University of Toronto
Department of Economics

Do high-frequency measures of volatility improve forecasts of return distributions?

By John M Maheu and Thomas H McCurdy

Do high-frequency measures of volatility improve forecasts of return distributions?

JOHN M. MAHEU and THOMAS H. MCCURDY[*]

June 2008

Abstract

Many finance questions require a full characterization of the distribution of returns. We propose a bivariate model of returns and realized volatility (RV), and explore which features of that time-series model contribute to superior density forecasts over horizons of 1 to 60 days out of sample. This term structure of density forecasts is used to investigate the importance of: the intraday information embodied in the daily RV estimates; the functional form for $\log(RV)$ dynamics; the timing of information availability; and the assumed distributions of both return and $\log(RV)$ innovations. We find that a joint model of returns and volatility that features two components for $\log(RV)$ provides a good fit to S&P 500 and IBM data, and is a significant improvement over an EGARCH model estimated from daily returns.

Keywords:

RV, multiperiod, out-of-sample, term structure of density forecasts, observable SV

1 Introduction

Many finance questions require a full characterization of the distribution of returns. Examples include option pricing which uses the forecast density of the underlying spot asset, or value-at-risk which focuses on a quantile of the forecasted distribution. Once we move away from the simplifying assumptions of Normally-distributed returns or quadratic utility, portfolio choice also requires a full specification of the return distribution.

The purpose of this paper is to study the accuracy of forecasts of return densities produced by alternative models. Specifically, we focus on the value that high frequency measures of volatility,[1] provide in characterizing the forecast density of returns. We propose new bivariate models of returns and realized volatility and explore which features of those time-series models contribute to superior density forecasts over multiperiod horizons out of sample.

Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold, and Labys (2001), Andersen, Bollerslev, Diebold, and Ebens (2001), Andreou and Ghysels (2002), Barndorff-Nielsen and Shephard (2002a), Barndorff-Nielsen and Shephard (2002b), and Meddahi (2002), among others,[2] have established the theoretical and empirical properties of the estimation of quadratic variation for a broad class of stochastic processes in finance. Although theoretical advances continue to be important, part of the research in this new field has focused on the time-series properties and forecast improvements that realized volatility provides. Examples include Andersen, Bollerslev, Diebold, and Labys (2003), Andersen, Bollerslev, and Diebold (2007), Andersen, Bollerslev, and Meddahi (2004), Ghysels and Sinko (2006), Ghysels, Santa-Clara, and Valkanov (2006), Koopman, Jungbacker, and Hol (2005), Maheu and McCurdy (2002, 2007), Martens, van Dijk, and de Pooter (2003), and Taylor and Xu (1997).

Few papers have studied the benefits of incorporating RV into the return distribution. Andersen, Bollerslev, Diebold, and Labys (2003), and Giot and Laurent (2004) consider the value of RV for forecasting and for Value-at-Risk. These approaches decouple the return and volatility dynamics and assume that RV is a sufficient statistic for the conditional variance of returns. Ghysels, Santa-Clara, and Valkanov (2005) find that high frequency measures of volatility identify

---

[1]Unless necessary for clarity, we use the term volatility to refer to either variance or standard deviation.
[2]Recent reviews include Andersen, Bollerslev, and Diebold (2004), Barndorff-Nielsen and Shephard (2007).

a risk-return tradeoff at lower frequencies. Their filtering approach to volatility measurement does not provide a law of motion for volatility and therefore multiperiod forecasts cannot be computed in that setting.

RV is an *ex post* measure of volatility and in general may not be equivalent to the conditional variance of returns. We propose bivariate models based on two alternative ways in which RV is linked to the conditional variance of returns. Since our system provides a law of motion for both return and RV at the daily frequency, multiperiod forecasts of returns and RV or the density of returns are available. The dynamics of the conditional distribution of RV will have a critical impact on the quality of the return density forecasts.

Our benchmark model is an EGARCH model of returns. This model is univariate in the sense that it is driven by one stochastic process which directs the innovations to daily returns. It does not allow higher-order moments of returns to be directed by a second stochastic process. Nor does it utilize any intraday information.

Two types of functional forms for the bivariate models of returns and RV are proposed. The first model uses a heterogeneous autoregressive (HAR) specification (Corsi (2003), Andersen, Bollerslev, and Diebold (2007)) of $\log(RV)$. A second model allows different components of $\log(RV)$ to have different decay rates (Maheu and McCurdy (2007)).

We also consider two ways to link RV to the variance of returns. First, we impose the cross-equation restriction that the conditional variance of daily returns is equal to the conditional expectation of daily RV. Second, motivated by Bollerslev, Kretschmer, Pigorsch, and Tauchen (2007) who model returns, bipower variation and realized jumps in a multivariate setting,[3] we also investigate a specification of our bivariate component model for which the variance of returns is assumed to be synonymous with RV. We label this case 'observable stochastic volatility' and explore whether this assumption improves the term structure of density forecasts. We also compare specifications with non-Normal versus Normal innovations for both returns and $\log(RV)$.

As in our benchmark EGARCH model, all of our bivariate models allow for so-called leverage or asymmetric effects of past negative versus positive return innovations. Our bivariate

---

[3]For definition and development of bipower variation and realized jumps see, for example, Barndorff-Nielsen and Shephard (2004).

models allow for mean reversion in RV. This allows us to evaluate variance targeting for these specifications.

Our main method of model comparison uses the predictive likelihood of returns. This is the forecast density of a model evaluated at the realized return; it provides a measure of the likelihood of the data being consistent with the model. Intuitively, better forecasting models will have higher predictive likelihood values. Therefore our focus is on the relative accuracy of the models in forecasting the return density out of sample. The forecast density of the models is not available in closed form; however, we discuss accurate simulation methods that can be used to evaluate the forecast density and the predictive likelihood.

An important feature of our approach is that we can directly compare traditional volatility specifications, such as EGARCH, with our bivariate models of return and RV since we focus on a common criteria – forecast densities of returns. We generate a predictive likelihood for each out-of-sample data point and for each forecast horizon. For each forecast horizon, we can compute the average predictive likelihood where the average is computed over the fixed number of out-of-sample data points. A term structure of these average predictive likelihoods allows us to investigate the relative contributions of RV over short to long forecast horizons.

Our empirical applications to S&P 500 (Spyder) and IBM returns reveal the importance of intraday return information, the timing of information availability, and non-Normal innovations to both returns and $\log(RV)$. The main features of our results are as follows. Bivariate models that use high frequency intraday data provide a significant improvement in density forecasts relative to an EGARCH model estimated from daily data. Two-component specifications for $\log(RV)$ provide similar or better performance than HAR alternatives; both dominate the less flexible single-component version. A bivariate model of returns with Normal innovations and observable stochastic volatility directed by a 2-component, exponentially decaying function of $\log(RV)$ provides good density forecasts over a range of out-of-sample horizons for both data series. We find that adding a mixture of Normals or GARCH effects to the innovations of the $\log(RV)$ part of this specification is not statistically important for our sample of S&P 500 returns, while the addition of the mixture of Normals provides a significant improvement for IBM.

This paper is organized as follows. The next section introduces the data used to construct daily returns and daily RV. It also discusses the measurement of volatility, the adjustments to realized volatility to remove the effects of market microstructure, and a benchmark model which is based on daily return data. Our bivariate models of returns and RV, based on high-frequency intraday data, are introduced in Section 3. The calculation of density forecasts and the predictive likelihood are discussed in Section 4; results are presented in Section 5. Section 6 concludes.

2   Data and Realized Volatility Estimation

2.1   Data

We investigate a broadly diversified equity index (the S&P 500) and an individual stock (IBM). For the former we use the Standard & Poor's Depository Receipt (Spyder) which is a tradable security that represents ownership in the S&P 500 Index. The ticker symbol is SPY. Since this asset is actively traded, it avoids the stale price effect associated with using the S&P 500 index at high frequencies. Transaction price data associated with both the Spyder and IBM are obtained from the New York Stock Exchange's Trade and Quotes (TAQ) database.

Our data samples cover the period January 2, 1996 to August 29, 2007 for the Spyder and January 4, 1993 to August 29, 2007 for IBM. The shorter sample for the Spyder data was chosen based on volume of trading, for example there were many 5-minute periods with no transactions during the first years after the Spyder started trading in 1993, and a structural break in the Spyder $\log(RV)$ data in the mid 1990s (Liu and Maheu (2008)). The average number of transactions per day for the 1996-2007 sample of Spyder data[4] was $32,971$ but the volume of trades has increased substantially over the sample – especially from 2005 forward. In contrast, the average number of transactions per day for IBM shares has been more stable over our 1993-2007 sample, averaging 6011 transactions per day with a substantial increase from late 2006.

---

[4]Henceforth, unless necessary for clarity, we refer to the Spyder data as S&P 500.

After removing errors from the transaction data,[5] a 5-minute grid[6] from 9:30 to 16:00 EST was constructed by finding the closest transaction price before or equal to each grid-point time. From this grid, 5-minute continuously compounded (log) returns are constructed. These returns are scaled by 100 and denoted as $r_{t,i}, i = 1, ..., I$, where I is the number of intraday returns in day t. For our 5-minute grid, normally $I = 78$ although the market closed early on a few days. This procedure generated $228,394$ 5-minute returns corresponding to 2936 trading days for the S&P 500; and $286,988$ 5-minute returns corresponding to 3693 trading days for IBM.

2.2 Realized Volatility Estimation

A traditional proxy for *ex post* latent volatility has been *daily* squared returns. As shown by Andersen and Bollerslev (1998), this measure of volatility is very noisy and of limited use in assessing features of volatility such as its time-series properties.

Better estimates of *ex post* latent volatility are available. The increment of quadratic variation is a natural measure of *ex post* variance over a time interval. A popular estimator of this type is realized variance or realized volatility (RV) computed as the sum of squared returns over this time interval. As shown by Andersen, Bollerslev, Diebold, and Labys (2001), as the sampling frequency is increased, the sum-of-squared returns converges to the quadratic variation over a fixed time interval for a broad class of models. Thus RV is a consistent estimate of *ex post* variance for that period. The asymptotic distribution of RV has been studied by Barndorff-Nielsen and Shephard (2002a) who provide conditions under which RV is also an unbiased estimate.

Given the intraday returns, $r_{t,i}, i = 1, ..., I$, an unadjusted RV estimator of daily *ex post* variance is

$$RV_{t,u} = \sum_{i=1}^{I} r_{t,i}^2. \tag{2.1}$$

However, in the presence of market-microstructure dynamics, RV can be a biased and inconsistent estimator for quadratic variation (Bandi and Russell (2005) and Zhang, Mykland, and Aït-Sahalia

---

[5]Data were collected with a TAQ correction indicator of 0 (regular trade) and when possible a 1 (trade later corrected). We also excluded any transaction with a sale condition of Z, which is a transaction reported on the tape out of time sequence, and with intervening trades between the trade time and the reported time on the tape. We also checked any price change that was larger than 3% and removed obvious errors.

[6]Volatility signature plots using grids ranging from 1 minute to 195 minutes are available on request.

(2005)). Therefore, we consider several adjustments to our estimates based on Hansen and Lunde (2006) and gauge their statistical performance in our model comparisons.[7]

Hansen and Lunde (2006) suggest the use of Bartlett weights to rule out negative values for RV. Following this approach, a corrected RV estimator is

$$RV_{t,ACqb} = \omega_0\hat{\gamma}_0 + 2\sum_{j=1}^{q}\omega_j\hat{\gamma}_j, \quad \hat{\gamma}_j = \sum_{i=1}^{I-j}r_{t,i}r_{t,i+j}, \tag{2.2}$$

in which the weights follow a Bartlett scheme $\omega_j = 1 - \frac{j}{q+1}$, $j = 0, 1, ..., q$. We consider $q = 1, 2, 3$.

In summary, we construct daily RV with no adjustment, that is RV_u as in equation (2.1), and, alternatively, RV_{ACqb} using a kernel-based adjustment for market-microstructure dynamics, as in equation (2.2). Daily returns, r_t, are computed as the logarithmic difference of the price for the last grid point of the day and the price for the first grid point following market opening. These returns are scaled by 100.

Table 1 displays summary statistics for returns and alternative RV estimates computed from a 5-minute grid. If we take the sample variance of daily returns as a benchmark estimate of volatility in which no market microstructure effects are present, and compare this to the sample mean of RV, we see a clear bias for unadjusted RV. With respect to removing bias, it appears that a Bartlett adjustment with $q = 3$ is necessary for the S&P 500 (Spyder) data, whereas an adjustment with $q = 1$ is adequate for the IBM data. This conclusion is supported by autocorrelation analyses of the 5-minute returns data, as revealed by the autocorrelation functions with associated confidence bounds in Figure 1 for the S&P 500 and IBM respectively. For the remainder of our paper, unless otherwise stated, we use $RV_t \equiv RV_{t,ACqb}$, with $q = 3$ for the S&P 500 and $q = 1$ for the IBM data. The noise-to-signal ratio, computed following Hansen and Lunde (2006), is 0.15% for our sample of Spyder data and 0.05% for the IBM sample. Figures 2 and 3 display daily returns plus the square-root of adjusted RV for the S&P 500 and IBM sample periods respectively.

---

[7]For alternative approaches to dealing with market microstructure dynamics see Aït-Sahalia, Mykland, and Zhang (2005, 2006), Bandi and Russell (2006), Barndorff-Nielsen, Hansen, Lunde, and Shephard (2004, 2006), Curci and Corsi (2006), Oomen (2005), Zhang (2004) and Zhou (1996).

2.3 Benchmark Specification

One way to ascertain whether or not high-frequency (intraperiod) information contributes to improved forecasts of return distributions, is to compare density forecasts from our bivariate specifications of returns and $\log(RV)$ with those from a traditional EGARCH model[8] (Nelson (1991)) specified as:

$$r_t = \mu + \epsilon_t, \ \ \epsilon_t = \sigma_t u_t \ \ u_t \sim NID(0,1), \tag{2.3}$$

$$\log(\sigma_t^2) = \omega + \beta \log(\sigma_{t-1}^2) + \gamma u_{t-1} + \alpha|u_{t-1}|. \tag{2.4}$$

Note that this EGARCH model conditions on daily (standardized) return innovations rather than RV which incorporates intraday information.

3 Joint Return-RV Models

As discussed in the Introduction, an integrated model of returns and realized volatility is needed to deal with common questions in finance which require a forecast density of returns for multiple horizons.[9] In this section, we introduce two alternative joint specifications of daily returns and realized volatility. These bivariate models are distinguished by alternative assumptions about RV dynamics. We also consider versions of these bivariate models with non-Normal return and $\log(RV)$ innovations, as well as a version with an alternative assumption concerning available information about RV. In each case, cross-equation restrictions link the variance of returns and our realized volatility specification.

3.1 Linking Variance of Returns to RV

Corollary 1 of Andersen, Bollerslev, Diebold, and Labys (2003) shows that, under empirically realistic conditions, the conditional expectation of quadratic variation (QV_t) is equal to the con-

---

[8]Note that we found the EGARCH to have the best performance among a variety of other GARCH specifications that were explored.

[9]Similar to most of the existing literature, although see Meddahi (2003), we ignore aggregation issues linking an underlying diffusion to dynamics of daily returns and RV. We focus on the empirical value of using daily RV in daily return models.

ditional variance of returns, that is, $E_{t-1}(QV_t) = \text{Var}_{t-1}(r_t) \equiv \sigma_t^2$. If RV is an unbiased estimator of quadratic variation,[10] it follows that the conditional variance of returns can be linked to RV as $\sigma_t^2 = E_{t-1}RV_t$ where the information set is defined as $\Phi_{t-1} \equiv \{r_{t-1}, RV_{t-1}, r_{t-2}, RV_{t-2}, ..., r_1, RV_1\}$. Assuming that RV has a log-Normal distribution, that restriction takes the form

$$\sigma_t^2 = E_{t-1}RV_t = \exp\left(E_{t-1}\log(RV_t) + \frac{1}{2}\text{Var}_{t-1}(\log(RV_t))\right). \tag{3.1}$$

3.2 HAR-RV Specifications

We begin with a bivariate specification for daily returns and RV in which conditional returns are driven by Normal innovations and the dynamics of $\log(RV_t)$ are captured by Heterogeneous AutoRegressive (HAR) functions of lagged $\log(RV_t)$. Corsi (2003) and Andersen, Bollerslev, and Diebold (2007) use HAR functions in order to parsimoniously capture long-memory dependence. Motivated by that work, we define

$$\log(RV_{t-h,h}) \equiv \frac{1}{h}\sum_{i=0}^{h-1}\log(RV_{t-h+i}), \quad \log(RV_{t-1,1}) \equiv \log(RV_{t-1}). \tag{3.2}$$

For example, $\log(RV_{t-22,22})$ averages $\log(RV)$ over the most recent 22 days, that is, from $t-22$ to $t-1$, $\log(RV_{t-5,5})$ over the most recent 5 days, etc.

This leads to our bivariate specification for daily returns and RV with the dynamics of $\log(RV_t)$ modeled as an asymmetric HAR function of past $\log(RV)$. This bivariate system is summarized as follows:

$$r_t = \mu + \epsilon_t, \ \epsilon_t = \sigma_t u_t, \ u_t \sim NID(0,1) \tag{3.3}$$

$$\log(RV_t) = \omega + \phi_1\log(RV_{t-1}) + \phi_2\log(RV_{t-5,5}) + \phi_3\log(RV_{t-22,22})$$

$$+ \ \gamma u_{t-1} + \eta v_t, \ v_t \sim NID(0,1). \tag{3.4}$$

This bivariate specification of daily returns and RV imposes the cross-equation restriction that

---

[10]We assume that any stochastic component in the intraperiod conditional mean is negligible compared to the total conditional variance. It is also straightforward to estimate a bias term.

relates the conditional variance of daily returns to the conditional expectation of daily RV, as in equation (3.1). If the model is stationary, the unconditional mean of $\log(RV)$ is $\omega/(1-\phi_1-\phi_2-\phi_3)$. This can be used to enforce variance targeting. Joint estimation of the bivariate system in equations (3.3), (3.4) and (3.1) is by maximum likelihood.

Since our applications are to equity returns, it is important to allow for asymmetric effects in volatility. To facilitate comparisons with the benchmark EGARCH model, our parameterization in equation (3.4) includes an asymmetry term, γu_{t-1} associated with the standardized return innovation, u_{t-1}. The impact coefficient for a negative innovation to returns will be $-\gamma$, whereas the impact of a positive innovation will be γ. Typically, $\hat{\gamma} < 0$, which means that a negative innovation to returns implies a higher conditional variance for next period. Unlike EGARCH, our parameterization does not propagate the asymmetry further into future volatility.

In-sample fit of GARCH models have generally favored return innovations with tails that are fatter than those implied by a Normal distribution. Therefore, we evaluate whether or not that result obtains for our bivariate models of returns and RV. That is, we also try replacing equation (3.3) with

$$r_t = \mu + \epsilon_t, \ \epsilon_t = \sigma_t u_t, \ \ u_t \sim t_\nu(0, 1), \tag{3.5}$$

in which t_ν denotes a t-distribution with mean 0, variance 1, and ν degrees of freedom. The remainder of the bivariate dynamic system for this case is the same as above. We compare this bivariate system with t-distributed return innovations to that with Normally-distributed innovations, not only for in-sample fit, but also for the term structure of out-of-sample density forecasts.

3.3 Component-RV Specifications

This bivariate specification for daily returns and RV has conditional returns driven by Normal innovations but now the dynamics of $\log(RV_t)$ are captured by two components (2Comp) with different decay rates, as in Maheu and McCurdy (2007). In particular, this bivariate system can

be summarized as follows:

$$r_t = \mu + \epsilon_t, \; \epsilon_t = \sigma_t u_t, \;\; u_t \sim NID(0,1) \tag{3.6}$$

$$\log(RV_t) = \omega + \sum_{i=1}^{2} \phi_i s_{i,t} + \gamma u_{t-1} + \eta v_t, \;\; v_t \sim NID(0,1) \tag{3.7}$$

$$s_{i,t} = (1 - \alpha_i)\log(RV_{t-1}) + \alpha_i s_{i,t-1}, \;\; 0 < \alpha_i < 1, \; i = 1,2. \tag{3.8}$$

Again, we impose the cross-equation restriction that relates the conditional variance of daily returns to the conditional expectation of daily RV as in equation (3.1). For this specification of our bivariate model, the dynamics of daily $\log(RV)$ are parameterized as the component model specified in equations (3.7) and (3.8) which replace the HAR function in equation (3.4).

Note that this dynamic model for $\log(RV)$ includes the sum of two components that are allowed to decay at different rates associated with their parameters α_i. That is, for component i, a small value of α_i puts more weight on $\log(RV_{t-1})$, and less weight on the past $s_{i,t-1}$. Conversely, an α_i close to 1 puts less weight on recent observations and more weight on past $s_{i,t-1}$ which smooths the data. Startup values of $s_{i,0}$ are set to the average of $\log(RV)$ based on the reserved presample of data.

Although infinite exponential smoothing provides parsimonious estimates, it possesses several drawbacks. For instance, it does not allow for mean reversion in volatility; and, as Nelson (1990) has shown in the case of squared returns or squared innovations to returns, the model is degenerate in its asymptotic limit. To circumvent these problems, but still retain parsimony, our dynamic model for $\log(RV_t)$, given by equation (3.7), weights each component i by the parameter $0 < \phi_i < 1$ and adds an intercept, ω. Note that when the model is stationary, variance forecasts will mean revert to $\omega/(1 - \phi_1 - \phi_2)$.[11] This result can be used to do variance targeting and eliminate the parameter ω from the model.[12] This model implies an infinite expansion in $\log(RV_{t-j})$ with coefficients of $\phi_1(1 - \alpha_1)\alpha_1^{j-1} + \phi_2(1 - \alpha_2)\alpha_2^{j-1}$, $j = 1,2,....$[13]

In order to evaluate the potential importance of t-distributed return innovations for this bivariate specification, we replace equation (3.6) with equation (3.5), and jointly estimate with

---

[11] Assuming $\phi_1 + \phi_2 < 1$.

[12] That is, set $\omega = \text{mean}(\log(RV))(1 - \phi_1 - \phi_2)$.

[13] Expanding (3.8) gives $s_{i,t} = (1 - \alpha_i)\sum_{n=0}^{\infty} \alpha_i^n \log(RV_{t-1-n})$.

equations (3.7), (3.8) and (3.1).

3.4 Component-RV Specifications with Observable SV

Motivated by Bollerslev, Kretschmer, Pigorsch, and Tauchen (2007), we also present results for an alternative assumption about available information in which we replace equation (3.1) with $\sigma_t^2 \equiv RV_t$. This is like taking a stochastic volatility model in which volatility is latent and assuming it is equal to observable RV, hence we label this *observable SV* (OSV).

This bivariate system, with the stronger information assumption, is specified as follows:

$$r_t = \mu + \epsilon_t, \; \epsilon_t = \sqrt{RV_t}u_t, \;\; u_t \sim NID(0,1) \tag{3.9}$$

$$\log(RV_t) = \omega + \sum_{i=1}^{2}\phi_i s_{i,t} + \gamma u_{t-1} + \eta v_t, \;\; v_t \sim NID(0,1) \tag{3.10}$$

$$s_{i,t} = (1-\alpha_i)\log(RV_{t-1}) + \alpha_i s_{i,t-1}, \;\; 0 < \alpha_i < 1, \; i = 1,2. \tag{3.11}$$

which we label 2Comp-OSV.

3.5 Extensions

We consider two extensions to the previous model. The first sets $\eta = 1$, and replaces the innovation v_t in (3.10) with a mixture of two Normals. It has density

$$v_t \sim \begin{cases} N(0, \sigma_{v,1}^2) & \text{with probability } \pi \\ N(0, \sigma_{v,2}^2) & \text{with probability } 1 - \pi \end{cases} \tag{3.12}$$

and allows $\log(RV_t)$ to have a fat-tailed distribution.

The second extension is to include GARCH dynamics. In this case, η in (3.10) has a time subscript and follows the GARCH(1,1) model

$$\eta_t^2 = \kappa_0 + \kappa_1[\log(RV_{t-1}) - E_{t-2}\log(RV_{t-1})]^2 + \kappa_2\eta_{t-1}^2. \tag{3.13}$$

where $\log(RV_{t-1}) - E_{t-2}\log(RV_{t-1})$ denotes the innovation to $\log(RV)$ at time $(t-1)$.

4 Density Forecasts

Our focus is on the return distribution and we evaluate each of the models ability to provide accurate density forecasts of daily returns. A popular approach to assess the accuracy of a model's density forecasts is the predictive likelihood or logarithmic score (Amisano and Giacomini (2007), Bao, Lee, and Saltoglu (2006), and Weigend and Shi (2000)). This approach evaluates the model's density forecast at the realized return. This is generally done for a one-step-ahead forecast density as multiperiod density forecasts are often not available in closed form. In this paper we advocate multiperiod forecasts since they provide more information to discern among models. The details of the multiperiod predictive likelihood and how to calculate it are described below.

The average predictive likelihood over the out-of-sample observations $t = \tau + k_{max}, ..., T - k$, is

$$D_{M,k} = \frac{1}{T - \tau - k_{max} + 1} \sum_{t=\tau+k_{max}-k}^{T-k} \log f_{M,k}(r_{t+k}|\Phi_t, \theta), \quad k \geq 1, \tag{4.1}$$

where $f_{M,k}(x|\Phi_t, \theta)$ is the k-period ahead predictive density for model M, given Φ_t and parameter θ, evaluated at the realized return $x = r_{t+k}$. Intuitively, models that better account for the data produce larger $D_{M,k}$.

As we will see below for our application to S&P 500, $T = 2936, \tau = 1200, k_{max} = 60$ so that $\tau + k_{max} - 1 = 1259$. $D_{M,k}$ is computed for each k using the out-of-sample returns $r_{1260}, ..., r_{2936}$. That is, if $k = 1$, $D_{M,1}$ is computed using out-of-sample returns $r_{1260}, ..., r_{2936}$. For $k = 2$, $D_{M,2}$ is computed using the same out-of-sample returns, etc. This gives us a *term structure of average predictive likelihoods*, $D_{M,1}, ..., D_{M,60}$, to compare the performance of alternative models, M, over an *identical set of out-of-sample data points*.

To assess the statistical differences in $D_{M,k}$ for two models we present Diebold and Mariano (1995) test statistics based on the work of Amisano and Giacomini (2007). Under the null hypothesis of equal performance based on predictive likelihoods of horizen k for models A and B, $t_{A,B}^k = (D_{A,k} - D_{B,k})/(\hat{\sigma}_{AB,k}/\sqrt{T - \tau - k_{max} + 1})$ is asymptotically standard Normal.

$\hat{\sigma}_{AB,k}$ is the Newey-West long-run sample variance (HAC) estimate for $d_t = \log f_{A,k}(r_{t+k}|\Phi_t, \hat{\theta}) - \log f_{B,k}(r_{t+k}|\Phi_t, \hat{\theta})$. $\hat{\theta}$ denotes the maximum likelihood estimate for the respective model. Due to the overlapping nature of the density forecasts for $k > 1$ we set the lag-length in the Newey-West variance estimate to the integer part of $[k \times 0.15]$.[14] A large positive (negative) test statistic is a rejection of equal forecast performance and provides evidence in favor of model A (B). As with the predictive likelihoods, a term structure of associated test statistics $t^k_{A,B}$, $k = 1, ..., k_{max}$ are presented in the Results section.

4.1   Computations

For all $k > 1$ the term $f_{M,k}(r_{t+k}|\Phi_t, \theta)$ will be unknown for the models we consider. However, given that we have fully specified the law of motion for daily returns and RV, we can accurately estimate this quantity by standard Monte Carlo methods. A conventional approach to estimate the forecast density would be to simulate the model out k periods a large number of times and apply a kernel density estimator to these realizations. However, using the kernel density estimator to estimate the forecast density ignores the fact that, in our applications, conditional on the variance we know the distribution. The use of conditional analytic results has been referred to as Rao-Blackwellization and is a standard approach to reduce the variance of a Monte Carlo estimate (Robert and Casella (1999)). This is particularly useful in density estimation which is our context.

To illustrate consider our basic benchmark EGARCH model in (2.3). Note that in this univariate case the information set, Φ_t, just includes past returns. Our estimate is

$$f_{M,k}(r_{t+k}|\Phi_t, \theta) = \int f(r_{t+k}|\mu, \sigma^2_{t+k}) p(\sigma^2_{t+k}|\Phi_t) d\sigma^2_{t+k} \qquad (4.2)$$

$$\approx \frac{1}{N} \sum_{i=1}^{N} f(r_{t+k}|\mu, \sigma^{2(i)}_{t+k}), \quad \sigma^{2(i)}_{t+k} \sim p(\sigma^2_{t+k}|\Phi_t) \qquad (4.3)$$

where $f(r_{t+k}|\mu, \sigma^{2(i)}_{t+k})$ is a Normal density with mean μ and variance σ^2_{t+k}, evaluated at return r_{t+k}; and $\sigma^{2(i)}_{t+k}$ is simulated out N times according to the EGARCH specification, $p(\sigma^2_{t+k}|\Phi_t)$,

---

[14]Our results are generally stronger (stronger rejections of the null hypothesis) for smaller lag-lengths.

which is conditional on time t quantities σ_t^2, u_t, and $\hat{\theta}$, the maximum likelihood estimate of the parameter vector based on Φ_t.

For the joint models of returns and RV, we do a similar exercise to compute the predictive likelihood for returns. In this case, we simulate out both the return and RV dynamics, which implicitly integrates out the unknown σ_{t+k}^2. For each simulation of $RV_{t+1}^{(i)}, ..., RV_{t+k-1}^{(i)}$, $i = 1, ..., N$, we can compute $\sigma_{t+k}^{2(i)} = E_{t+k-1}RV_{t+k}^{(i)}$ using (4.1).[15] A numerical standard error can be used to access accuracy of $\hat{f}_{M,k}(r_{t+k}|\Phi_t, \theta)$ and $\hat{D}_{M,k}$. [16] In our application we found $N = 10000$ to provide sufficient accuracy. For example, the numerical standard error is typically well below 1% of $\hat{D}_{M,k}$. Note that for all of our bivariate models the dynamics of the conditional distribution of RV will have a critical impact on the quality of the return density forecasts.

5 Results

Our first results are out-of-sample density forecasts evaluated using predictive likelihoods. The S&P 500 sample starts at 1996/01/02, the first out-of-sample density forecast begins at 2000/12/26 ($t = 1260$) and ends at 2007/8/29 ($t = 2936$), for a total of 1677 density forecasts for each k. We summarize these out-of-sample forecasts by averaging the associated 1677 predictive likelihoods for each k and then plotting their term structure for the forecast horizons $k = 1, ..., 60$, that is, from 1 to 60 days out of sample. Note that the IBM sample starts at 1993/01/04, the first density forecast begins at 1997/12/24 ($t = 1260$), and ends at 2007/8/29 ($t = 3693$), for a total of 2434 density forecasts for each k. Full sample parameter estimates for the best models are discussed at the end of the section. Model estimation conditions on the first 24 observations.

Our empirical work considered many different models, including different innovation distributions for returns, the value of variance targeting for $\log(RV)$, different functional forms for $\log(RV)$, and a variety of GARCH specifications estimated using daily returns. We note the

---

[15]Recall that the observable SV specification in Section 3.4 sets $\sigma_{t+k}^{2(i)} = RV_{t+k}^{(i)}$.

[16]To calculate a numerical standard error for $\hat{D}_{M,k}$: let v^2 denote the sample variance of the draws of $f(r_{t+k}|\mu, \sigma_{t+k}^{2(i)})$, then the numerical standard error for $\hat{f}_{M,k}(r_{t+k}|\Phi_t, \theta)$ is $\nu/\sqrt{N}$. Using the delta rule to calculate $\hat{Var}(\log \hat{f}_{M,k}(r_{t+k}|\Phi_t, \theta))$; the numerical standard error of $\hat{D}_{M,k}$ is $\sqrt{\sum_{t=\tau+k_{max}-k}^{T-k} \hat{Var}(\log \hat{f}_{M,k}(r_{t+k}|\Phi_t, \theta))}/(T - \tau - k_{max} + 1)$.

following general results: models with variance targeting were dominated by the unrestricted version of the model; HAR and component models that link the conditional variance of returns to RV_t by (3.1) always performed better with t-innovations to returns;[17] 2-component models were always better than single-component versions; EGARCH was the best model based on daily returns data. In the following summary of results, we focus on the top models in different categories.

Our empirical applications to S&P 500 and IBM returns reveal the importance of intraday information, the timing of information availability, and non-Normal innovations to both returns and $\log(RV)$. Figures 4 and 6 compare the term structures of density forecasts for the best models of each type for the S&P 500 and IBM respectively; Figures 5 and 7 evaluate the robustness of those results to further generalizations. The second plot in each figure displays a corresponding term structure of Diebold-Mariano test statistics for equal forecast performance for selected models.

Note that all of the average predictive likelihood term structures display a negative slope. This is because the conditioning information is most useful for small k. As we forecast further and further out of sample, the value of the current information diminishes. All of our models are stationary so that multiperiod forecast densities converge to the unconditional distribution. Using the same data points to evaluate the predictive likelihood for different k, we can see how accuracy of forecasts deteriorate for longer horizons.

Two main conclusions can be gleaned from Figure 4. Firstly, high-frequency intraday data provide a significant improvement in density forecasts relative to an EGARCH model estimated from daily data. The same conclusion about the value of high-frequency data can be drawn from the IBM sample, as shown in Figure 6. Secondly, both the 2-component and the HAR specification dominate a single-component version of equations (3.7) and (3.8) for the dynamics of $\log(RV)$. Note that the advantage of the more flexible functional forms (either 2-component or HAR) increase the further out we forecast.

The three best bivariate specifications are the 2Comp-OSV, 2Comp and HAR. For the S&P

---

[17]We did consider t-innovations for returns in the observable SV models, but estimation supported a Normal distribution since the degree of freedom parameter always moved to extremely large values.

500, the latter two do equally well; for IBM forecasts the 2Comp specifications are better than HAR. The additional information assumed by the observable stochastic volatility (OSV) assumption, although very important with respect to in-sample fit as shown below, is only significant with respect to density forecasts for long horizons (beyond 45 days) for the S&P 500. The OSV assumption does not improve density forecasts for the IBM case, as shown by the Diebold-Mariano test statistics in Figure 6 for '2Comp-OSV vs 2Comp'.

Figures 5 and 7 evaluate the robustness of the best bivariate specification, for the S&P 500 and IBM respectively, to generalizations of the distributional assumption for $\log(RV)$. In particular, as discussed in Section 3.5, we generalize equation (3.10) to allow either a mixture-of-Normals or a GARCH parameterization of the variance of $\log(RV)$. It is clear from Figure 5 that neither of these generalizations significantly improve the out-of-sample density forecasts for our S&P 500 sample. On the other hand, Figure 7 suggests that a mixture-of-Normals parameterization of the variance of $\log(RV)$ improves density forecasts relative to the Normally-distributed alternative for the IBM sample.

Table 2 provides full-sample model estimates for two of the best bivariate specifications for S&P 500 data. Estimates for the 2Comp-OSV model are reported in column 2 of the table. This specification imposes the restriction $\phi_1 = \phi_2$ which produced the best forecasts. The 3rd column of the table reports estimates for a model which replaces the OSV informational assumption with the assumption used by Maheu and McCurdy (2007), that is, relating the conditional variance of daily returns to the conditional expectation of daily RV, as in equation (3.1). In this case, t-distributed return innovations, as in equation (3.5), dominate Normal return innovations.

Based on the in-sample loglikelihood, the 2Comp-OSV specification dominates the 2Comp specification. However, as shown in Figure 4, there is not a large difference with respect to out-of-sample density forecasts. This is also evident from comparing the parameter estimates in Table 2. Except for the return intercept, and the fact that the return innovations have fatter tails for the 2Comp model than for the 2Comp-OSV version, the parameter estimates are similar.

The main features of our results are as follows. Bivariate models that use high-frequency intraday data provide a significant improvement in density forecasts relative to an EGARCH

model estimated from daily data. Two-component specifications for the dynamics of $\log(RV)$ provide similar or better performance than HAR alternatives; both dominate the less flexible single-component version. A bivariate model of returns with Normal innovations and observable stochastic volatility directed by a 2-component, exponentially decaying function of $\log(RV)$ provides good density forecasts over a range of out-of-sample horizons for both data series. We find that adding a mixture of Normals or GARCH effects to the innovations of the $\log(RV)$ part of this specification is not statistically important for S&P 500, while the addition of the mixture of Normals provides a significant improvement for IBM.

6   Conclusion

This paper proposes alternative joint specifications of daily returns and RV which link RV to the variance of returns and exploit the benefits of using intraperiod information to obtain accurate measures of volatility. Our focus is on out-of-sample forecasts of the return distribution generated by our bivariate models of return and RV. We explore which features of the time-series models contribute to superior density forecasts over horizons of 1 to 60 days out of sample.

Our main method of model comparison uses the predictive likelihood of returns, the forecast density evaluated at the realized return, which provides a measure of the likelihood of the data being consistent with the model. An identical set of return observations is used to compute a term structure of test statistics over a range of forecast horizons, so that the average predictive likelihoods are not only comparable across models but also over different forecast horizons for a particular model.

Two alternative joint specifications of daily returns and realized volatility were investigated. These two bivariate models are distinguished by alternative assumptions about RV dynamics. The first model uses a heterogenous autoregressive (HAR) specification of $\log(RV)$. The second model allows components of $\log(RV)$ to have different decay rates. Both of these bivariate models allow for asymmetric effects of past negative versus positive return innovations. Both models are stationary and consistent with mean reversion in RV. We also investigate an observable SV assumption (OSV) for the timing of information availability.

Using the predictive likelihood, we find that high-frequency intraday data is important for density forecasts relative to using daily data as in our benchmark EGARCH specfication. Secondly, a flexible function form (either two components or HAR) is very important for the dynamics of $\log(RV)$. The OSV assumption marginally improves density forecasts at long horizons for the S&P 500 but is essentially similar for the IBM data. A bivariate model of returns with Normal innovations and observable stochastic volatility directed by a 2-component, exponentially decaying function of $\log(RV)$ provides good density forecasts over a range of out-of-sample horizons for both data series.

Table 1: Summary Statistics: Daily Returns and Realized Volatility

	Mean	Variance	Skewness	Kurtosis	Min	Max
SPY						
r_t	-0.018	0.967	0.080	6.180	-7.504	8.236
RV_u	1.210	2.640	6.932	84.936	0.055	33.217
RV_{AC1b}	1.079	2.373	7.670	96.439	0.047	30.789
RV_{AC2b}	1.013	2.115	7.530	88.588	0.043	25.227
RV_{AC3b}	0.978	2.054	8.071	102.635	0.036	26.329
IBM						
r_t	-0.037	2.602	0.074	3.898	-11.699	11.310
RV_u	2.825	9.161	5.145	54.879	0.150	58.270
RV_{AC1b}	2.623	9.433	6.051	75.409	0.132	65.069
RV_{AC2b}	2.558	9.875	6.377	82.091	0.114	66.594
RV_{AC3b}	2.531	10.095	6.362	81.024	0.010	65.235

r_t are daily returns constructed from open and close prices. RV_u is constructed from raw 5-minute returns with no adjustment. $RV_{ACqb}, q = 1, 2, 3$ are computed as:

$$RV_{ACqb} \;=\; \hat{\gamma}_0 + 2\sum_{j=1}^{q} \omega_j \hat{\gamma}_j, \quad \hat{\gamma}_j = \sum_{i=1}^{M-j} r_{t,i} r_{t,i+j}, \quad \omega_j = 1 - \frac{j}{q+1}, \; j = 0, 1, ..., q$$
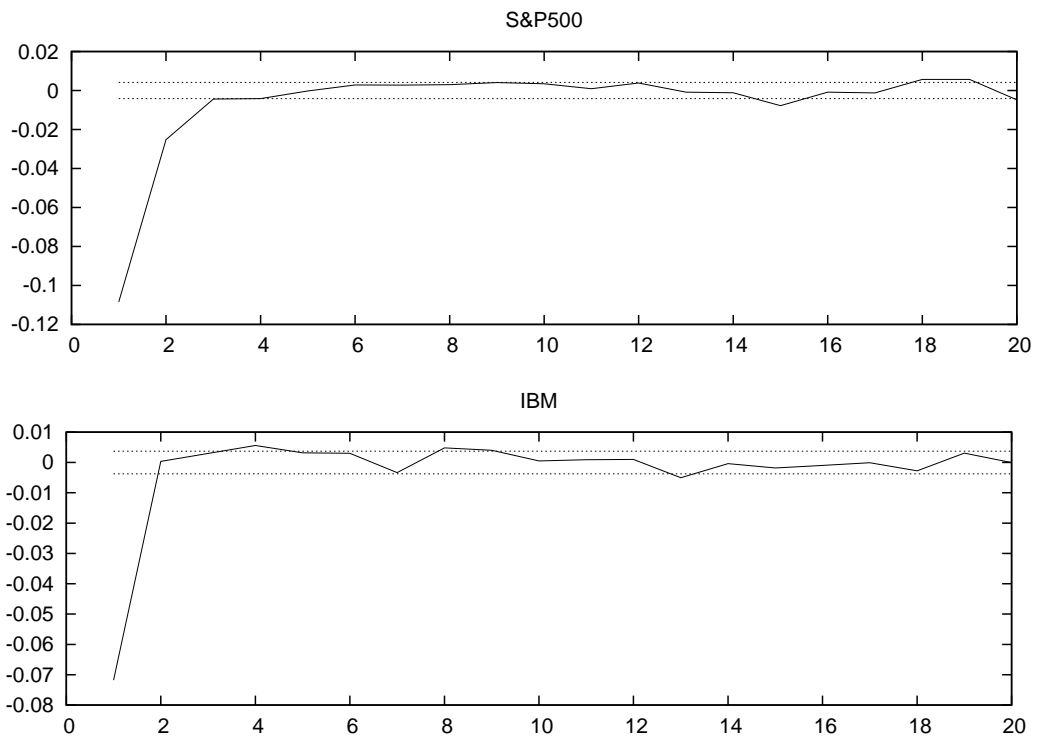
Figure 1: ACF of 5-Minute Return Data
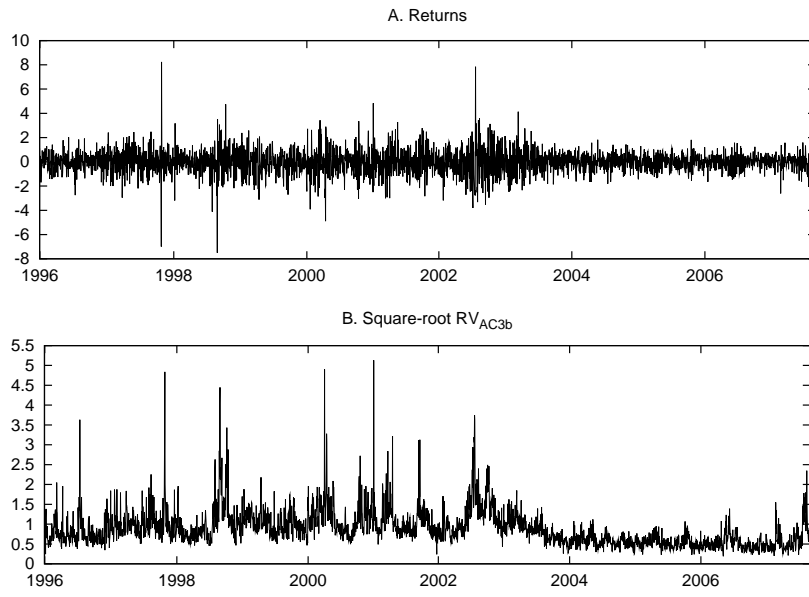
Figure 2: S&P 500 (Spyder) Daily Time Series

A. Returns

B. Square-root RV_{AC3b}

Figure 3: IBM Daily Time Series
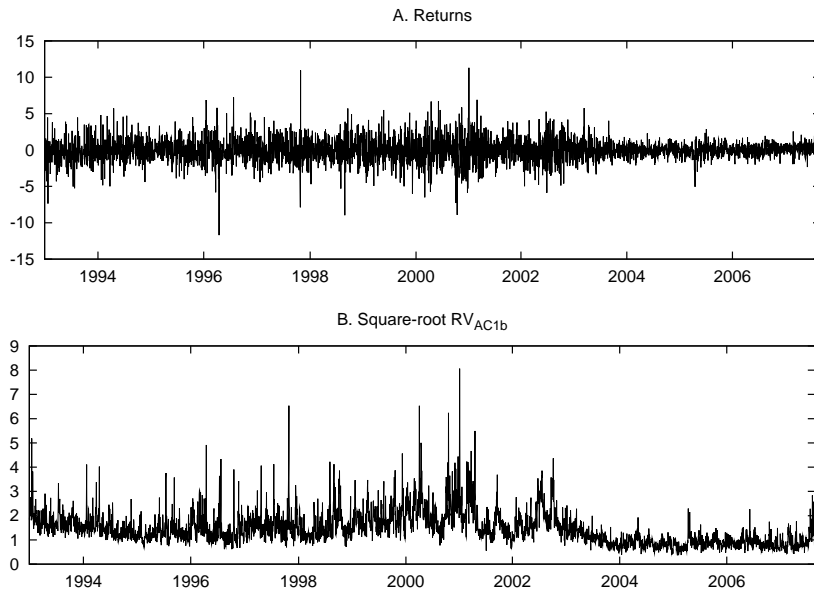
A. Returns

B. Square-root RV_{AC1b}

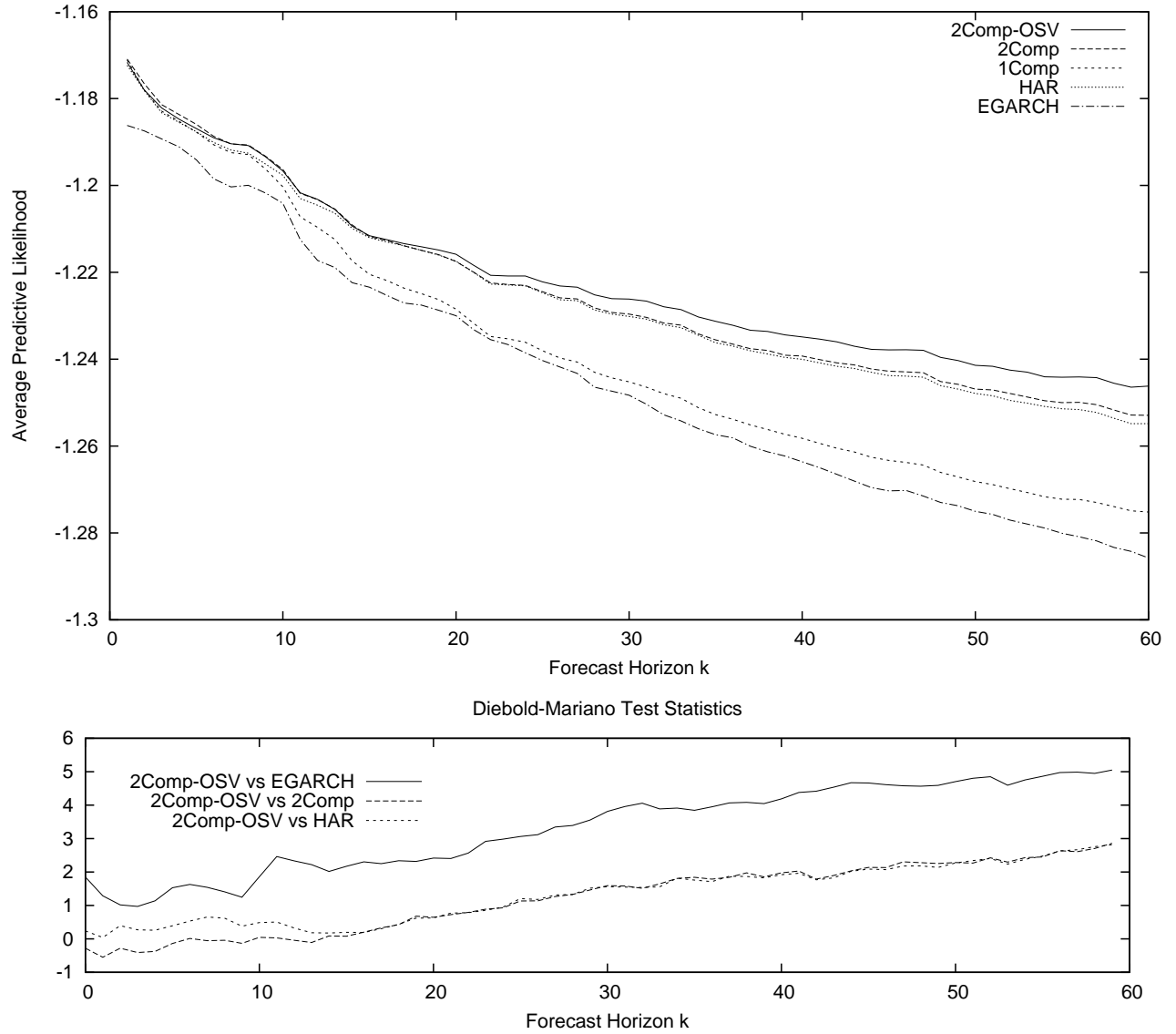Figure 4: S&P 500, Joint Models versus EGARCH

23

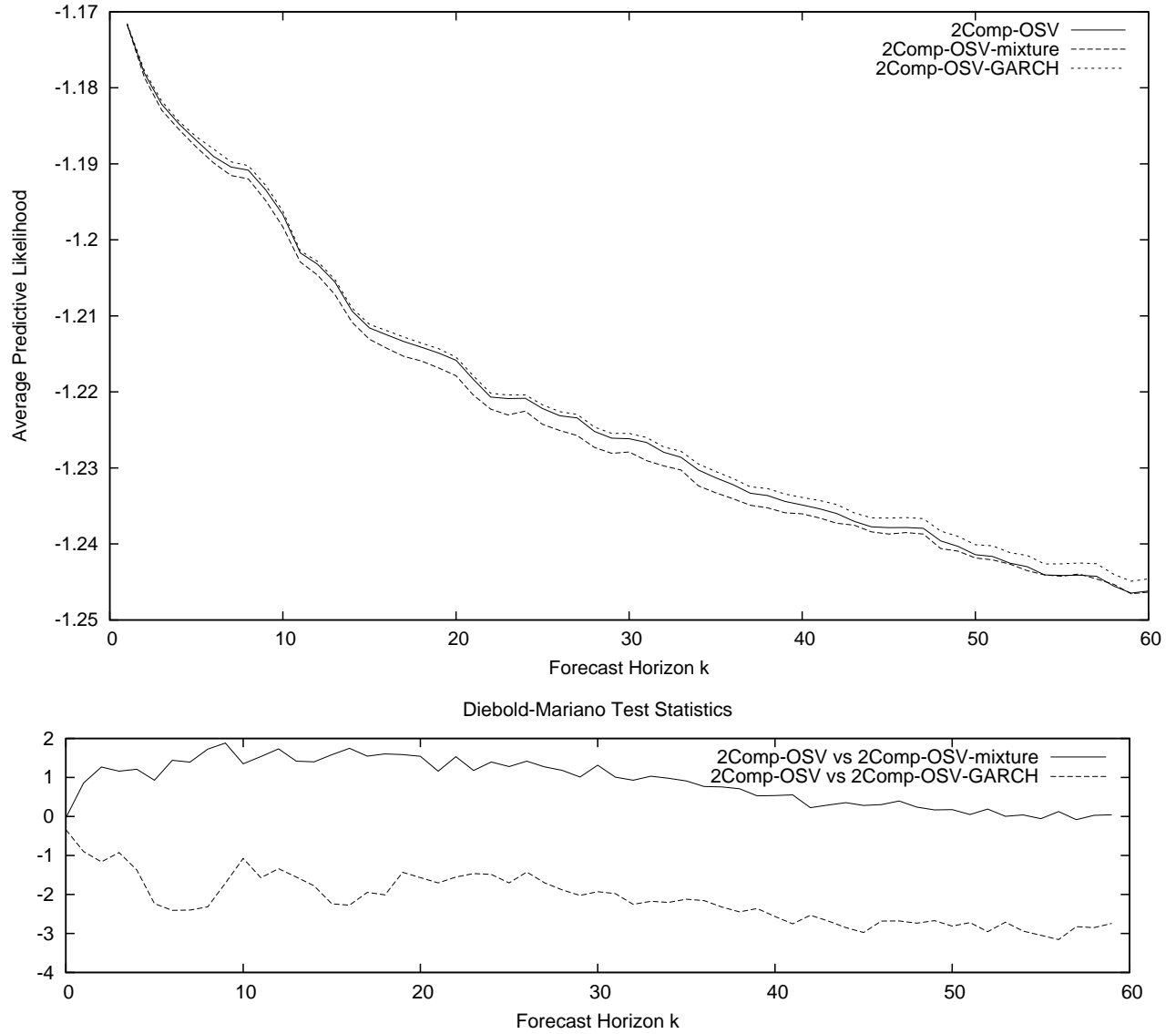Figure 5: S&P 500, Robustness to non-Normal Innovations to $\log(RV)$
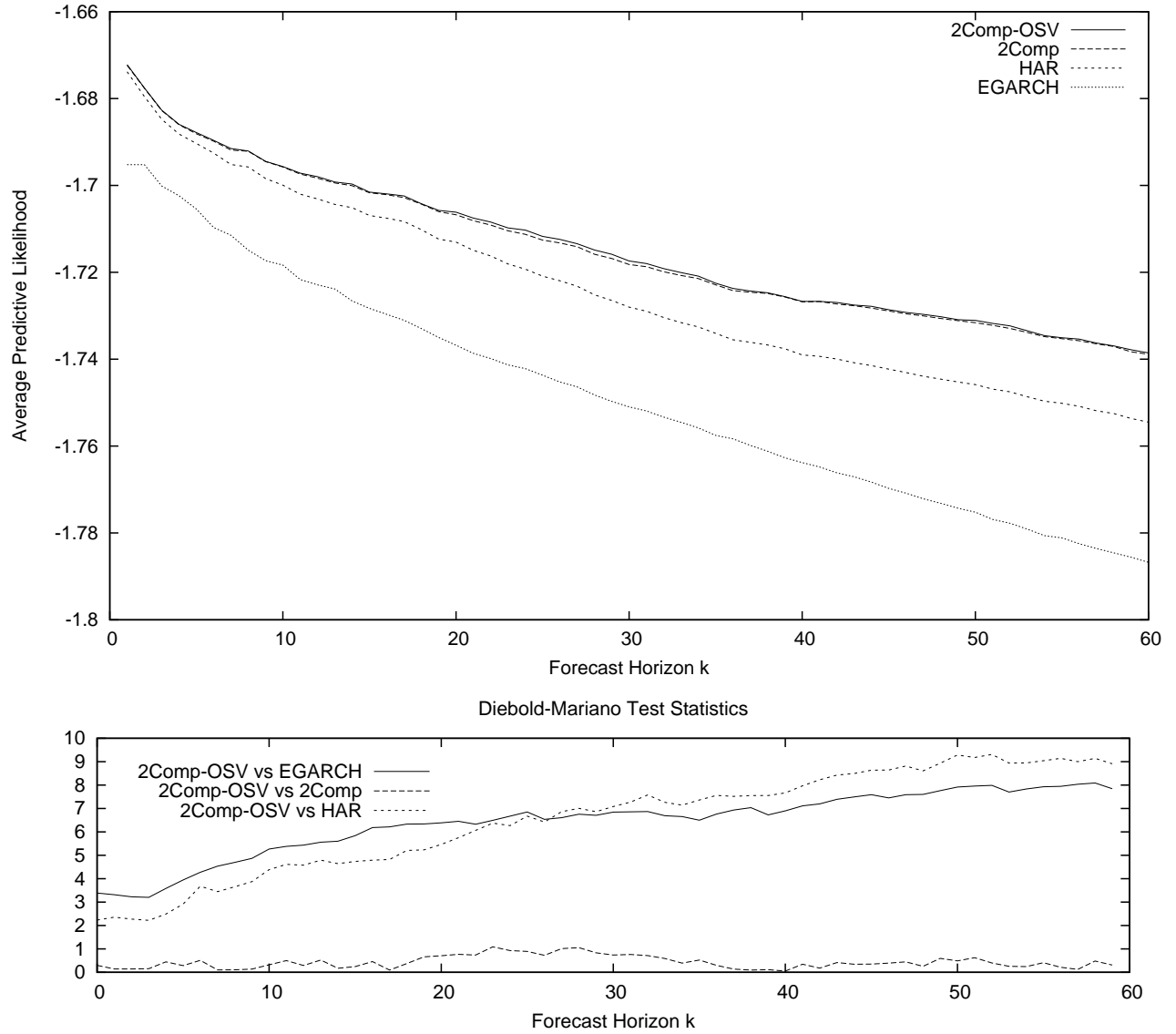
Figure 6: IBM, Joint Models versus EGARCH

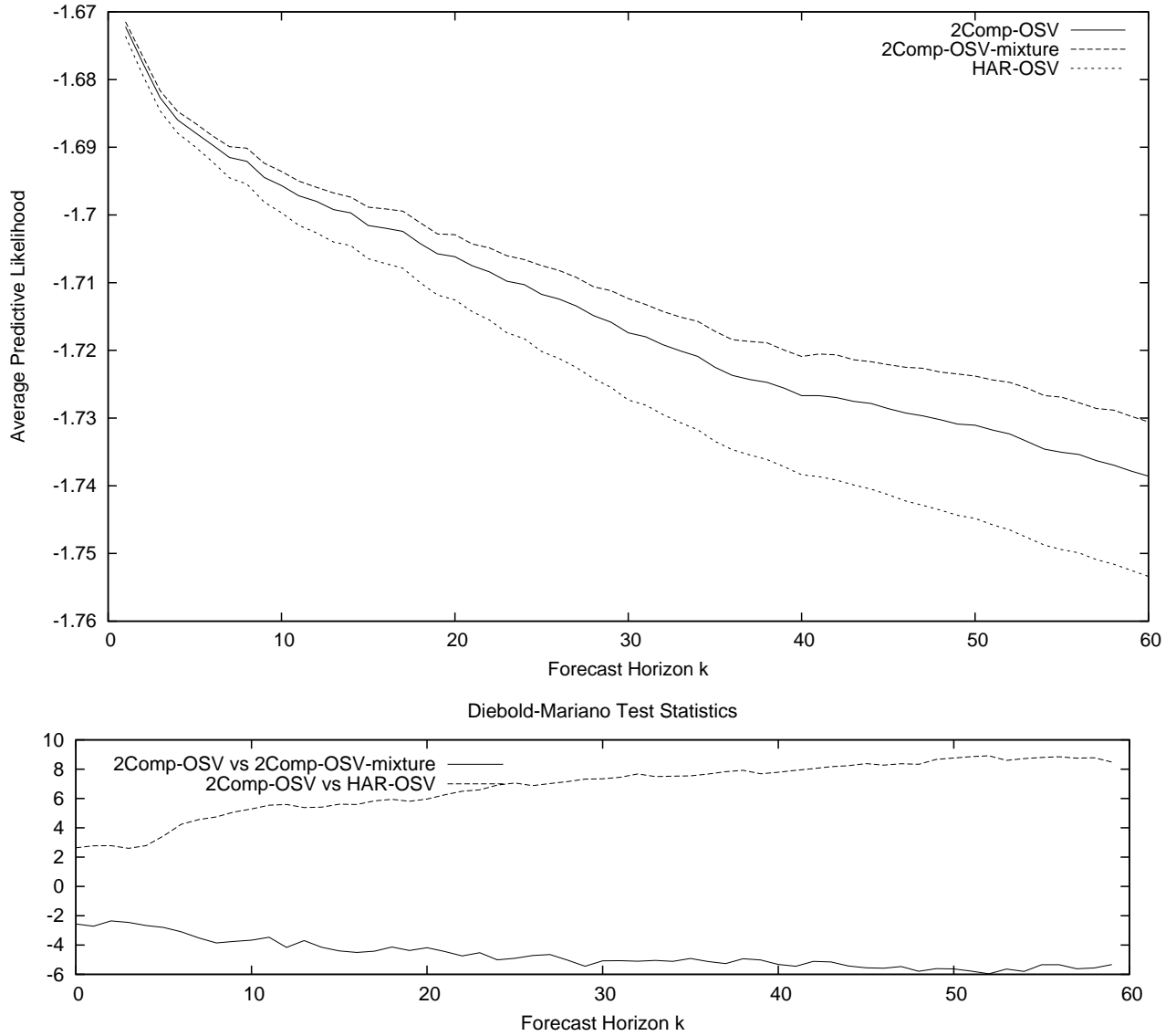Figure 7: IBM, Robustness to non-Normal Innovations to $\log(RV)$

Table 2: S&P 500 Model Estimates

2Comp-OSV Model

$$r_t = \mu + \epsilon_t, \;\; \epsilon_t = \sqrt{RV_t}u_t, \;\; u_t \sim D(0,1)$$

$$\log(RV_t) = \omega + \sum_{i=1}^{2} \phi_i s_{i,t} + \gamma u_{t-1} + \eta v_t, \;\; v_t \sim NID(0,1),$$

$$s_{i,t} = (1-\alpha_i)\log(RV_{t-1}) + \alpha_1 s_{i,t-1}, \;\; i=1,2.$$

2Comp Model

$$r_t = \mu + \epsilon_t, \;\; \epsilon_t = \sigma_t u_t, \;\; u_t \sim t_v(0,1)$$

$$\sigma_t^2 = \exp\left(E_{t-1}\log(RV_t) + \frac{1}{2}Var_{t-1}(\log(RV_t)) \right)$$

$$\log(RV_t) = \omega + \sum_{i=1}^{2} \phi_i s_{i,t} + \gamma u_{t-1} + \eta v_t, \;\; v_t \sim NID(0,1),$$

$$s_{i,t} = (1-\alpha_i)\log(RV_{t-1}) + \alpha_1 s_{i,t-1}, \;\; i=1,2.$$

Parameter	$u_t \sim N(0,1)$ 2Comp-OSV	$u_t \sim t_\nu(0,1)$ 2Comp
μ	0.038	-0.018
	(0.011)	(0.014)
ω	-0.026	-0.025
	(0.012)	(0.013)
ϕ_1	0.476	0.402
	(0.007)	(0.147)
ϕ_2	0.476	0.543
	()	(0.154)
α_1	0.888	0.911
	(0.017)	(0.045)
α_2	0.435	0.508
	(0.037)	(0.105)
γ	-0.129	-0.141
	(0.010)	(0.011)
η	0.531	0.528
	(0.009)	(0.009)
$1/\nu$		0.089
		(0.016)
lgl	-5646.725	-5916.342

References

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

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2006): "Ultra High-Frequency Volatility Estimation with Dependent Microstructure Noise," Manuscript, Princeton University, University of Chicago and University of Illinois.

Amisano, G., and R. Giacomini (2007): "Comparing Density Forecasts via Weighted Likelihood Ratio Tests," *Journal of Business and Economic Statistics*, 25(2), 177–190.

Andersen, T., T. Bollerslev, and F. X. Diebold (2007): "Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility," *Review of Economics and Statistics*, 89, 701–720.

Andersen, T., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 529–626.

Andersen, T. G., and T. Bollerslev (1998): "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39(4), 885–905.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2004): "Parametric and Nonparametric Volatility Measurement," in *Handbook of Financial Econometrics*, ed. by L. Hansen, and Y. Ait-Sahalia. North Holland.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001): "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43–76.

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

Andersen, T. G., T. Bollerslev, and N. Meddahi (2004): "Analytic Evaluation of Volatility Forecasts," *International Economic Review*, 45, 1079–1110.

Andreou, E., and E. Ghysels (2002): "Rolling-sample Volatility Estimators: Some New Theoretical, Simulation and Empirical Results," *Journal of Business & Economic Statistics*, 20, 363–376.

Bandi, F. M., and J. R. Russell (2005): "Microstructure Noise, realized volatility, and optimal sampling," Manuscript, GSB, University of Chicago.

——— (2006): "Separating microstructure noise from volatility," *Journal of Financial Economics*, 79, 655–692.

Bao, Y., T.-H. Lee, and B. Saltoglu (2006): "Comparing Density Forecast Models," forthcoming *Journal of Forecasting*.

BARNDORFF-NIELSEN, O., P. HANSEN, A. LUNDE, AND N. SHEPHARD (2004): "Regular and modified kernel-based estimators of integrated variance: The case with indpendent noise," Working Paper, http://ssrn.com/abstract=620203.

BARNDORFF-NIELSEN, O. E., P. R. HANSEN, A. LUNDE, AND N. SHEPHARD (2006): "Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise," Working paper Department of Economics, Stanford University.

BARNDORFF-NIELSEN, O. E., AND N. SHEPHARD (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.

——— (2004): "Power and Bipower Variation with Stochastic Volatility and jumps," *Journal of Financial Econometrics*, 2, 1–48.

——— (2007): "Variation, Jumps, Market Frictions and High-Frequency Data in Financial Econometrics," in *Advances in Economics and Econometrics. Theory and Applications, Ninth World Congress*, ed. by R. Blundell, T. Persson, and W. Newey. Cambridge University Press.

BOLLERSLEV, T., U. KRETSCHMER, C. PIGORSCH, AND G. TAUCHEN (2007): "A Discrete-Time Model for Daily S&P500 Returns and Realized Variations: Jumps and Leverage Effects," forthcoming Journal of Econometrics.

CORSI, F. (2003): "A Simple Long Memory Model of Realized Volatility," manuscript, University of Southern Switzerland.

CURCI, G., AND F. CORSI (2006): "Discrete sine transform for multi-scales realized volatility measures," NCCR FINRISK Working Paper No. 44.

DIEBOLD, F. X., AND R. S. MARIANO (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13(3), 252–263.

GHYSELS, E., P. SANTA-CLARA, AND R. VALKANOV (2005): "There is a Risk-Return Tradeoff After All," *Journal of Financial Economics*, 76, 509–548.

——— (2006): "Predicting Volatility: Getting the Most Out of Return Data Sampled at Different Frequencies," *Journal of Econometrics*, 131, 445–475.

GHYSELS, E., AND A. SINKO (2006): "Volatility Forecasting and Microstructure Noise," Manuscript, Department of Economics, University of North Carolina.

GIOT, P., AND S. LAURENT (2004): "Modelling daily Value-at-Risk using realized volatility and ARCH models," *Journal of Empirical Finance*, 11, 379–398.

HANSEN, P. R., AND A. LUNDE (2006): "Realized Variance and Market Microstructure Noise," *Journal of Business & Economic Statistics*, 24(2), 127–161.

KOOPMAN, S. J., B. JUNGBACKER, AND E. HOL (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.

Liu, C., and J. M. Maheu (2008): "Are There Structural Breaks in Realized Volatility?," *Journal of Financial Econometrics*, 6(3), 326–360.

Maheu, J. M., and T. H. McCurdy (2002): "Nonlinear Features of FX Realzied Volatility," *Review of Economics and Statistics*, 84(4), 668–681.

Maheu, J. M., and T. H. McCurdy (2007): "Components of Market Risk and Return," *Journal of Financial Econometrics*, 5(4), 560–590.

Martens, M., D. van Dijk, and M. de Pooter (2003): "Modeling and Forecasting S&P500 Volatility: Long Memory, Structural Breaks and Nonlinearity," Econometric Institute, Erasmus University Rotterdam.

Meddahi, N. (2002): "A Theoretical Comparison between Integrated and Realized Volatility," *Journal of Applied Econometrics*, 17, 479–508.

——— (2003): "ARMA Representation of Integrated and Realized Variances," *The Econometrics Journal*, 6, 334–355.

Nelson, D. B. (1990): "ARCH Models as Diffusion Approximations," *Journal of Econometrics*, 45, 7–39.

——— (1991): "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–370.

Oomen, R. C. A. (2005): "Properties of Bias-Corrected Realized Variance under Alternative Sampling Schemes," *Journal of Financial Econometrics*, 3, 555–577.

Robert, C. P., and G. Casella (1999): *Monte Carlo Statistical Methods*. Springer, New York.

Taylor, S., and X. Xu (1997): "The Incremental Information in One Million Foreign Exchange Quotations," *Journal of Empirical Finance*, 4, 317–340.

Weigend, A. S., and S. Shi (2000): "Predicting Daily Probability Distributions of S&P500 Returns," *Journal of Forecasting*, 19, 375–392.

Zhang, L. (2004): "Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach," Discussion Paper, Carnegie-Mellon University.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005): "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100(472), 1394–1411.

Zhou, B. (1996): "High-Frequency Data and Volatility in Foreign Exchange Rates," *Journal of Business & Economic Statistics*, 14, 45–52.