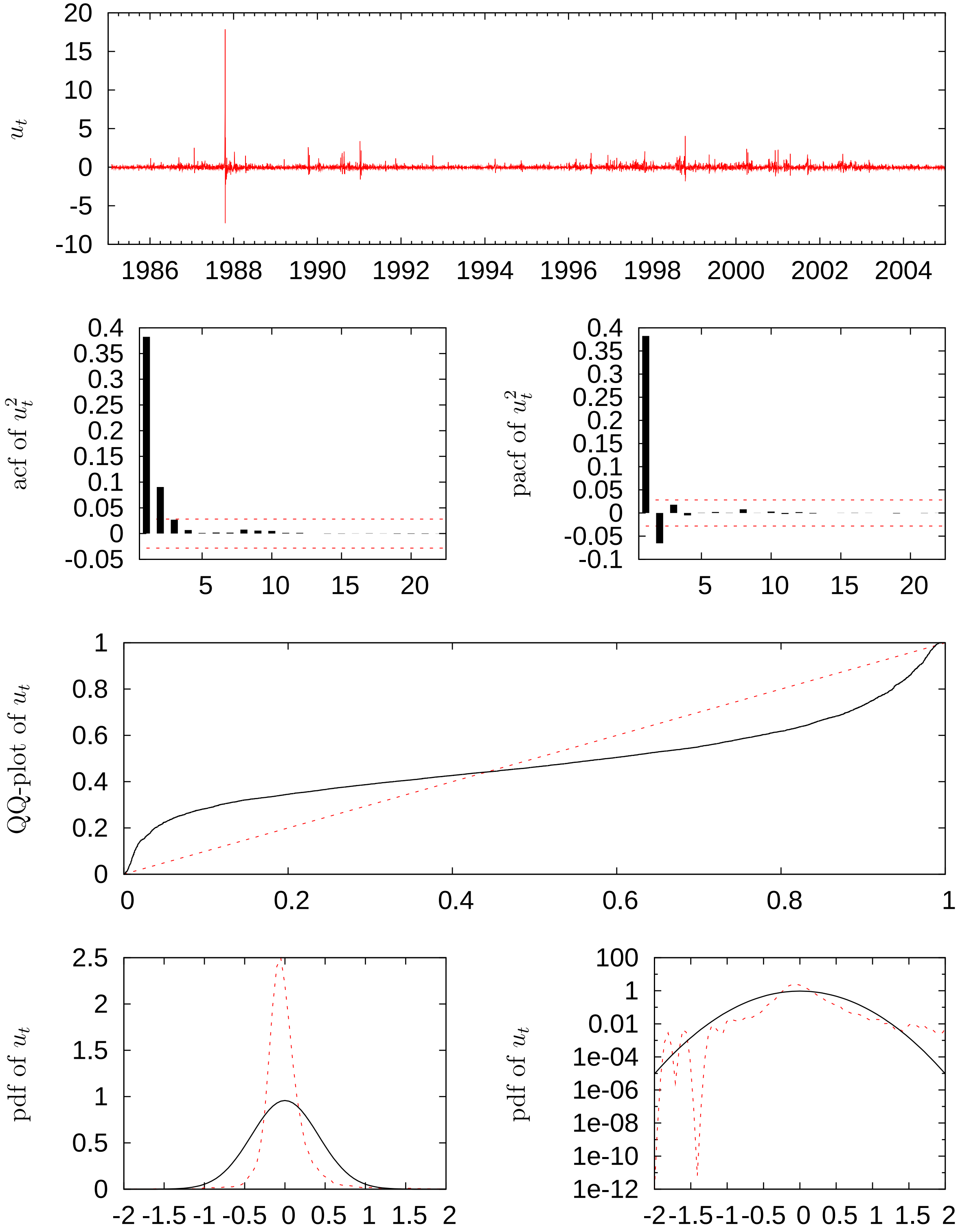

Figure 2: Residual analysis of the ARFIMA(0,d,3) model with Gaussian innovations. Shown are the time series of the residuals (upper panel), the sample autocorrelation functions (acf) and partial autocorrelation functions (pacf) of the squared residuals (second panel), the quantile–quantile plot (third 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).

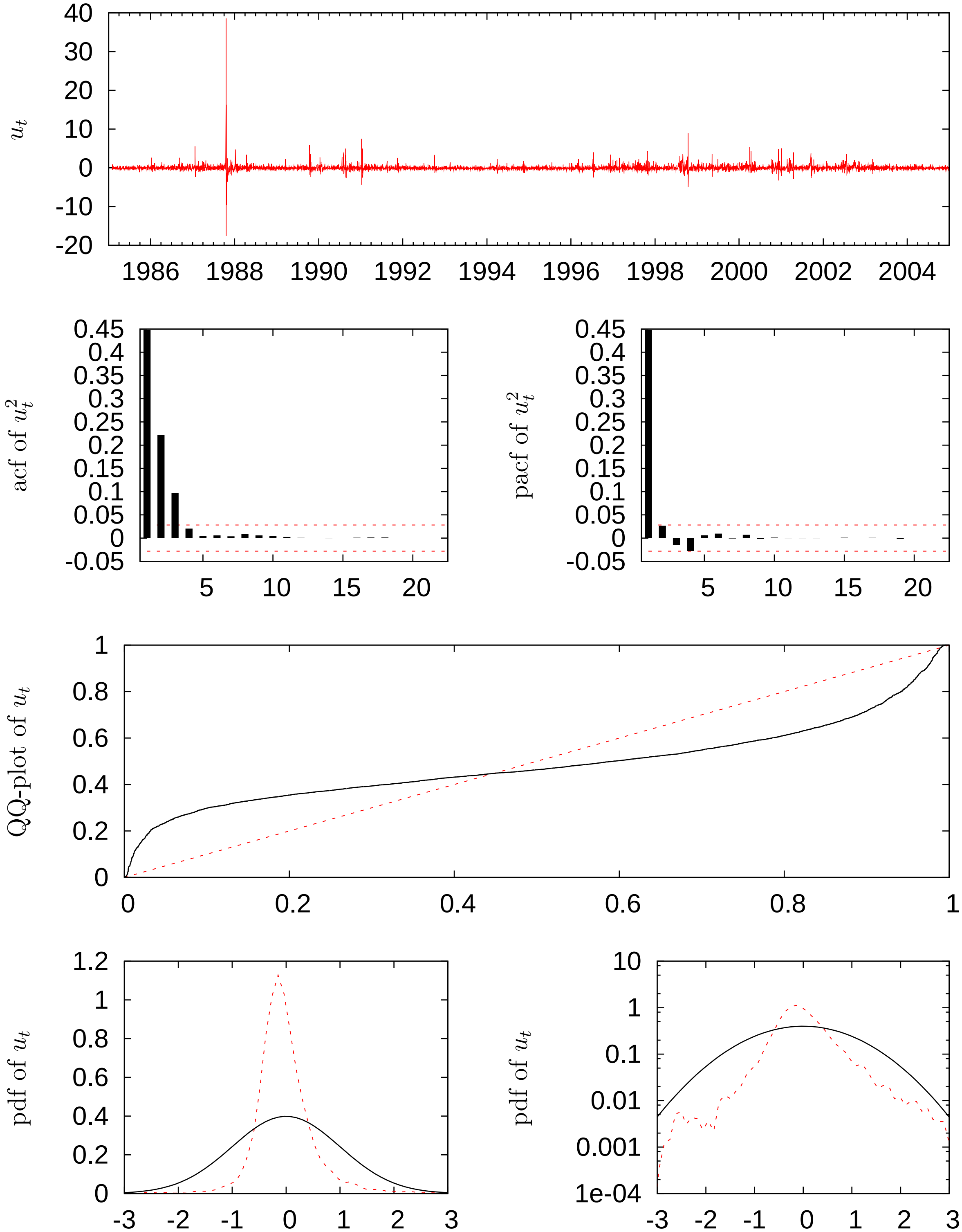

Figure 3: Residual analysis of the pure HAR model with Gaussian innovations. Shown are the time series of the residuals (upper panel), the sample autocorrelation functions (acf) and partial autocorrelation functions (pacf) of the squared residuals (second panel), the quantile–quantile plot (third panel) 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).

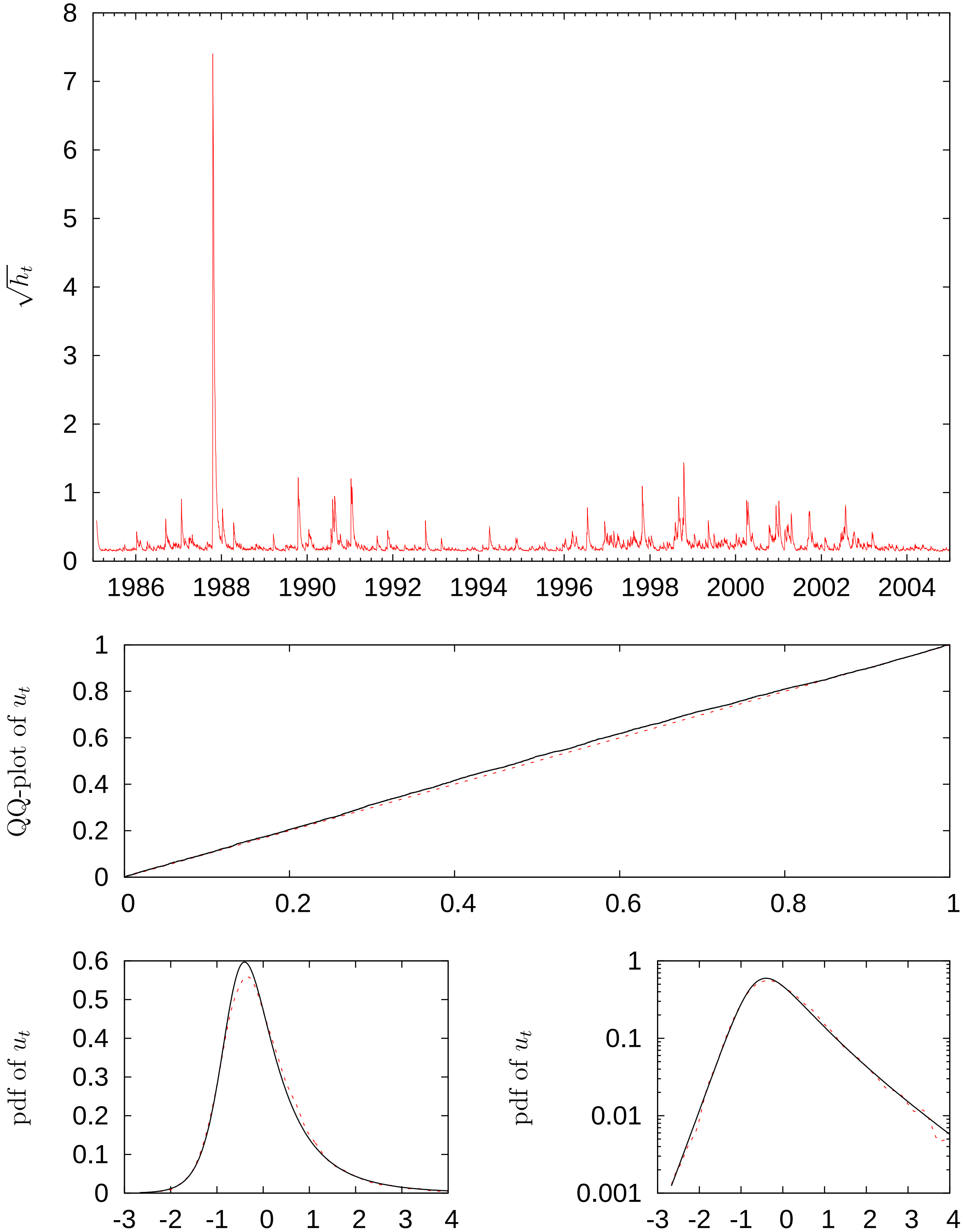

Figure 4: Diagnostics of the HAR–GARCH(1,1)–NIG model. The upper panel presents the GARCH–filtered volatility–of–realized–volatility series; the middle panel shows the quantile–quantile plot of the residuals, the bottom panel presents the kernel density estimates of the residuals (dashed line) and the estimated NIG density (solid line) in level (left) and log scales (right).

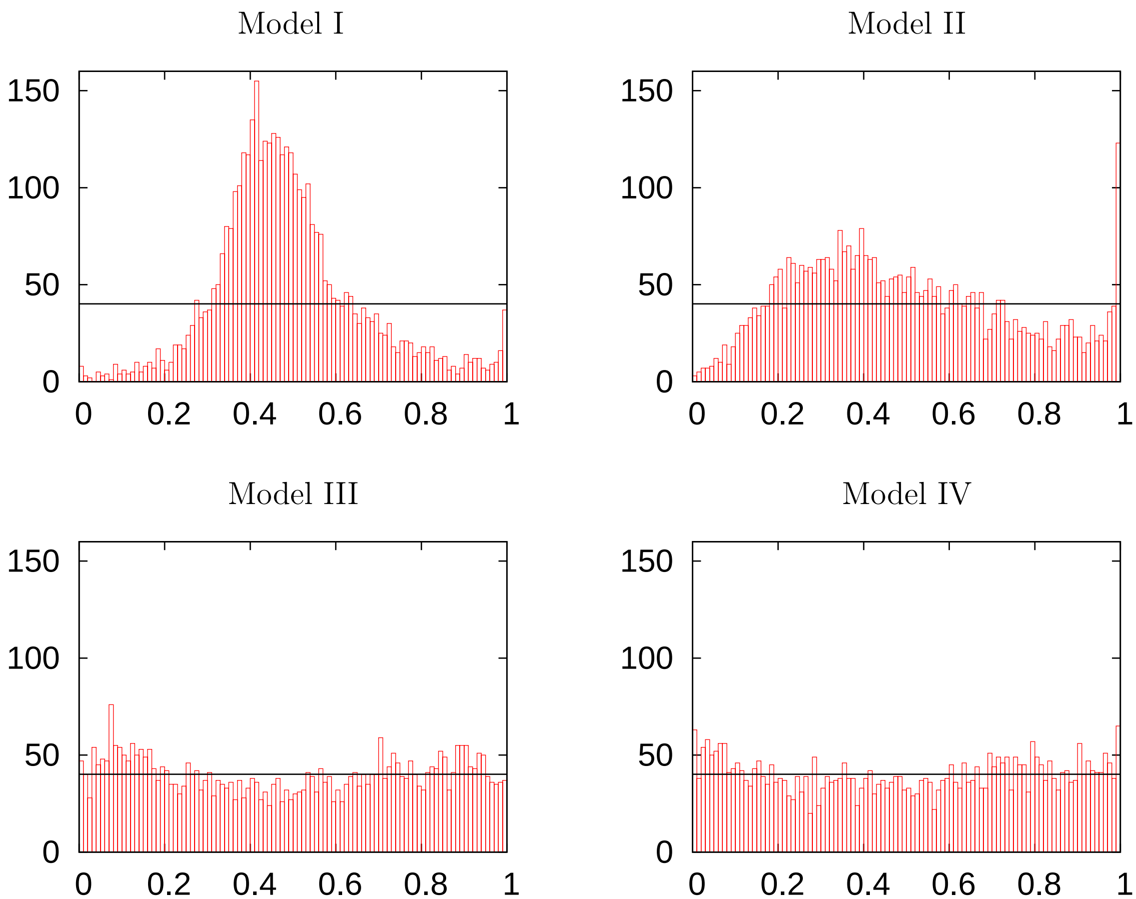

Figure 5: Probability integral transforms of density forecasts based on the four models: Model I refers to the standard HAR model with Gaussian innovations; Model II also includes GARCH effects; Model III is a standard HAR model with standardized NIG innovations; and Model IV corresponds to the HAR–GARCH(1,1) model with standardized NIG innovations.

References

Agiakloglou, C., Newbold, P., and 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., and 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. and 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., and Diebold, F. X. (2005), "Roughing it Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility," Tech. rep., NBER.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and 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., and Labys, P. (2001b), "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55.

— (2003), "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

Andersen, T. G., Bollerslev, T., and Meddahi, N. (2002), "Correcting the Errors : A Note on Volatility Forecast Evaluation Based on High-Frequency Data and Realized Volatilities," Tech. rep., University of Montreal.

Areal, N. M. P. C. and Taylor, S. J. (2002), "The Realized Volatility of FTSE-100 Futures Prices," *Journal of Futures Markets*, 22, 627–648.

Bai, X., Russell, J. R., and 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., and Tieslau, M. A. (1996), "Analysing Inflation by the Fractionally Integrated ARFIMA-GARCH Model," *Journal of Applied Econometrics*, 11, 23–40.

Bandi, F. M. and Russell, J. R. (2005), "Microstructure Noise, Realized Volatility, and Optimal Sampling," Tech. rep., University of Chicago.

Barndorff-Nielsen, O. E. and 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.

— (2002a), "Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, Series B*, 64, 253–280.

— (2002b), "Estimating Quadratic Variation Using Realized Variance," *Journal of Applied Econometrics*, 17, 457–477.

— (2003), "Realised Power Variation and Stochastic Volatility Models," *Bernoulli*, 9, 243–265.

— (2004a), "Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics," *Econometrica*, 72, 885–925.

— (2004b), "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–37.

— (2005a), "Econometrics of Testing for Jumps in Financial Economics using Bipower Variation," *Journal of Financial Econometrics*, forthcoming.

— (2005b), "How Accurate is the Asymptotic Approximation to the Distribution of Realised Volatility?" in *Identification and Inference for Econometric Models. A Festschrift for Tom Rothenberg*, eds. Andrews, D., Powell, J., Ruud, P., and Stock, J., Cambridge: Cambridge University Press, pp. ??–??

Bollerslev, T., Kretschmer, U., Pigorsch, C., and Tauchen, G. (2005), "A Discrete-Time Model for Daily S&P500 Returns and Realized Variations: Jumps and Leverage Effects," Tech. rep., Duke University.

Comte, F. and Renault, E. (1998), "Long-Memory in Continuous-Time Stochastic Volatility Models," *Mathematical Finance*, 8, 291–323.

Corsi, F. (2004), "A Simple Long Memory Model of Realized Volatility," Tech. rep., University of Southern Switzerland.

Corsi, F., Zumbach, G., Müller, U. A., and Dacorogna, M. (2001), "Consistent High-Precision Volatility from High-Frequency Data," *Economic Notes*, 30, 183–204.

Curci, G. and 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.

Diebold, F. X., Gunther, T. A., and Tay, A. S. (1998), "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review*, 39, 863–883.

Doornik, J. A. and 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.

Geweke, J. and Porter-Hudak, S. (1983), "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–238.

Hosking, J. R. M. (1981), "Fractional Differencing," *Biometrica*, 68, 165–176.

Koopman, S. J., Jungbacker, B., and 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.

Martens, M., Dijk, D. v., and de Pooter, M. (2004), "Modeling and Forecasting S&P 500 Volatility: Long Memory, Structural Breaks and Nonlinearity," Tech. rep., Erasmus University Rotterdam.

Martens, M. and Zein, J. (2004), "Predicting Financial Volatility: High-Frequency Time-Series Forecasts vis-à-vis Implied Volatility," *Journal of Futures Markets*, 24, 1005–1028.

Merton, R. C. (1980), "On Estimating the Expected Return on the Market : An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

Mittnik, S., Paolella, M. S., and Rachev, S. T. (1998), "Unconditional and Conditional Distributional Models for the Nikkei Index," *Asia-Pacific Financial Markets*, 5, 99–128.

— (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., and 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.

Pong, S., Shackleton, M. B., Taylor, S. J., and Xu, X. (2004), "Forecasting Currency Volatility: A Comparison of Implied Volatilities and AR(FI)MA Models," *Journal of Banking & Finance*, 28, 2541–2563.

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. and Wang, T. (2003), "Realized Volatility in the Futures Markets," *Journal of Empirical Finance*, 10, 321–353.

Verhoeven, P. and McAleer, M. (2004), "Fat Tails and Asymmetry in Financial Volatility Models," *Mathematics and Computers in Simulation*, 64, 351–361.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005), "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411.

CFS Working Paper Series:

No.	Author(s)	Title
2005/24	**Neil A. Doherty** **Alexander Muermann**	Insuring the Uninsurable: Brokers and Incomplete Insurance Contracts
2005/25	**Franklin Allen** **Elena Carletti**	Credit Risk Transfer and Contagion
2005/26	**Michael Haliassos** **Michael Reiter**	Credit Card Debt Puzzles
2005/27	**Luigi Guiso** **Paola Sapienza** **Luigi Zingales**	Trusting the Stock Market
2005/28	**Tullio Jappelli** **Luigi Pistaferri**	Intertemporal Choice and Consumption Mobility
2005/29	**Luigi Guiso** **Tullio Jappelli**	Awareness and Stock Market Participation
2005/30	**Axel A. Weber** **Guenter W. Beck**	Price Stability, Inflation Convergence and Diversity in EMU: Does One Size Fit All?
2005/31	**Guenter W. Beck** **Axel A. Weber**	Inflation Rate Dispersion and Convergence in Monetary and Economic Unions: Lessons for the ECB
2005/32	**Sumit Agarwal** **Souphala Chonsisengphet** **Chunlin Liu** **Nicholas S. Souleles**	Do Consumers Choose the Right Credit Contracts?
2005/33	**Fulvio Corsi** **Uta Kretschmer** **Stefan Mittnik** **Christian Pigorsch**	The Volatility of Realized Volatility

Copies of working papers can be downloaded at http://www.ifk-cfs.de